

Klotz, Stefan

Working Paper

Ökonometrische Modelle mit raumstruktureller Autokorrelation: Eine kurze Einführung

Diskussionsbeiträge - Serie II, No. 328

Provided in Cooperation with:

Department of Economics, University of Konstanz

Suggested Citation: Klotz, Stefan (1997) : Ökonometrische Modelle mit raumstruktureller Autokorrelation: Eine kurze Einführung, Diskussionsbeiträge - Serie II, No. 328, Universität Konstanz, Sonderforschungsbereich 178 - Internationalisierung der Wirtschaft, Konstanz

This Version is available at:

<https://hdl.handle.net/10419/101595>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

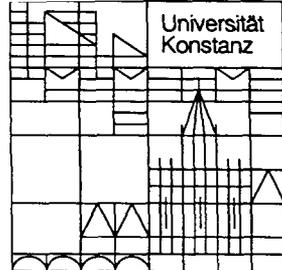
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Sonderforschungsbereich 178
„Internationalisierung der Wirtschaft“

Diskussionsbeiträge

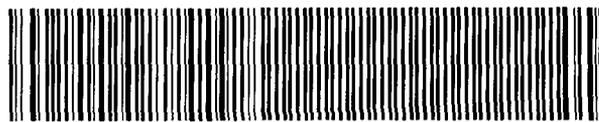
Juristische
Fakultät

Fakultät für Wirtschafts-
wissenschaften und Statistik

Stefan Klotz

Ökonometrische Modelle
mit raumstruktureller Autokorrelation:
Eine kurze Einführung

W 113 (328)



24. FEB. 1997 Weltwirtschaft
Kiel

W 113 C 328)

mi se sig 5a

Postfach 5560
D-78434 Konstanz

Serie II — Nr. 328
Januar 1997

**Ökonometrische Modelle mit
raumstruktureller Autokorrelation:
Eine kurze Einführung**

Stefan Klotz

739151

Serie II - Nr. 328

Januar 1997

ÖKONOMETRISCHE MODELLE MIT RAUMSTRUKTURELLER AUTOKORRELATION: EINE KURZE EINFÜHRUNG

STEFAN KLOTZ¹

Universität Konstanz
Postfach 5560 D124
D-78434 Konstanz

klotz@enterprise.wiwi.uni-konstanz.de

Zusammenfassung:

Querschnittsdaten aus benachbarten Raumgebieten, wie in der Regionalökonomie verwendet, weisen neben Heteroskedastie oft auch gegenseitige Abhängigkeiten auf. Die wechselseitige Beeinflussung wird in der raumstrukturellen Ökonometrie (Spatial Econometrics) meist explizit durch Autokorrelation entweder im Fehlerterm oder in der endogenen Variablen modelliert. Für die Schätzung entsprechender Prozesse ist KQ nicht anwendbar, während Maximum Likelihood-Schätzer rechentechnische Probleme verursachen, so daß GMM-Verfahren vorzuziehen sind. Vorliegender Beitrag gibt einen Überblick über gängige Test- und Schätzverfahren und verdeutlicht deren Eigenschaften in endlichen Stichproben mit Hilfe von Monte Carlo-Studien.

Abstract:

Cross section data originating from neighboring-spatial units - frequently used in Regional Sciences - is not only affected by heteroscedasticity but often also mutually dependent. In Spatial Econometrics, this is often taken into account by explicitly modelling autocorrelated error terms or dependent variables. OLS is not suitable for the estimation of such processes whereas ML may cause computational burden so the GMM methodology is favorable. This paper gives a survey of several tests and estimation procedures used for spatial data and examines their small sample properties presenting some Monte Carlo studies.

¹ Für Unterstützung, wertvolle Kritik und Anregungen zu dieser Arbeit danke ich Winfried Pohlmeier, Thies Büttner, Frank Gerhard und Joachim Inkmann. Ferner geht mein Dank für finanzielle Unterstützung an die Deutsche Forschungsgemeinschaft. Frühere Versionen wurden in Workshops der Universität Konstanz sowie des Zentrums für Europäische Wirtschaftsforschung (ZEW, Mannheim) vorgestellt. Alle verbliebenen Fehler sind ausschließlich mir anzulasten.

Gliederung

1. Einführung in Thema und Aufbau der Arbeit	2
2. Grundlegende Begriffe und Eigenschaften der Raumstrukturen	4
3. Modelle mit raumstruktureller Autokorrelation	7
4. Eigenschaften einiger Schätzverfahren bei raumstrukturellen Abhängigkeiten	12
5. Tests auf Anwesenheit raumstruktureller Abhängigkeiten	17
6. Schätzverfahren nach der Verallg. Momentenmethode (GMM)	19
7. Zusammenfassung und Ausblick	23
Anhang	24
Literaturverzeichnis	29

1. Einführung

Bei der Betrachtung von Querschnittsdaten ist es üblich, etwaige Heteroskedastie zu berücksichtigen. Wesentlich seltener wird die Möglichkeit in Betracht gezogen, daß die einzelnen Beobachtungen voneinander abhängig sein könnten. Handelt es sich bei den Beobachtungseinheiten um räumlich benachbarte Gebietseinheiten, dürfte die Annahme der Unabhängigkeit in der Regel unzutreffend sein, da die einzelnen Gebiete z.B. durch willkürliche Grenzziehungen oder einfach durch Mobilität der Wirtschaftssubjekte und -objekte interagieren. Daraus ist die Notwendigkeit abzuleiten, den Raum als explizite Größe in ökonomische und ökonometrische Überlegungen einzubeziehen, wie dies mit der Größe "Zeit" längst üblich ist.

Nachdem im ökonomischen Kontext nicht der Raum an sich, sondern die in ihm befindlichen Strukturen (politischer, ökonomischer, sozialer etc. Art) von Interesse sind, soll in bewußter Anlehnung an den Begriff der Zeitreihe von *Raumstrukturen* die Rede sein, wenn die räumliche Anordnung von Wirtschaftssubjekten bzw. -objekten zu einer Erklärung ökonomischer Tatbestände beitragen kann. Die vorliegende Arbeit zeigt auf, daß gerade bei ökonomischen bzw. ökonometrischen Fragestellungen Raumstrukturen eine wichtige Rolle spielen können.

Trotz seines Schattendaseins hat dieser als *Spatial Econometrics* firmierende Zweig der Ökonometrie inzwischen eine große Anzahl von Schätzmethoden und Testverfah-

ren entwickelt. In dieser Arbeit wird ein Überblick über das bedeutendste Teilgebiet gegeben, die Autokorrelation im Raum.

Eine sicherlich wünschenswerte scharfe Eingrenzung der Ökonometrie der Raumstrukturen ist nicht eindeutig möglich. Zum einen können raumstrukturelle Zusammenhänge in etlichen Fällen auch mit herkömmlicher Methodologie erfaßt werden, zum anderen sind raumstrukturelle Techniken bisweilen in Problemstellungen hilfreich, bei denen der Raum keine entscheidende Rolle spielt. In Anlehnung an die inzwischen allgemein anerkannte Definition² für *Spatial Econometrics*, wird in dieser Arbeit

Raumstrukturelle Ökonometrie folgendermaßen definiert:

Alle Methoden und Techniken, welche für raumstrukturelle Problemstellungen angemessene Modellspezifikationen, Schätzverfahren, Hypothesentests oder auch Prognosen herbeiführen können, gehören zum Gebiet der Raumstrukturellen Ökonometrie.

Eine Problemstellung habe dabei raumstrukturellen Charakter, wenn für Phänomene einer bestimmten Region erklärende Faktoren in anderen Regionen (z.B. Städte, Bundesländer, ...) zu finden sind, so daß Interdependenzen über bzw. durch räumliche Entfernungen bestehen. Die explizite Modellierung des Raumes ist dadurch regelmäßig nötig.

Darüber hinaus können Methoden der raumstrukturellen Ökonometrie auch Anwendung auf nichtgeographische, im Abstrakten definierte Räume finden. Im folgenden soll deshalb der Begriff „Regionen“ auch Beobachtungen miteinschließen, die wie räumliche Einheiten aufeinander einwirken. Ein Beispiel hierfür sind Unternehmen, deren Innovationsaktivitäten und -erfolge abhängig von der Ähnlichkeit ihrer Produkte oder Produktionsprozesse, darstellbar als Nachbarschaften im Produktraum, aufeinander einwirken.

Andere, durchaus zentrale Begriffe in diesem Gebiet sind nicht eindeutig geklärt. Nachdem Begriffsprägungen nur für die englische Sprache vorliegen, werden in dieser Arbeit eigenständige deutsche Begriffe formuliert. Die in der Literatur vorherrschenden englischen Bezeichnungen werden zusätzlich genannt.

Die Arbeit ist wie folgt aufgebaut:

Abschnitt 2 faßt begriffliche Grundlagen der Raumstrukturen zusammen. Dabei zeigt sich, daß eine Darstellung durch abgewandelte und erweiterte Konzepte aus der Zeitreihenökonomie weitgehend möglich ist. Es erfolgt eine Unterteilung in die beiden Problemfelder „Konditionale Heterogenität“ und „Wechselseitige Abhängigkeit“. Da letztere im Zentrum dieser Arbeit steht, werden Aspekte der Heterogenität nur kurz skizziert.

² Dabei handelt es sich um folgende Definition in Anselin (1988), S.10, auf welche sich seitdem die meisten Publikationen beziehen: "(...) I will consider the field of spatial econometrics to consist of those methods and techniques that, based on a formal representation of the structure of spatial dependence and spatial heterogeneity, provide the means to carry out the proper specification, estimation, hypothesis testing, and prediction for models in regional science." Problematisch ist die fehlende Erläuterung des Begriffes "regional science" bei Anselin und ihn zitierenden Autoren.

Raumstrukturelle Abhängigkeit kann durch verschiedene Spezifikationen dargestellt werden, die sich in Begriff und Form ebenfalls an Zeitreihenprozesse anlehnen. Neben diesen werden in Abschnitt 3 die Schwierigkeiten diskutiert, welche bei der Konstruktion raumstruktureller Gewichtungsmatrizen auftreten.

Verschiedene Schätzmethoden für raumstrukturelle Prozesse werden in Abschnitt 4 erörtert. Dabei wird gezeigt, daß kein Standardinstrumentarium der ökonometrischen Schätztheorie bei Raumstrukturen zufriedenstellende Eigenschaften aufweist. Ein Beispiel für geeignete Verfahren, auf raumstrukturelle Abhängigkeiten zu testen, skizziert Abschnitt 5.

In Abschnitt 6 wird die Verallgemeinerte Momentenmethode als Schätzalternative aufgeführt. An einem Vorschlag von KELEJIAN/ROBINSON (1993) werden der Ansatz und die Vorteile des Verfahrens verdeutlicht.

Eine kurze Zusammenfassung der wichtigsten Ergebnisse und ein Ausblick auf weitere Möglichkeiten der raumstrukturellen Ökonometrie schließt in Abschnitt 7 die Arbeit ab.

2. Grundlegende Begriffe und Eigenschaften der Raumstrukturen

2.1 Definition zentraler Begriffe Raumstruktureller Prozesse

Es erweist sich aufgrund der Verwandtschaft der beiden Arten von Dimension immer wieder als nützlich, für raumstrukturelle Problemstellungen auf die zahlreichen Konzepte der Zeitreihenökonomie zurückzugreifen. So lassen sich zentrale Begriffe ganz ähnlich definieren. Allerdings ist es nicht damit getan, den Zeitindex t durch einen Raumindex r zu ersetzen, weil der Raum komplexere Strukturen bereithält. Erstens zwingt der Raum zur Betrachtung einer zweiten Dimension.³ Zweitens existiert im Raum keine eindeutig vorgegebene Betrachtungsrichtung, wie sie der Zeitablauf vorgibt. Daraus folgt, daß der Abstand zweier Punkte oder gar zweier Flächen im Raum nicht so einfach zu messen ist wie entlang einer Zeitachse.

Zunächst soll zur Vereinfachung der Formulierung für diese Arbeit festgelegt werden, daß unter dem Begriff *Region* eine räumliche Einheit verstanden werden soll. Die einzelnen Regionen überschneiden sich dabei nicht und teilen das betrachtete Gebiet vollständig untereinander auf.

Ein elementares Konzept der Zeitreihenbetrachtung besteht darin, daß die Perioden zur Untersuchung der Abhängigkeit von den Vergangenheitswerten in systematischen Abständen geordnet werden. Dazu bedient man sich meist der sogenannten *Verzögerungs-* oder *Lagoperatoren*:

$$L^k y_t = y_{t-k} \quad (2 - 1)$$

³ Zumindest bei der Betrachtung nichtgeographischer Räume können ohne weiteres sogar mehr als zwei Dimensionen auftreten.

L^k verschiebt also von der Beobachtung zum Zeitpunkt t auf diejenige von k Perioden zuvor. Ein vergleichbares Konzept von *Räumlichen Verzögerungen* (*Spatial Lags*) ist wünschenswert, jedoch wegen der obigen beiden Gründe nicht problemlos übertragbar. Unterstellen wir jedoch Regionen identischer Größe, die in einem regelmäßigen Muster angeordnet sind, wie z.B. in einer Wabenstruktur, können Nachbarschaftsgrade k eingeführt werden. Ein solcher liege vor, wenn der kürzeste denkbare Pfad von einer Region zur anderen über $(k - 1)$ weitere Regionen führt. Auch wenn das Problem der nicht festgelegten Betrachtungsrichtung bestehen bleibt, sind hierdurch die meisten Konzepte und Begriffe der Zeitreihenbetrachtung auf den raumstrukturellen Kontext übertragbar.

Wir gehen aus von einem in R Regionen gewonnenen $R \times 1$ -Beobachtungsvektor y , dessen Elemente grundsätzlich auch Reihenvektoren mehrerer Merkmale sein dürfen. Dann nennen wir den Prozeß für y_r *stationär*,⁴ wenn folgende Bedingungen erfüllt sind⁵:

$$\begin{aligned} E(y_r) &= \mu_r = \mu \quad , \quad V(y_r) = \sigma_r^2 = \sigma^2 \quad , \\ E[(y_r - \mu)(y_s - \mu)] &= \sigma^2 \rho_k \quad \text{mit } r, s = 1, \dots, R; \end{aligned} \quad (2 - 2)$$

Der Index r sowie der Hilfsindex s zählen dabei (wie in der gesamten folgenden Arbeit) die Regionen von 1 bis R ; k gibt die räumliche Verzögerung zwischen Region r und s wieder und bestimmt den Wert ρ_k der Korrelation zwischen ihnen. In dieser Spezifikation ist der Sinn einer Analogie zur Zeitreihenbetrachtung offensichtlich. Komplizierter wird die Beschreibung des Prozesses, wenn man die unregelmäßigen Muster im Raum besser berücksichtigen möchte und statt der diskret gemessenen räumlichen Verzögerungen die Korrelation von einem stetigen Distanzmaß abhängig macht:

$$\begin{aligned} E[(y_r - \mu)(y_s - \mu)] &= \sigma^2 \rho_{r,s} \quad , \\ \text{mit } \rho_{r,s} &= \rho[d(r,s)] \quad , \quad d > 0 \quad , \quad \text{und } \frac{\partial \rho}{\partial d} < 0 \quad . \end{aligned} \quad (2 - 3)$$

Die Korrelation zweier Regionen nimmt also mit der Entfernung d zwischen ihnen ab. Die zutreffende Form der Distanzfunktion $d(r,s)$ muß für den betreffenden Einzelfall bestimmt werden.⁶ Daneben muß auch Klarheit über die von der Distanz abhängige, monoton fallende Korrelationsfunktion $\rho(d)$ gewonnen werden, die meist *Entfernungsbewertungs-Funktion* (*Distance-Decay-Function*) genannt wird. Dann muß gelten, daß für zwei Punktepaare, zwischen denen derselbe Abstand gemessen wird, auch eine identische Kovarianz herrschen muß, damit es sich um einen stationären Prozeß handeln kann. Aus der unvermeidlichen Wahl der Distanzfunktion resultiert eine gewisse Beliebigkeit in der Frage der Stationarität.

⁴ Stationarität wird hier also als Kovarianzstationarität definiert. Ein engerer Stationaritätsbegriff im Sinne einer gemeinsamen Verteilung erscheint im raumstrukturellen Kontext nicht sinnvoll, vgl. KLOTZ(1996).

⁵ In der gängigen Literatur finden sich unterschiedliche Definitionen, da jeder Autor einen eigenen Spezialfall heranzieht. Vgl. z.B. CLIFF/ORD (1973), S: 143, CLIFF ET AL. (1975), S. 50, ANSELIN (1988), S.42f.

⁶ Das Problem, eine geeignete Distanzfunktion zu finden, ist eng verwandt mit der Suche nach einer geeigneten Gewichtungsmatrix. Diese Frage wird in Abschnitt 3.3 ausführlich diskutiert.

Die bisher verwendete allgemeine Form der Distanzfunktion schließt die Möglichkeit der Asymmetrie ein: Der Einfluß von Region r auf Region s kann ein anderer sein als der von s auf r . Symmetrie hingegen impliziert, daß (bei Flächenbetrachtungen) ein eindimensionales Distanzmaß verwendet werden darf, ohne daß dabei Information verloren ginge. Ein symmetrischer, also richtungsunabhängiger, und außerdem stationärer Prozeß heißt *isotrop*. Der ausführlicheren Diskussion in 3.3 vorgreifend sei bereits hier bemerkt, daß Symmetrie keineswegs selbstverständlich ist, da z.B. ein kleines Land von einem großen stärker beeinflusst wird als umgekehrt.

2.2 Raumstrukturelle Heterogenität und raumstrukturelle Abhängigkeit

Die Effekte, welche das Einbeziehen der räumlichen Dimensionen in eine Modellierung auslösen kann, werden, zurückgehend auf Anselin⁷, durch Abgrenzung der Aspekte *raumstruktureller Heterogenität (Spatial Heterogeneity)* einerseits und *raumstruktureller Abhängigkeit (Spatial Dependence)* andererseits systematisiert.

Von *raumstruktureller Heterogenität* sprechen wir, wenn die funktionalen Formen und/oder Parameter der Modellgleichungen für die einzelnen Regionen nicht konstant sind, sondern sich mit ihrer räumlichen Lage ändern. Darunter fällt z.B. auch räumlich induzierte Heteroskedastie. In der Regel hängen Charakteristika der einzelnen Regionen von ihrer Distanz zu einer Einflußquelle ab; zu schätzen sind dann entsprechende Entfernungsbewertungs-Funktionen.

Der Begriff *raumstrukturelle Abhängigkeit* faßt die zweite Sorte räumlicher Effekte zusammen. Er drückt aus, daß gerade bei räumlichen Querschnitten keine Unabhängigkeit der Beobachtungen voneinander unterstellt werden kann. Ganz offensichtlich müssen wir gerade bei ökonomischen Daten annehmen, daß die einzelnen Regionen einander beeinflussen, da die meisten der wichtigen ökonomischen Größen wie Geldkapital, Arbeitskräfte, Waren etc. (mehr oder weniger) mobil sind. Unter Abhängigkeit soll in dieser Arbeit *raumstrukturelle Autokorrelation* verstanden werden. Diese liege vor, falls für wenigstens ein Paar von Beobachtungen y_r, y_s (mit $r, s \in \{1, \dots, R\}$) aus den R untersuchten Regionen eine Kovarianz ungleich Null ist.⁸

Da nur die raumstrukturelle Autokorrelation, jedoch nicht die Heterogenität Thema dieser Arbeit ist, seien nur die häufigsten Modellklassen dieses Teilgebietes⁹ genannt. Eine verbreitete Behandlung der Heterogenität ist die *raumstrukturelle Expansionsmethode (Spatial Expansion Method)*, welche zuerst von Casetti vorgeschlagen¹⁰ wurde. Weiter werden in diesem Kontext aus der Zeitreihenökonomie übernommene Modelle wie v.a. *Random Coefficient Variation* diskutiert.¹¹

⁷ vgl. ANSELIN (1988), Kap.2.

⁸ Die Definition geht auf CLIFF/ORD (1973), S. 2, zurück und wird am häufigsten herangezogen.

⁹ für eine Einführung: ANSELIN (1988), Kap. 9.

¹⁰ vgl. CASETTI (1972). Eine empfehlenswerte Darstellung des Grundkonzepts sowie eine kurze Diskussion findet sich außerdem bei ANSELIN (1988), Kap. 9.3. Erweiterungen des Modells und zusätzliche Eigenschaften werden in CASETTI/POON (1995) beschrieben.

¹¹ Solche Alternativen erörtern z.B. ARORA/BROWN (1977). Anwendungsbeispiele finden sich in KAU/LEE/SIRMANS (1986).

3. Modelle mit raumstruktureller Autokorrelation

Autokorrelation im Raum kann auf sehr verschiedene Arten entstehen. Ökonomische Gesetzmäßigkeiten (z.B. Anbieter wie Supermärkte oder Tankstellen in verschiedenen Landkreisen, welche alle um die Kundschaft dieser Landkreise werben) kommen als Ursache ebenso in Frage wie über Grenzen hinweg wirkende Umwelteinflüsse oder Nachahmungseffekte, z.B. im Konsum- oder Investitionsverhalten¹².

Daneben können auch bei räumlichen Messungen auftretende Datenprobleme die Abhängigkeiten scheinbar verstärken oder sogar künstlich erzeugen. Das Grundproblem besteht dabei in der korrekten Abgrenzung zwischen den betrachteten Regionen, nachdem die politischen Grenzen gerade von Gemeinden oft wenig Rücksicht auf die tatsächlichen (wirtschaftlichen) Verflechtungen nehmen. Selbst wenn also in sich homogene und untereinander verschiedene Regionen existieren mögen, ist es oft wegen der nicht kongruenten politischen Grenzen unmöglich, eben diese Regionen auch zu messen.¹³ Ein weiterer Aspekt besteht darin, daß sich Individuen oft nicht einer einzigen Region zuordnen lassen, weil sie z.B. in der einen arbeiten und der anderen wohnen. Solche Effekte gleichen sich meist nicht zwischen den Regionen aus; man denke an die bundesdeutschen Stadtstaaten.

Modellieren lassen sich diese Autokorrelationen, indem man entweder Abhängigkeiten im Fehlerterm konstruiert, und/oder direkt die Abhängigkeit der endogenen Größen (auch) voneinander einführt. Aufgrund der Verwandtschaft zu ähnlichen Konzepten in der Zeitreihenökometrie bietet es sich an, von *raumstrukturellen Moving-Average-Prozessen* (*Spatial MA* oder *SpMA*) bzw. von *raumstrukturellen Autoregressiven Prozessen* (*Spatial AR* oder *SpAR*) zu sprechen, auch wenn etliche Autoren diese Bezeichnungen noch nicht übernommen haben bzw. keine Unterscheidungen zwischen verschiedenen Arten von Autokorrelation treffen.¹⁴

3.1 Raumstrukturelle Autoregression (SpAR)

Direkte Abhängigkeiten der Beobachtungen aus unterschiedlichen Regionen untereinander sind oft begründbar. Sie können positiv (z.B. bei Bruttoinlandsprodukten durch internationalen Handel, oder bei der Bereitstellung öffentlichen Güter wegen Nachahmungseffekten), aber auch negativ sein (z.B. verschiedene Anbieter bemühen sich regionenübergreifend um Kunden). Solche Zusammenhänge lassen sich durch autoregressive Prozesse darstellen. Die gebräuchlichste Formulierung¹⁵ geht auf WHITTLE (1954) zurück:

¹² CASE (1991) findet keine beobachtbare Nachahmung des Konsummusters benachbarter Haushalte.

¹³ Für eine ausführliche Diskussion räumlicher Datenprobleme vgl. HAINING (1995).

¹⁴ Schätztechnische Aspekte werden erst in Abschnitt 4 behandelt.

¹⁵ Ein alternativer Ansatz (Conditional SpAR) hat sich nicht durchsetzen können. Für einen Überblick vgl. BESAG (1974).

$$y_r = \sum_{s \neq r} w_{r,s} y_s + \varepsilon_r ; \quad (3 - 1)$$

$$\text{wobei: } E(\varepsilon_r) = 0; E(\varepsilon_r \varepsilon_s) = 0; V(\varepsilon_r) = \sigma_r^2;$$

Dabei spiegeln die $w_{r,s}$ die räumlichen Strukturen wider, indem sie den Spill-Over von s auf r angeben. Diese Gewichtungsfaktoren seien vorgegeben. Man beachte, daß keine Isotropie angenommen wurde, die Wirkung der Schocks mag also richtungsabhängig sein. Die diffizile Bestimmung dieser Gewichte wird in Unterabschnitt 3.3 diskutiert.

Dieser *Simultane Autoregressive Prozeß* (*Simultaneous Spatial AR-Scheme* oder *SSpAR*, meist einfach *SpAR*) wird gewöhnlich in Matrixschreibweise notiert:

$$y = \rho W y + \mu + \varepsilon; \quad \text{mit} \quad W = \begin{pmatrix} 0 & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & 0 & & \\ \vdots & & \ddots & \vdots \\ w_{R,1} & w_{R,2} & \cdots & 0 \end{pmatrix}. \quad (3 - 2)$$

mit y , μ und ε als $R \times 1$ -Vektoren der Beobachtungen, Konstanten bzw. Fehlerterme. (Im Regressionsmodell wird μ durch $X\beta$ ersetzt.) Die relativen Größen der Wechselwirkung werden in der $R \times R$ -Gewichtungsmatrix W vorgegeben, welche im allgemeinen keine Symmetrie aufweist. Somit ist der die absolute Stärke der Wechselwirkung angegebende Autokorrelationsparameter ρ nicht eindeutig identifiziert, solange die Elemente der Gewichtungsmatrix nicht normiert sind. In der ersten Darstellung hatten die Gewichte $w_{r,s}$ beide Funktionen übernommen.

Für die Varianz dieses Prozesses gilt

$$V(y) = \Omega = (I - \rho W)^{-1} V(\varepsilon) (I - \rho W')^{-1}. \quad (3 - 3)$$

3.2 Raumstrukturelle Autokorrelation im Fehlerterm

Unbeobachtbare Variablen können auf mehrere benachbarte Regionen Einflüsse ausüben oder selbst raumstrukturell autokorreliert sein. Beispiele dafür können Umweltbedingungen (Klima, Bodenqualität, ...) sein, aber auch die Bestimmungsfaktoren von FuE-Aktivitäten der Unternehmen einer Branche bei Spill-Overs. Solche Situationen motivieren zu einer Formulierung, in welcher die Beobachtung y_r einer Region r auch durch die Störterme ε_i anderer der R Regionen erklärt werden:

$$y_r = \mu_r + \varepsilon_r + \sum_{s \neq r} m_{r,s} \varepsilon_s, \quad (3 - 4)$$

$$E(\varepsilon_r) = 0; E(\varepsilon_r^2) = \sigma_r^2; E(\varepsilon_r \varepsilon_s) = 0 \quad \forall r \neq s.$$

Dabei stellen die $m_{r,s}$ ebenso Gewichtungsfaktoren mit den gleichen Eigenschaften dar wie die $w_{r,s}$ beim SpAR-Modell. Unterstellt man homoskedastische Störterme mit $\sigma_r = \sigma \quad \forall r$, gilt für die Varianz des Prozesses:

$$\begin{aligned}
V(y_r) &= \gamma = E((y_r - \mu_r)^2) = \\
&= E(\varepsilon_r^2 + m_{r,1}^2 \varepsilon_1^2 + \dots + m_{r,R}^2 \varepsilon_R^2 + 2m_{r,1} \varepsilon_r \varepsilon_1 + \dots) \\
&= \sigma^2 (1 + m_{r,1}^2 + \dots + m_{r,R}^2).
\end{aligned} \tag{3 - 5}$$

Daraus folgt offenbar, daß die Stationarität nur gesichert ist, falls

$$\sum m_{r,s}^2 < \infty, \tag{3 - 6}$$

denn eine unendliche Varianz würde die Stationaritätsannahme verletzen, daß die Varianz unabhängig von der Position im Raum ist. Die Kovarianz läßt sich mit

$$Cov(y_r, y_s) = \gamma_{r,s} = \sigma^2 \left(m_{r,s} + m_{s,r} + \sum_{i \neq r,s} m_{r,i} m_{s,i} \right) \tag{3 - 7}$$

angeben.¹⁶ Der Kovarianzausdruck (3 - 7) ist übrigens nicht, wie gewohnt, auch für den Fall $y_r = y_s$, also die Varianz, definiert.

Analog zu (3 - 2) lautet die Darstellung dieses Prozesses in Matrixform:

$$y = \mu + (I + \theta M)\varepsilon. \tag{3 - 8}$$

Der skalare Parameter θ hat hier die gleiche Funktion wie der MA-Parameter bei einer Zeitreihendarstellung; er bestimmt die Stärke der gegenseitigen Beeinflußung.

Neben dieser Darstellung gibt es weitere Vorschläge, raumstrukturelle Abhängigkeiten im Fehlerterm zu formulieren. Eine interessante Spezifikation mit einer einfacheren Kovarianzstruktur, wobei getrennte Störterme für die jeweilige Region bzw. deren Außenwirkung auftreten, schlagen CLIFF/ORD (1973)¹⁷ vor.

Weitgehend durchgesetzt hat sich jedoch, den Fehlerterm analog zur Autoregression in der Endogenen zu formulieren:

$$\begin{aligned}
y &= E(y) + \varepsilon, \\
\text{wobei } \varepsilon &= \theta M \varepsilon + v, \quad E(vv') = \sigma^2 I,
\end{aligned} \tag{3 - 9}$$

weshalb diese Darstellung unter *Autoregression im Fehlerterm* geführt wird.

3.3 Die Konstruktion geeigneter raumstruktureller Gewichtungsmatrizen

Aus den bisherigen Darstellungen dieses Abschnittes wurde deutlich, welche zentrale Rolle *Gewichtungsmatrizen* bei raumstrukturellen Prozessen spielen. Da sie jedoch als exogene Größe in eine Schätzung eingehen, ist die Frage nach der korrekten Bestimmung der Gewichte von zentraler Bedeutung. Unstrittig ist lediglich, daß die Haupt-

¹⁶ Die Herleitung findet sich in Anhang A1.

¹⁷ vgl. dort S.150f. und S.161.

diagonale der Gewichtungsmatrix bei jeder Autokorrelation unbesetzt bleibt, da sonst eine Selbstbezüglichkeit dargestellt würde.

Die einfachste Form von Gewichtungsmatrizen sind *Nachbarschaftsmatrizen* (*Contiguity Matrices*). Viele Autoren orientieren sich dabei an regelmäßigen Mustern, wobei meist eine Anordnung gleich großer, z.B. wabenförmiger oder quadratischer Regionen, wie bei einem Schachbrett, betrachtet wird. Entsprechend den Bewegungen, welche den Schachfiguren erlaubt sind, werden in der Literatur Typen von Nachbarschaften unterschieden. Im *Turm*-Fall gelten zwei Regionen als benachbart, die eine gemeinsame Kante teilen. Stoßen die beiden Regionen nur in einer Ecke aneinander, spricht man vom *Läufer*-Fall. Das *Königin*-Kriterium hingegen fordert, daß einer der beiden genannten Fälle vorliegt. Natürlich können diese grundlegenden Typen variiert werden.¹⁸ So kann z.B. eine *Doppelturm*-Nachbarschaft definiert werden, wenn die betreffenden beiden Regionen waagrecht oder senkrecht nur durch genau eine dritte Region getrennt werden. Gemeinsam ist diesen Konzepten, daß der jeweilige Gewichtungsfaktor den Wert "Eins" annimmt, falls das Nachbarschaftskriterium erfüllt ist, und ansonsten "Null" wird.

Der Vorteil dieser Konzeption liegt darin, daß sie exakte Ordnungsrelationen festzulegen gestattet, und sich auch bei nicht unmittelbar benachbarten Regionen deren Korrelation bestimmen läßt. Abstufungsgrade von Nachbarschaften sind unterscheidbar und erlauben somit eine weitgehende Analogie zur Zeitreihenbetrachtung. Im übrigen erhält man auf diese Weise eine symmetrische Gewichtungsmatrix. Eine theoretische Schwäche besteht in folgender Unklarheit: Soll höhere Korrelation angenommen werden, wenn die Nachbarschaft über mehrere Strecken festgelegt werden kann? (So ist auf einem Schachbrett das Feld B4 von A1 aus über drei verschiedene Kombinationen von einem Läufer- und zwei Turmbewegungen zu erreichen.) Fraglich ist, ob die Zahl der wählbaren Strecken einbezogen werden soll; dann kann das Problem auftreten, daß für weiter entfernte Regionen eine höhere Korrelation errechnet wird als für näher liegende. Des weiteren offenbart ein Blick auf eine Europakarte das für die praktische Anwendung entscheidende Erschwernis: die Grenzen zwischen den Staaten verlaufen keineswegs entlang der vier Himmelsrichtungen. Eine Läufer-Nachbarschaft in Europa finden zu wollen, ist beinahe aussichtslos.¹⁹ Bei dieser Gelegenheit macht es sich offensichtlich bemerkbar, daß die meisten Untersuchungen zu Raumstrukturen bislang in den USA stattfanden, wo die politischen Grenzen eine einfachere Gestalt aufweisen als auf dem Alten Kontinent.

Aus den genannten Gründen erscheint es naheliegend, auf die Analogie zum Schachbrett oder ähnlichen Mustern zu verzichten, und Nachbarschaft so zu definieren, daß eine gemeinsame Grenze zwischen den betrachteten Regionen vorliegt. Aber auch dieses Kriterium verursacht Komplikationen. So sind Rußland und Polen wegen des Gebietes von Königsberg (Kaliningrad) direkte Nachbarn, aber die eigentlichen Staatsgebiete sind 500 km voneinander entfernt. Liechtenstein, Italien, Ungarn und die Ukraine sind nach diesem Kriterium einige der gleichgewichteten Nachbarn zweiten Grades von Deutschland, was in den meisten Anwendungen kaum sinnvoll sein kann.

¹⁸ Eine Übersicht über gängige Kriterien bieten KELEJIAN/ROBINSON (1995), S.95.

¹⁹ Die vier österreichischen Bezirke Feldkirchen, St. Veit, Klagenfurt-Stadt sowie Klagenfurt-Land bieten das einzige dem Autor bekannte Beispiel.

Außerdem bleibt unklar, wie Seewege bewertet werden sollen: Ist Großbritannien mit niemandem benachbart, zumindest mit Irland, oder seit der Fertigstellung des Kanaltunnels auch mit Frankreich? Solche Fragen stellen sich im übrigen auch auf der Ebene kleinerer Verwaltungseinheiten, z.B. bei Betrachtung der Schweizer Kantone.

Die raumstrukturellen Korrelationsparameter können nur sinnvoll interpretiert und zwischen verschiedenen Betrachtungen verglichen werden, wenn die Gewichtungsmatrizen standardisiert werden. Dies geschieht durch die sogenannte *Reihennormalisierung*: Die Gewichte einer Reihe werden alle durch ihre Summe dividiert, was bei dem einfachen Nachbarschaftskriterium bedeutet, daß für eine Region r alle Nachbarn das Gewicht ($1 / \text{Anzahl der Nachbarn}$) erhalten. Diese Annahme, jede Region unterliege dem gleich starken Einfluß von außen, ist jedoch ebenfalls zu bezweifeln, wenn man z.B. ein kleines Land inmitten des Beobachtungsgebietes wie die Schweiz vergleicht mit einem großen wie Spanien, das am Rande Europas liegt. Sicherlich wird die Schweiz zumindest in einigen Teilbereichen wie dem Arbeitsmarkt ungleich mehr von der Entwicklung jenseits ihrer Grenzen beeinflusst. Unlogisch erscheint auch eine weitere aus dem Konzept abzuleitende Aussage. So soll die Wirkung z.B. Italiens auf Österreich bedingt durch die Aufteilung der CSFR in Tschechien und die Slowakei abgenommen haben (von einem Sechstel auf ein Siebtel, also immerhin um knapp zehn Prozent!). Im übrigen verliert die Gewichtungsmatrix durch die Reihennormalisierung ihre Symmetrieeigenschaft.

Zusätzlich ist noch zu bedenken, daß bei einer derartigen Spezifikationsweise recht unterschiedliche Raumstrukturen zu einer identischen Nachbarschaftsmatrix führen können. Aus solchen Überlegungen heraus wird vorgeschlagen, diese binäre Betrachtungsweise zu ersetzen durch einen Ansatz, welcher die Gewichtungsfaktoren abhängig macht von der Entfernung der beiden Regionen voneinander. Dies geschieht dann über eine bereits beschriebene Entfernungsbewertungs-Funktion²⁰. Schwierigkeiten ergeben sich nun eher bei flächenmäßig ausgebreiteten Regionen. Geometrische Konzepte wie Entfernung der Schwerpunkte oder die kürzeste Verbindung zwischen beiden Staatsgebieten können fehlleiten, wenn die ökonomischen Aktivitäten eines Landes nicht gleichmäßig verteilt sind. Ein Beispiel dafür ist Schweden.

Ein aus der Außenhandelstheorie bekanntes Konzept ist in diesem Kontext erwägenswert, nämlich, die Distanz in Transportkosten zu messen. Bei etlichen Anwendungen dürfte dies tatsächlich der relevante Maßstab sein. Man sollte jedoch nicht aus dem Auge verlieren, daß Wechselwirkungen vorliegen: Die Verkehrsinfrastruktur wird auch nur dort ausgebaut, wo entsprechende Aktivitäten dies lohnenswert erscheinen lassen. Gegebenenfalls müssen auch politische Konstellationen berücksichtigt werden, welche Handelshemmnisse nach sich ziehen. Man denke nur an die frühere Situation zwischen EG und den damaligen COMECON-Staaten. Auch soziale oder historische Phänomene können eine Rolle spielen. So sind die kulturelle und ökonomische Bande zwischen Österreich und der italienischen Provinz "Alto Adige" bis heute nicht durchschnitten. Verschiedene Distanzkonzepte lassen sich auch kombinieren, indem man einen Vektor der verschiedenen Entfernungsaspekte aufstellt.²¹ Mit einem auf die jeweilige Anwendung ausgerichteten Bewertungsvektor multipliziert, läßt sich

²⁰ vgl. Abschnitt 2.1, Formel (2 - 3).

²¹ Ein entsprechender Vorschlag stammt von ISARD (1969), S.821 ff.

eine ausgefeilte Entfernungsmeßzahl angeben, welche *Effektive Distanz* (*Effective Distance*) genannt wird. Allerdings muß neben dem hohen Aufwand auch die Schwierigkeit gesehen werden, Größen wie "historische Verbundenheit" kardinal auszudrücken.

Auch wenn ein angemessenes Distanzkonzept gefunden ist, sind weitere Gestaltungsmöglichkeiten für die Gewichtungsmatrix denkbar. Ein populärer Vorschlag²² besteht darin, zusätzlich zur Distanz die relative Länge der gemeinsamen Grenze einzubeziehen. Das Gewicht kann z.B. durch die multiplikative Verknüpfung einer Entfernungsbewertungs-Funktion mit dem Anteil der betrachteten Region an der eigenen Grenze bestimmt werden. Vor allem für Anwendungen, bei welchen der "kleine Grenzverkehr" von Arbeitskräften oder Waren wichtig ist, ist dies eine interessante Spezifikation. Zu beachten ist, daß aus der relativen Betrachtung der Grenzlängen auch ohne Reihennormalisierung eine asymmetrische Gewichtungsmatrix resultiert.²³

Ein weiteres Dilemma: Selten bietet sich ein Raumgebiet zur Untersuchung an, welches an seinen Rändern keinen Einflüssen von außen unterliegt. Falls sich diese nicht anderweitig einbeziehen lassen, ist anzunehmen, daß diese Abhängigkeiten in den geschätzten Korrelationen der Regionen des Untersuchungsgebietes aufgefangen werden, und so Verzerrungen auslösen.

Bei allem Bemühen, eine möglichst zutreffende Gewichtungsmatrix zu erstellen, sollte nicht vergessen werden, daß eine vollbesetzte Matrix bei Schätzungen erhebliche Rechenprobleme verursachen kann. Deshalb mag es sinnvoll sein, einen Gewichtungsfaktor erst dann nach einem Entfernungskonzept ungleich "Null" zu wählen, wenn er ein bestimmtes Nachbarschaftskriterium erfüllt (z.B. Nachbarschaft von höchstens Grad zwei). Eine solche *sparsame* Gewichtungsmatrix ist besonders bei ML-Schätzungen wichtig.²⁴

Allerdings erweckt ein Blick auf die gängige Literatur den Eindruck, daß bislang selten Augenmerk auf eine optimale Gewichtungsmatrix gelegt wurde. Dabei kann ein noch so geschicktes Verfahren zur Schätzung der absoluten Stärke gegenseitiger Abhängigkeit Fehler nicht mehr korrigieren, welche bei einer wenig sorgfältigen Spezifikation der Gewichte, also der relativen Abhängigkeit, gemacht wurden.

4. Eigenschaften einiger Schätzverfahren bei raumstrukturellen Abhängigkeiten

Aufgrund der mehrdimensionalen Natur von raumstrukturellen Prozessen gehen wünschenswerte, aus dem Zeitreihenkontext bekannten Eigenschaften verschiedener Schätzverfahren teilweise verloren. *Maximum-Likelihood*-Schätzer (ML-Schätzer) weisen bei bestimmten Spezifikationen rechentechnische Probleme auf, welche wegen der Gewichtungsmatrix mit der Anzahl der Regionen zunehmen. *Kleinst-Quadrat*-Schätzer (LS-Schätzer) hingegen sind bei Autokorrelation in der Endogenen verzerrt

²² vgl. CLIFF/ORD (1973), S.12f.

²³ Weitere Vorschläge in dieser Richtung finden sich bei ANSELIN (1988), S.20f.

²⁴ vgl. Unterabschnitt 4.3.

und inkonsistent. Die folgende Darstellung beschränkt sich auf solche Prozesse. Allerdings kann auch Autokorrelation im Fehlerterm, abhängig von der genauen Spezifikation, zur Inkonsistenz des verwendeten Schätzers führen. In jedem Fall sind dessen Effizienzeigenschaften zu überprüfen.

4.1 OLS-Schätzung ohne Berücksichtigung von Raumstrukturen

Zunächst ist es nützlich festzuhalten, daß ein für ein umfassendes Gebiet (z.B. die Bundesrepublik) formuliertes Modell nicht ohne weiteres auf die Regionen (z.B. Landkreise) übertragen werden kann, aus denen das Gebiet besteht. Dies ist nur möglich, wenn die Regionen sich nicht gegenseitig beeinflussen, weil der Parametervektor β_{Reg} des Regionenmodells nur Reaktionen auf exogene Vorgänge in der betroffenen Region selbst angibt. Im Falle der Abhängigkeit unterscheidet sich β_{Reg} vom Parametervektor β_{Geb} des gebietsbezogenen Modells um den Faktor $(1 - \rho)$.

Davon zu unterscheiden ist der Fall, daß ein Schätzmodell auf der Regionenebene raumstrukturelle Autokorrelation fälschlicherweise nicht berücksichtigt. Dann werden die Steigungsparameter, abhängig vom wahren Autokorrelationsparameter ρ , im Absolutwert überschätzt. Verantwortlich für diesen Bias ist der Umstand, daß jede Region aufgrund der Zyklizität raumstruktureller Korrelation sich über ihre Nachbarregionen wieder selbst beinflusst. Das Ausmaß der Verzerrung kann wie folgt abgeschätzt werden²⁵:

$$\text{Wahres Modell: } y = \rho W y + X \beta_0 + \varepsilon$$

$$\text{mit } \beta_0 = (\alpha, \beta)' \quad \text{und } X = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_R \end{pmatrix}.$$

$$\hat{\beta}_0 = (X' X)^{-1} X' y \Rightarrow$$

$$E(\hat{\alpha}) \approx \frac{\alpha}{1 - \rho} + \frac{\rho - \rho^2}{1 - \rho} \frac{1}{w} \bar{X} \beta$$

$$E(\hat{\beta}) \approx \left(1 + \frac{\rho^2}{w(1 - \rho)} \right) \beta,$$

(4 - 1)

wobei w : (Median der Anzahl an Nachbarregionen + 1).

Für eine empirische Bestätigung wurde in einer Monte Carlo-Studie ein wahrer SpAR-Prozeß wie in (4 - 1) unterstellt. Zur Gewichtung diente eine Nachbarschaftsmatrix von 93 politischen Bezirken Österreichs. Der Fehlerterm ist unabhängig standardnormalverteilt, die X_r wurden ebenfalls unabhängig normalverteilt gezogen mit $\mu_x = (.25; .1)$; $\Omega_x = I$; und außerdem $\beta = (2.0; 5.0)$.

²⁵ Für eine ausführliche Herleitung vgl. Anhang A2.

Für verschiedene ρ wurden je 1000 Replikationen durchgeführt; für jede Durchführung ein kompletter Datensatz zufallsbestimmt und dann mit OLS geschätzt. Die Resultate finden sich in *Tabelle 1*. Die angegebenen erwarteten Verzerrungen $E(\hat{\beta}_i - \beta_i)$ wurden mit der Formel aus (4 - 1) bestimmt und mit der tatsächlich gemessenen Abweichung $(\hat{\beta}_i - \beta_i)$ verglichen.

Tabelle 1: Monte Carlo-Studie: Verzerrung bei nicht spezifiziertem raumstrukturellen AR-Prozeß unter Verwendung einer reihennormalisierten Nachbarschaftsmatrix österreichischer Bezirke

ρ	$E(\hat{\beta}_i - \beta_i)$	$\hat{\beta}_i - \beta_i$	RMAD	$E(\hat{\alpha} - \alpha)$	$\hat{\alpha} - \alpha$	RMAD
-.30	.0257	.0237	.0330	-.2565	-.2558	.0419
-.25	.0167	.0165	.0315	-.2167	-.2080	.0371
-.20	.0100	.0108	.0293	-.1767	-.1768	.0324
-.15	.0053	.0060	.0265	-.1357	-.1387	.0297
-.10	.0022	.0027	.0253	-.0931	-.0938	.0252
-.05	.0005	.0009	.0245	-.0483	-.0405	.0241
.00	0	.0002	.0238	0	.0032	.0247
.05	.0005	.0000	.0243	.0521	.0570	.0265
.10	.0022	.0024	.0257	.1089	.1155	.0314
.15	.0053	.0052	.0265	.1712	.1711	.0369
.20	.0100	.0097	.0302	.2600	.2440	.0448
.25	.0167	.0153	.0326	.3167	.3179	.0547
.30	.0257	.0241	.0361	.4029	.4134	.0649

Für Österreich gilt $w = 1 + 4 = 5$. Die Standardfehler bewegen sich in der Größenordnung 10^{-3} . Beim Steigungsparameter ist das arithmetische Mittel der Verzerrungen der beiden β angegeben. Der RMAD (Mittlere absolute Abweichung vom wahren Wert relativ zu diesem) vermittelt einen Eindruck von der Schätzgüte. Weitere Erläuterungen siehe Text.

4.2 OLS-Schätzung des raumstrukturellen AR-Prozesses

Die raumstrukturelle OLS-Schätzung ist auch bei trivial formulierten Prozessen verzerrt. Die einfachste SpAR-Spezifikation in Matrixschreibweise stellt das Modell von Whittle aus (3 - 2) mit $\mu = 0$ dar:²⁶

$$y = \rho W y + \varepsilon . \quad (4 - 2)$$

Dabei seien Fehlerterme mit "White Noise"-Eigenschaften unterstellt. Mit $(W y)$ als erklärender Variable lautet der OLS-Schätzer für den Autokorrelationsparameter hier

$$\hat{\rho} = \left[(W y)' (W y) \right]^{-1} (W y)' y . \quad (4 - 3)$$

²⁶ Für eine ähnliche Darstellung vgl. ANSELIN (1988), S.57ff.

Dessen Erwartungswert²⁷

$$E(\hat{\rho}) = \rho + E\left[\left((Wy)'(Wy)\right)^{-1} (Wy)' \varepsilon\right] \quad (4 - 4)$$

ist ungleich ρ , sobald Autokorrelation vorliegt, weil bei $\rho \neq 0$ gilt: $E[(Wy)'\varepsilon] \neq 0$.²⁸ Mit der Korrelation zwischen Fehlerterm und Regressor ist eine Basisannahme des OLS-Verfahrens verletzt²⁹, die auch die Konsistenzeigenschaften beeinträchtigt. Der OLS-Schätzer wäre zumindest asymptotisch konsistent, wenn

$$\begin{aligned} \text{plim} \frac{1}{R} \left((Wy)'(Wy) \right) &= Q < \infty \\ \wedge \text{plim} \frac{1}{R} \left((Wy)' \varepsilon \right) &= 0 \end{aligned} \quad (4 - 5)$$

gälte. Während die erste Prämisse, bei welcher Q eine endliche und nichtsinguläre Matrix darstellt, zumindest durch entsprechende Restriktionen auf ρ und W erfüllt werden kann, wird die zweite Bedingung verletzt, sobald raumstrukturelle Autokorrelation vorliegt. Dies liegt in dem aus der Gewichtungsmatrix (bei $\rho \neq 0$) resultierenden Term begründet, welcher mit der Summe der Fehlertermquadrate multipliziert wird, so daß der Gesamtausdruck nicht Null werden kann.

Diese Resultate sind intuitiv nicht überraschend, da die Spezifikation den aus der Standardökonomie bekannten Simultaneous Equations-Modellen sehr verwandt ist, bei denen mit OLS die gleichen Schwierigkeiten auftreten. Ebenso wie bei diesen stellt die Instrumentvariablenschätzung eine Lösungsalternative auch für SpAR-Prozesse dar.³⁰ Dies wird bei dem in Abschnitt 6 aufgeführten GMM-Schätzer berücksichtigt.

ORD (1975) schlägt einen modifizierten LS-Schätzer mit einer die Unverzerrtheit sichernden Momentenrestriktion vor. Allerdings räumt der Autor selbst die starke Ineffizienz dieses Verfahrens ein, welche vom wahren ρ abhängig ist und mit steigender Autokorrelation zunimmt. Bei einem $\rho = 0.5$ z.B. ist die Varianz eines alternativen ML-Schätzers weniger als nur halb so hoch wie die des Ord'schen Schätzers.

4.3 ML-Schätzung des raumstrukturellen AR-Prozesses

Die ML-Schätzung raumstruktureller AR-Prozesse weist die gleichen asymptotischen Eigenschaften auf wie allgemeine ML-Schätzer. Problematisch ist jedoch, daß die erforderliche nichtlineare Optimierung zu durch die Umformungen der Gewichtungsmatrix verursachten Rechenproblemen führen kann. Die Bedingungen erster und zweiter Ordnung für ein repräsentatives Modell werden angegeben.

²⁷ Nach Einsetzen von $y = \rho Wy + \varepsilon$.

²⁸ Nachweis: vgl. Anhang A3.

²⁹ Für ein instruktives einfaches Modell mit lediglich zwei Regionen vgl. PAELINCK/KLAASSEN (1979), S.51f., oder KLOTZ (1996), Kap. 5.1.

³⁰ Eine Darstellung raumstruktureller IV-Schätzung findet sich bei ANSELIN (1988), Kap.7.1.

Um zu veranschaulichen, daß Spezifikationserweiterungen im Fehlerterm möglich sind, ohne die Schätzung entscheidend zu erschweren, modifizieren wir das Modell aus 3.2 insoweit, als bei den Störtermen konditionale Heteroskedastie zugelassen wird: Für den Störterm der Region r gelte $\varepsilon_r = \varepsilon_r(z_r \times \alpha)$.³¹ Dabei ist z_r' ein $P \times 1$ -Vektor mit P (geographischen) Merkmalen der Region r , und α der dazugehörige Parametervektor. Außerdem definieren wir einen Vektor μ homoskedastischer und unabhängiger Störungen mit $\mu = (\Omega^{-1/2} \varepsilon)$, wobei Ω die nur auf der Diagonalen besetzte Kovarianzmatrix zu ε sei. Dann gilt nach entsprechenden Umformungen:

$$\mu = \Omega^{-1/2} [(I - \rho W)y - X\beta]. \quad (4 - 6)$$

Die Loglikelihoodfunktion lautet entsprechend:

$$LL = \text{const} - \frac{1}{2} \ln|\Omega| + \ln|I - \rho W| - \frac{1}{2} \mu' \mu, \quad (4 - 7)$$

$$\text{wobei } \mu' \mu = [(I - \rho W)y - X\beta]' \Omega^{-1} [(I - \rho W)y - X\beta].$$

Diese Loglikelihoodfunktion ist nun durch Differenzieren nach den gesuchten Parametern β , ρ , und α_p (dem p -ten Element des Vektors α) zu maximieren. Dies entspricht einer Minimierung der gewichteten quadratischen Fehler, korrigiert um die Bestandteile der Jacobimatrix. Die partiellen Ableitungen ergeben, gleich Null gesetzt, die Bedingungen erster Ordnung für die ML-Parameterschätzer:³²

$$0 = \frac{\partial LL}{\partial \beta} = X' \Omega^{-1/2} \mu.$$

$$0 = \frac{\partial LL}{\partial \rho} = -\text{tr}(I - \rho W)^{-1} + \mu' \Omega^{-1/2} W y. \quad (4 - 8)$$

$$0 = \frac{\partial LL}{\partial \alpha_p} = -\frac{1}{2} \text{tr} \Omega^{-1} H_p + \frac{1}{2} \mu' \Omega^{-1/2} H_p \mu.$$

H_p steht dabei für die Diagonalmatrix mit den Elementen $(\partial h(z\alpha) / \partial \alpha_p)$. Diese Gleichungen sind nichtlinear und deshalb im allgemeinen nicht analytisch lösbar. Auch numerische Methoden führen nicht immer zum Ziel, wenn die auftretende Gewichtungsmatrix nicht sparsam genug ist und/oder zu viele Regionen umfaßt.³³

Bestimmte, aber im allgemeinen unproblematischen Regularitätsbedingungen³⁴ sichern die asymptotischen Eigenschaften des ML-Schätzers.

Die hier sechs verschiedenen Elemente I_{ij} der Informationsmatrix lauten:

³¹ In ANSELIN (1988) findet sich eine ausführlichere Diskussion der ML-Schätzung seines Modells auf S.61ff.

³² Die Herleitung hierzu und zu den Elementen der Informationsmatrix in Anhang A4.

³³ vgl. einen kurzen Erfahrungsbericht zu ML-Schätzungen mit hoher Regionenzahl in KELEJIAN/PRUCHA (1995), 2.Abschnitt.

³⁴ vgl. z.B. ANSELIN (1988), S. 63.

$$\begin{aligned}
I_{\beta\beta} &= X' \Omega^{-1} X , \\
I_{\beta\rho} &= X' \Omega^{-1} W (I - \rho W)^{-1} X \beta , \\
I_{\beta\alpha} &= 0 , \\
I_{\rho\rho} &= \text{tr} \left(W (I - \rho W)^{-1} \right)^2 + \text{tr} \left[\Omega \left(W (I - \rho W)^{-1} \right)' \Omega^{-1} \left(W (I - \rho W)^{-1} \right) \right] \\
&\quad + \left(W (I - \rho W)^{-1} X \beta \right)' \Omega^{-1} \left(W (I - \rho W)^{-1} X \beta \right) , \\
I_{\rho\alpha} &= \text{tr} \Omega^{-1} H_p W (I - \rho W)^{-1} , \\
I_{\alpha_p \alpha_q} &= \frac{1}{2} \text{tr} \Omega^{-2} H_p H_q .
\end{aligned} \tag{4 - 9}$$

Ein in ORD (1975) vorgestelltes Verfahren kann die Berechnung der Determinanten im Rahmen einer ML-Schätzung erleichtern. Dabei seien $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_R$ die Eigenwerte der Gewichtungsmatrix W . Dann gilt:

$$|I - \rho W| = \prod_{r=1}^R (1 - \rho \lambda_r) . \tag{4 - 10}$$

Der Vorteil besteht darin, daß die Eigenwerte, im Gegensatz zur Determinante, bei den einzelnen Iterationsschritten zur Optimierung der Schätzung von ρ nicht mehr neu berechnet werden müssen; die multiplikative Berechnung mit Hilfe der rechten Gleichungsseite und dem jeweiligen ρ genügt.

5. Tests auf Anwesenheit raumstruktureller Abhängigkeiten

Auf die Existenz raumstruktureller Abhängigkeiten kann getestet werden, indem in einem nistenden Modell, welches Raumstrukturen einbezieht, die betreffenden Korrelationskoeffizienten ρ bzw. θ auf Signifikanz überprüft werden. Die Nullhypothese "Es existieren keine raumstrukturellen Abhängigkeiten"³⁵ kann dort in Abhängigkeit vom t -Wert des betreffenden Koeffizienten auf einem bestimmten Signifikanzniveau überprüft werden (Wald-Test-Prinzip). Im letzten Abschnitt wurde jedoch deutlich, wie schwierig sich raumstrukturelle Schätzungen gestalten können. Deshalb erscheint es als sinnvoll, eine Regression ohne raumstrukturelle Prozesse durchzuführen, und dann auf raumstrukturelle Autokorrelation in den dabei entstandenen Residuen zu testen. Erst wenn der Test nahelegt, die Nullhypothese zu verwerfen, ist die geeignete Modellierung für den räumlichen Einfluß zu suchen.

Aus den in der Literatur diskutierten Testverfahren soll beispielhaft die Moran- I -Statistik herausgegriffen werden. Sie stellt historisch den ersten auf raumstrukturelle

³⁵ Zu beachten ist, daß sich die Gegenhypothese explizit auf die vermutete Raumstruktur bezieht.

Effekte konstruierten Test dar und wird, mitunter modifiziert, auch in den meisten aktuellen Veröffentlichungen verwendet.³⁶

Die Moran-*I*-Statistik wurde in MORAN (1950) für die Anwendung in regelmäßigen Gittermustern vorgeschlagen. Ihr Prinzip kann aber auf beliebige Raumstrukturen angewendet werden. Auch wenn die Statistik von Moran in Anlehnung an gewöhnliche Korrelationsmaße ad hoc vorgeschlagen wurde, läßt sie sich für verschiedene Spezifikationen explizit herleiten. BURRIDGE (1980) z.B. zeigt, daß sie identisch ist mit dem LM-Test eines Modells mit Autoregression im Fehlerterm.

Wir gehen aus von einer Regression mit *R* Regionen, bei der keine Größe für räumliche Zusammenhänge vorgesehen wurde. Wir erhalten den $R \times 1$ -Residuenvektor *e*. Bezeichnet *S* nun die Summe aller Elemente von einer unterstellten Gewichtungsmatrix *W*, lautet die Moran-*I*-Statistik in ihrer allgemeinen Form

$$I = \frac{R}{S} \frac{e' W e}{e' e} \quad (5 - 1)$$

Bei reihennormalisierter Gewichtungsmatrix vereinfacht sich der Ausdruck, da $R = S$ gilt. Dann hat er im übrigen die gleiche mathematische Form wie ein OLS-Schätzer von *We* auf *e*; wobei allerdings eine Erklärung nur in umgekehrter Reihenfolge sinnvoll erscheint. Über den Erwartungswert und die Varianz³⁷

$$E(I) = \frac{\text{tr}(MW)}{R - k} \quad \text{und}$$

$$V(I) = \frac{\text{tr}(MWMW') + \text{tr}(MW)^2 + \text{tr}^2(MW)}{(R - k)(R - k + 2)} - E^2(I) \quad (5 - 2)$$

$$(\text{wobei } M = I_R - X(X'X)^{-1}X',$$

W reihennormalisiert, *k* = Anzahl der Parameter)

erhalten wir durch Standardisierung von *I* eine asymptotisch normalverteilte Prüfgröße, welche somit einen Test der Nullhypothese erlaubt.

Nachdem für den Moran-*I*-Test eine Schätzung vorgeschaltet werden muß, ist die Frage erheblich, welche Schätzmethode am besten mit dieser Statistik harmoniert. Obwohl eine theoretische Untermauerung fehlt, haben "Monte Carlo"-Studien gezeigt, daß der Test mit OLS-Residuen die größte Macht erreicht.³⁸

Eigene Monte Carlo-Studien mit zufallsbestimmten Gewichtungsmatrizen verschiedener Größe zeigen auf, daß (in Abhängigkeit von ρ) eine relativ große Anzahl von Regionen nötig ist³⁹, um den Fehler 2. Art beim Moran-*I*-Test einzuschränken. Die Verbesserung der Macht dieses Testes mit steigender Regionenanzahl läßt sich auch im Vergleich der österreichischen Nachbarschaftsmatrix (93 Regionen) mit jener der westdeutschen Landkreise (327 Regionen) veranschaulichen (Tabelle 2):

³⁶ In der Literatur finden sich vor allem in jüngerer Zeit Vorschläge weiterer Testverfahren, die v.a. eine größere Macht aufweisen. Der neueste Beitrag ist ANSELIN ET AL. (1996).

³⁷ Zur Herleitung der Momente vgl. CLIFF/ORD (1973), S.200 ff.

³⁸ vgl. ANSELIN (1988), S. 101ff.

³⁹ Erst ab ca. 50 Regionen erscheint der Einsatz dieser Statistik überhaupt sinnvoll, ab ca. 300 Regionen liefert sie gute Ergebnisse.

Tabelle 2: Nichtablehnung der Nullhypothese "Keine raumstrukturelle Autokorrelation ($\rho_0 = 0$)" durch den Moran-I-Test in Abhängigkeit von der Regionenanzahl R

ρ	Nichtablehnungen in Prozent für ...	
	alte BR Deutschland ($R = 327$)	Österreich ($R = 93$)
-.30	19.1	84.3
-.20	43.1	84.5
-.10	66.8	88.5
00	93.2	93.6
.10	44.1	72.1
.20	2.7	31.0
.30	0.0	5.0

Signifikanzniveau: 95 % (zweiseitiger Test).

Bei $\rho = 0$ ist die Nichtablehnung der Nullhypothese korrekt.

Für nähere Angaben zur Durchführung der Studie vgl. Text in Unterabschnitt 4.1 bei *Tabelle 1*.

6. Schätzverfahren nach der Verallgemeinerten Momentenmethode

In der jüngeren Zeit gewinnt auch für raumstrukturelle Anwendungen das *GMM-Schätzprinzip* (*Generalized Method of Moments* oder *Verallgemeinertes Momentenprinzip*⁴⁰) an Aufmerksamkeit, welches in diesem Kontext asymptotisch unverzerrt ist und doch rechentechnisch beherrschbar bleibt. Ein weiterer Vorteil besteht darin, daß keine Verteilungsannahmen benötigt werden. Aufgrund der Freiheit in der Wahl der Momente sind sehr verschiedene GMM-Schätzer konstruierbar.

KELEJIAN/PRUCHA (1995) präsentieren ein GMM-Verfahren zur Schätzung eines Modells mit raumstruktureller Autokorrelation im Fehlerterm. Die Autoren stellen drei Momentengleichungen auf und schätzen, neben dem Parametervektor β , mit θ das Ausmaß der räumlichen Autokorrelation. In einer Monte Carlo-Studie ermitteln sie, daß ihr GMM-Schätzer trotz leichter Berechenbarkeit vergleichbare Effizienzeigenschaften aufweist wie der ML-Schätzer. Eine nähere Beschreibung findet sich in KLOTZ (1996).

Im folgenden wird ein dreistufiges Schätzverfahren von KELEJIAN/ROBINSON (1993) dargestellt, das Autoregression in der abhängigen Variablen (und im Fehlerterm) berücksichtigt.

Die Autoren gehen aus von einem recht allgemein gehaltenen Modell:

⁴⁰ Eine Darstellung der Grundzüge des GMM-Verfahrens findet sich z.B. in GOURIEROUX/MONFORT (1995), 9.5.

$$y = \rho W_1 y + X\beta + \varepsilon, \quad (6-1)$$

wobei $\varepsilon = W_2 u + v$.

Hierbei ist y ein $R \times 1$ -Vektor der beobachteten Variable in den R Regionen, X die $R \times k$ -Matrix der je k erklärenden Variablen pro Region mit dem zugehörigen Parametervektor β , sowie ρ der skalare Autokorrelationsparameter (wobei $|\rho| < 1$ angenommen wird). W_1 und W_2 stehen für exogene $R \times R$ -Gewichtungsmatrizen; dabei sei $(I - \rho W_1)$ nichtsingulär. Der Störvektor ε setzt sich additiv zusammen aus der regionspezifischen Störung v und den Schocks u , welche Spill-Over-Effekte von anderen Regionen darstellen. Diese Modellierung wurde in 3.2 nicht erwähnt und ist auch nicht üblich. Beide Fehlertermvektoren seien unabhängig sowohl voneinander als auch von X , und es gelte:⁴¹

$$\begin{aligned} E(u) &= E(v) = 0. \\ E(uu') &= \sigma_u^2 I_R. \\ E(vv') &= \sigma_v^2 I_R. \\ E(\varepsilon\varepsilon') &= \Omega_\varepsilon = \sigma_u^2 W_2 W_2' + \sigma_v^2 I_R. \end{aligned} \quad (6-2)$$

Es sei unterstrichen, daß dabei keine spezielle, insbesondere keine Normalverteilung angenommen wird, und daß ε im allgemeinen Heteroskedastie aufweist, es sei denn, die Elemente auf der Hauptdiagonale von $(W_2 W_2')$ wären alle identisch. Ein Parameter θ über die absolute Größe der raumstrukturellen Abhängigkeit im Fehlerterm wird hier deshalb unterdrückt, weil er im Verlauf der Schätzung neben σ_u nicht identifiziert werden könnte.

Umformuliert lautet das Modell bei unverändertem Fehlerterm ε

$$y = (I - \rho W_1)^{-1} X\beta + (I - \rho W_1)^{-1} \varepsilon. \quad (6-3)$$

Wie in Unterabschnitt 4.2 dargestellt, ist eine OLS-Schätzung für ein solches Modell nicht geeignet. Die Autoren schlagen deshalb einen dreistufigen GMM-Schätzer als Alternative vor. Die grundsätzliche Idee besteht darin, daß sich ein effizienter Schätzer für ρ und β formulieren läßt, wenn ein konsistenter Schätzer von Ω_ε zugrundeliegt. Diese erhält man, wenn in einer ersten Schätzstufe ρ und β konsistent, wenn auch ineffizient, geschätzt werden.

Dies ist möglich mit einer Version der Instrumentvariablenschätzung, bei welcher die in einer vorgeschobenen Schätzung erhaltenen geschätzten Erwartungswerte der abhängigen Variablen als Instrumente dienen. Die erste Stufe des GMM-Schätzers besteht also bereits aus zwei Schätzschritten und stützt sich dabei auf die Vorgehensweise der *Zweistufigen KQ-Methode (TSLS)*.⁴² Dies ist hier notwendig, weil der durch die Autokorrelation entstandene Regressor mit dem Störterm korreliert ist:

$$E[(W_1 y)\varepsilon'] = W_1 (I - \rho W_1)^{-1} \Omega_\varepsilon. \quad (6-4)$$

Man wählt eine Matrix H , so daß

⁴¹ Für Zwischenschritte vgl. Anhang A5.

⁴² vgl. GOURIEROUX / MONFORT (1995), Kap. 9.4.5.

$$H = (X, W_1 X^*) \quad \text{mit}$$

$$\lim_{R \rightarrow \infty} \frac{1}{R} H'H = \Omega_H \quad (6 - 5)$$

$$\text{rank}(H) = k + k^*; \quad k^* < k.$$

gilt. Dabei sei X^* eine Teilmatrix mit k^* Spalten aus X , und Ω_H nichtsingulär. Damit die Identifikation gesichert bleibt, soll die Regressorenmatrix auch mit der Erweiterung vollen Rang behalten. Leider erläutern die Autoren nicht, wie sie die k^* Spalten aus X auswählen. Es ist anzunehmen, daß diese Auswahl Einfluß auf die Eigenschaften der Schätzprozedur ausübt, insbesondere auf die Schätzvarianz, weshalb ein Test auf eine günstige Spaltenauswahl erwägenswert erscheint. Gemäß TSLS wird nun der andere Regressor, $W_1 y$, auf H geschätzt, mit α als Parametervektor und einem Fehlerterm v , woraus sich ein geschätztes $W_1 y$ ergibt:

$$W_1 y = H\alpha + v$$

$$\Rightarrow \hat{W}_1 y = H(H'H)^{-1} H'W_1 y. \quad (6 - 6)$$

Um im folgenden die Notation zu vereinfachen, wird das Modell (6 - 3) noch einmal umgeschrieben, indem die erklärenden Variablen zusammengefaßt werden:

$$y = Z\xi + \varepsilon. \quad (6 - 7)$$

Dabei sei $Z = (W_1 y, X)$ und $\xi' = (\rho, \beta')$. Entsprechend gilt $\hat{Z} = (\hat{W}_1 y, X)$. Der für ξ nun vorgeschlagene Schätzer lautet:

$$\hat{\xi} = (\hat{Z}'Z)^{-1} \hat{Z}'y. \quad (6 - 8)$$

Wie herkömmliche TSLS-Schätzer auch, ist dieser konsistent:

$$\text{plim}(\hat{\xi}) = \xi. \quad (6 - 9)$$

Von den für einen Beweis erforderlichen Annahmen soll lediglich erwähnt werden, daß zur Konsistenz des Schätzers sparsame Gewichtungsmatrizen erforderlich sind. "Sparsam" bedeutet im Kontext dieses Beweises, daß in jeder Reihe die Anzahl der Elemente, die ungleich Null sind, endlich bleibt, wenn die Anzahl der Regionen gegen unendlich strebt. Außerdem müssen die Regionen so geordnet werden können, daß ab einem gewissen Abstand zweier Regionen voneinander der gemeinsame Gewichtungsfaktor Null wird.

Mit Hilfe der konsistenten Schätzung von ξ kann nun der Vektor Φ der Störtermvarianzen ebenfalls konsistent geschätzt werden, wobei $\Phi' = (\sigma_u^2, \sigma_v^2)$. Diese Schätzung verwendet die Residuen e aus der Regression von (6- 8) mit dem TSLS-Schätzer von ξ . Mit z_r bzw. m_r als r -ter Reihe von Z bzw. W_2 und d_r als r -tem Diagonalelement von $W_2 W_2'$ gilt (vgl. dazu (6 - 2)):

$$E(\varepsilon_r^2) = \sigma_u^2 d_r + \sigma_v^2. \quad (6 - 10)$$

Entsprechend dieses Zusammenhangs werden nun die Residuen e , welche erwartungstreue Schätzer für die Störterme ε sind, auf die gesuchten Varianzen regressiert. Damit läßt sich ein OLS-Schätzer für Φ gewinnen:⁴³

$$\hat{\Phi} = (Q'Q)^{-1}Q'(e \bullet e). \quad (6 - 11)$$

Q ist eine $R \times 2$ -Matrix, deren linke Spalte die d_t enthält, während die rechte Spalte mit Einsen aufgefüllt wird. Also folgt (6 - 11) durch Anwendung der OLS-Vorgehensweise direkt aus (6 - 10). Die Autoren zeigen die gewünschte Konsistenzeigenschaft

$$\text{plim}(\hat{\Phi}) = \Phi. \quad (6 - 12)$$

Gemäß der letzten Zeile in (6 - 2) errechnet sich der konsistente Schätzer der Kovarianz durch

$$\hat{\Omega}_\varepsilon = \hat{\sigma}_u^2 W_2 W_2' + \hat{\sigma}_v^2 I_R. \quad (6 - 13)$$

Es sei angemerkt, daß im vorgestellten Prozedere nicht gesichert ist, ob beide geschätzten Varianzen größer Null sind. Um sicherzustellen, daß die geschätzte Kovarianzmatrix keine negative Varianzen enthält, sollten eine entsprechende Restriktion berücksichtigt werden.

Für den dritten Schritt, die eigentliche GMM-Schätzung, definieren die Autoren folgenden $(k+k^*) \times 1$ -Vektor f zur Festlegung der notwendigen Stichprobenmomente:

$$f = \frac{1}{R} H' \varepsilon. \quad (6 - 14)$$

Wegen der Unabhängigkeit des Fehlerterms mit den Matrizen W_1 und X , aus denen H konstruiert wurde, sind die Erwartungswerte der in f enthaltenen Stichprobendurchschnitte gleich Null. Weiter erhält man⁴⁴

$$\Omega_f = VC(f) = \frac{1}{R^2} (H' \Omega_\varepsilon H) \quad (6 - 15)$$

als Kovarianzmatrix von f . Die Kriteriumsfunktion q des vorgeschlagenen GMM-Schätzers lautet nun:

$$q(\xi) = \frac{1}{R} (f' \hat{\Omega}_f^{-1} f). \quad (6 - 16)$$

Nachdem die als Gewichtungsfaktor verwendete Kovarianzmatrix auf dem konsistenten zweiten Schätzschritt beruht, ist derjenige Parametervektor ξ_{GMM} der effiziente GMM-Schätzer, welcher die Kriteriumsfunktion q minimiert. Dieser lautet explizit

$$\begin{aligned} \hat{\xi}_{GMM} &= (Z' \hat{D} Z)^{-1} Z' \hat{D} y, \\ \text{mit } \hat{D} &= H (H' \hat{\Omega}_\varepsilon H)^{-1} H' \\ \text{und } \sqrt{R} (\hat{\xi}_{GMM} - \xi) &\xrightarrow{d} N \left(0, (Z' \hat{D} Z)^{-1} \right). \end{aligned} \quad (6 - 17)$$

⁴³ Das Hadamard-Produkt (elementweise Multiplikation zweier Vektoren gleicher Dimension) wird hier durch das Zeichen " \bullet " dargestellt.

⁴⁴ Zur Herleitung dieser und folgender Zusammenhänge vgl. Anhang A5.

Dieser im Vergleich zum ML-Verfahren rechentechnisch besser beherrschbare Schätzer ermöglicht mit seiner bekannter Kovarianzstruktur auch in kleinen Stichproben eine approximative Inferenz. Damit besteht für eine raumstrukturelle Betrachtung die Möglichkeit, direkt aus der Schätzung zu testen, ob überhaupt Autokorrelation im Raum vorliegt.

Die Autoren präsentieren zwar ein Anwendungsbeispiel für ihr Verfahren, legen aber keinen Vergleich vor, welche Resultate andere Schätzmethoden (z.B. ML-Verfahren) erzielt hätten. Die Überlegenheit gegenüber der OLS-Methode wurden für dieses Survey in einer Monte Carlo-Studie belegt (vgl. *Tabelle 3*): Die Verzerrung des Autokorrelationsparameters bei der Kleinstquadrateschätzung ist deutlich; insgesamt weist das GMM-Verfahren die wesentlich geringeren mittleren quadratischen Fehler auf.

Tabelle 3: Monte Carlo-Vergleich des GMM-Schätzers gegen OLS

	OLS			GMM		
	Abweichung	σ	MSE	Abweichung	σ	MSE
α	-.0181	.0024	3.3×10^{-4}	.0016	.0024	8.3×10^{-6}
β_1	-.0017	.0018	6.1×10^{-6}	-.0001	.0018	3.2×10^{-6}
β_2	-.0074	.0017	5.8×10^{-5}	-.0035	.0017	1.5×10^{-5}
ρ	.0155	.0006	2.4×10^{-4}	-.0007	.0006	0.8×10^{-6}

Verwendet wurde eine reihennormalisierte Nachbarschaftsmatrix westdeutscher Landkreise ($R = 327$; $\rho = .2$). Die angegebenen Werte für σ sind mit dem jeweiligen Verfahren geschätzt, wurden aber auf Plausibilität geprüft. Für nähere Angaben zur Durchführung der Studie vgl. Text in Unterabschnitt 4.1 bei *Tabelle 1*.

7. Zusammenfassung und Ausblick

In dieser Arbeit soll deutlich gemacht werden, daß sich das Konzept der Raumstrukturen in die herkömmliche Ökonometrie einzupassen vermag. Prozeßbeschreibungen, Test- und Schätzverfahren leiten sich aus dem bekannten Instrumentarium her. Es wird aufgezeigt, daß die Raumstruktur beträchtliche Einflüsse ausüben kann, weshalb bei regionalen Daten stets auf solche Zusammenhänge geachtet werden sollte..

Nachdem ökonomische Variablen über die Zeit selten konstant sind, wäre es darüber hinaus nur konsequent, auch Prozesse zu betrachten, die sich sowohl im Raum als auch in der Zeit bewegen. Überhaupt nimmt die Ausbreitung eines Phänomens über verschiedene Regionen hinweg meist eine gewisse Zeit in Anspruch; man denke nur an die Verzögerung, mit welcher modische Trends von einer Metropole in ländlichere Gebiete gelangen. Mit einer Paneldarstellung bieten sich auch neue Möglichkeiten, raumstrukturelle Modelle zu spezifizieren und zu schätzen. Anregungen in diese

Richtung liefert CASE (1991). In ihrer Formulierung stellen die einzelnen Regionen die Panelindividuen dar; für die jeweilige Region werden dann nicht Beobachtungen zu verschiedenen Zeitpunkten eingesetzt, sondern die Daten mehrerer Haushalte aus dieser Region. Dabei setzt sie neben einer Formulierung, welche die Raumstrukturen nicht berücksichtigt, sowohl ein Fixed Effects-Modell mit Regionendummies als auch ein Random Effects-Modell mit beiden Arten raumstruktureller Autokorrelation ein. Gemessen am Wert der Likelihoodfunktion verbessern bei ihr alle räumlichen Darstellungen, in unterschiedlichem Ausmaß, die Qualität der Schätzung deutlich.

Ein noch unbefriedigender Aspekt der Methodik besteht in der exogenen Vorgabe der Gewichtungsmatrix; eine Schätzung der relativen Abhängigkeiten wäre vorzuziehen. Allerdings sind die $[R*(R - 1)]$ Elemente nur mit wenigen Datensätzen bestimmbar, falls nicht enge Restriktionen getroffen werden.

Die zunehmenden Anstrengungen, Daten auf der Ebene kleinerer Gebietseinheiten erheben (z.B. die REGIO-Datenbank von EUROSTAT), vergrößert die Anwendungsgebiete der raumstrukturellen Ökonometrie. Ausbaufähig erscheint sie daneben insofern, als von ihr entwickelte Instrumentarien auch in anderen Gebieten einsetzbar sind, die Autokorrelationsstrukturen aufweisen. Solche Dependenzien dürften z.B. bei Finanzmärkten oder der Innovationsforschung zu finden sein.

Anhang

A1:

$$\begin{aligned}
 \text{Cov}(y_r, y_s) &= \gamma_{r,s} = E((y_r - \mu_r)(y_s - \mu_s)) = \\
 &= E\left(\left(\varepsilon_r + \sum_{r \neq i} m_{r,i} \varepsilon_i\right)\left(\varepsilon_s + \sum_{s \neq j} m_{s,j} \varepsilon_j\right)\right) = \\
 &= E\left(\varepsilon_r \varepsilon_s + \varepsilon_s \sum_{r \neq i} m_{r,i} \varepsilon_i + \varepsilon_r \sum_{s \neq j} m_{s,j} \varepsilon_j + \sum_{r \neq i} m_{r,i} \varepsilon_i \times \sum_{s \neq j} m_{s,j} \varepsilon_j\right) = \\
 &= E\left(0 + m_{r,s} \varepsilon_s^2 + m_{s,r} \varepsilon_r^2 + \sum_{i \neq r,s} m_{r,i} \cdot m_{s,i} \cdot \varepsilon_i^2\right) = \\
 &= \sigma^2 \left(m_{r,s} + m_{s,r} + \sum_{i \neq r,s} m_{r,i} \cdot m_{s,i}\right)
 \end{aligned}$$

A2

Wahrer Prozeß: $y = \rho W y + X_0 \beta_0 + \varepsilon$

$$E(\varepsilon \varepsilon') = \sigma^2 \cdot I_R$$

Bei einer OLS-Schätzung ohne Berücksichtigung raumstruktureller Autokorrelation

lautet der Schätzer hier:

$$\hat{\beta}_0 = (X_0' X_0)^{-1} X_0' y = (X_0' X_0)^{-1} X_0' (I - \rho W)^{-1} (X_0 \beta_0 + \varepsilon)$$

mit
$$\hat{\beta}_0 = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \quad \text{und} \quad X_0 = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{R1} & \dots & x_{Rk} \end{pmatrix} = (1_R \quad X)$$

Erwartungswert von $\hat{\beta}_0$:

$$\begin{aligned} E(\hat{\beta}_0) &= (X_0' X_0)^{-1} X_0' (I - \rho W)^{-1} X_0 \beta_0 + 0 \\ &= \begin{pmatrix} \frac{1}{R} 1_R' \left[I_R - (X(X'EX)^{-1} X'E) \right] \\ (X'EX)^{-1} X'E \end{pmatrix} (I_R - \rho W)^{-1} (1_R \quad X) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\ &\quad \text{mit } E = I_R - \frac{1}{R} 1_R 1_R' \end{aligned}$$

Zu $E(\hat{\alpha})$:

Da bei reihennormalisierter Gewichtungsmatrix $(I - \rho W)^{-1} 1_R = \frac{1}{1 - \rho} \cdot 1_R$, gilt für den

ersten Summanden:

$$\begin{aligned} &\frac{1}{R} 1_R' \left[I_R - X(X'EX)^{-1} X'E \right] (I - \rho W)^{-1} 1_R \cdot \alpha \\ &= \frac{1}{R} 1_R' I_R \frac{1}{1 - \rho} 1_R \alpha - \frac{1}{R} \frac{1}{1 - \rho} 1_R' X (X'EX)^{-1} X'E \underbrace{1_R \alpha}_{=0} = \frac{\alpha}{1 - \rho} \end{aligned}$$

Nachdem für $(I - \rho W)^{-1}$ auch $\left(I_R + \sum_{i=1}^{\infty} (\rho W)^i \right)$ geschrieben werden kann, wird der

zweite Summand:

$$\begin{aligned} &\frac{1}{R} 1_R' \left[I_R - X(X'EX)^{-1} X'E \right] (I - \rho W)^{-1} X \beta \\ &= \frac{1}{R} 1_R' I_R I_R X \beta + \frac{1}{R} 1_R' I_R \sum_i (\rho W)^i X \beta - \frac{1}{R} 1_R' X (X'EX)^{-1} X'E I_R X \beta \\ &\quad - \frac{1}{R} 1_R' X (X'EX)^{-1} X'E \sum_i (\rho W)^i X \beta \\ &= \frac{1}{R} 1_R' X \beta + \frac{1}{R} 1_R' \sum_i (\rho W)^i X \beta - \frac{1}{R} 1_R' X \beta - \frac{1}{R} 1_R' X (X'EX)^{-1} X'E \sum_i (\rho W)^i X \beta \end{aligned}$$

Somit gilt:
$$E(\hat{\alpha}) = \frac{\alpha}{1 - \rho} + T$$

mit
$$T = \frac{1}{R} 1_R' \sum_i (\rho W)^i X \beta - \frac{1}{R} 1_R' X (X'EX)^{-1} X'E \sum_i (\rho W)^i X \beta$$

Da die Regressoren in den einzelnen Regionen nicht korreliert sein sollen, wird der erste Summand von T etwa den Wert $\frac{\rho}{1-\rho} X\beta$ erreichen. Der zweite Summand setzt

sich zusammen aus $\bar{X} = \frac{1}{R} 1'_R X$, der Inversen der Regressoren-Kovarianzmatrix

$(X'EX)^{-1}$ und der Kovarianzmatrix zwischen den mittelwertbereinigten Regressoren bzw. deren raumstrukturell Verzögerten. Somit wird der durch $\bar{X}\beta$ ausgedrückte Mittelwert von y mit der von der Struktur der Gewichte abhängigen Eigenkorrelation jeder Region bewertet, welche durch die Zyklizität der raumstrukturellen Autokorrelation entsteht. Da die Hauptdiagonalen der Matrizen $(\rho W)^i$ erst für $i \geq 2$ ungleich

Null sind, kann dieser Summand den Wert $\frac{\rho^2}{1-\rho} \bar{X}\beta$ nicht überschreiten. Je höher die

Anzahl der Regionen, mit denen eine Region interagiert, umso deutlicher wird dieser Wert unterschritten. Eine gute Schätzung für Nachbarschaftsmatrizen liefert

$\frac{1}{w} \frac{\rho^2}{1-\rho} \bar{X}\beta$, wobei w für den um Eins erhöhten Median der Anzahl der Nachbarschaftsregionen steht.

Zu $E(\hat{\beta})$:

$$\begin{aligned} & (X'EX)^{-1} X'E(I-\rho W)^{-1}(1_R\alpha + X\beta) \\ &= \frac{1}{1-\rho} (X'EX)^{-1} X'E1_R\alpha + (X'EX)^{-1} X'E(I-\rho W)^{-1} X\beta \\ &= 0 + (X'EX)^{-1} X'EIX\beta + (X'EX)^{-1} X'E \sum_{i=1}^{\infty} (\rho W)^i X\beta \\ &= X\beta + (X'EX)^{-1} X'E \sum_i (\rho W)^i EX\beta_i \end{aligned}$$

Für den zweiten Summanden gelten die gleichen Überlegungen wie oben. Bei $\rho \neq 0$ wird $\hat{\beta}$ also, wenn auch geringfügig, absolut zu groß geschätzt.

A3

$$\begin{aligned} E((Wy)' \varepsilon) &= E(\varepsilon'(I-\rho W')^{-1} W' \varepsilon) = E\left(\text{tr}\left(\varepsilon'(I-\rho W')^{-1} W' \varepsilon\right)\right) = \\ & \text{tr}\left(E\left(\varepsilon\varepsilon'(I-\rho W')^{-1} W'\right)\right) = \sigma^2 \text{tr}\left((I-\rho W')^{-1} W'\right) \end{aligned}$$

Nachdem die Hauptdiagonale von W definitionsgemäß nur Nullen enthält, kann die Spur des Argumentes gleich Null werden, jedoch nur für $\rho = 0$.

q.e.d.

A4

$$\begin{aligned}\frac{\partial LL}{\partial \beta} &= \frac{\partial \left(-\frac{1}{2} \mu' \mu \right)}{\partial \beta} = -\frac{1}{2} \left(\mu' \frac{\partial \mu}{\partial \beta} + \frac{\partial \mu'}{\partial \beta} \mu \right) = \\ &= -\mu' \frac{\partial \mu}{\partial \beta} = \mu' \Omega^{-\frac{1}{2}} X.\end{aligned}$$

$$\begin{aligned}\frac{\partial LL}{\partial \rho} &= \frac{\partial \ln |I - \rho W|}{\partial \rho} + \frac{\partial \left(-\frac{1}{2} \mu' \mu \right)}{\partial \rho} = \\ &= \text{tr}(I - \rho W)^{-1} \frac{\partial (I - \rho W)}{\partial \rho} - \mu' \frac{\partial \mu}{\partial \rho} = \\ &= -\text{tr}(I - \rho W)^{-1} W + \mu' \Omega^{-\frac{1}{2}} W y.\end{aligned}$$

$$\begin{aligned}\frac{\partial LL}{\alpha_p} &= \frac{\partial \left(-\frac{1}{2} \ln |\Omega| \right)}{\alpha_p} + \frac{\partial \left(-\frac{1}{2} \mu' \mu \right)}{\partial \alpha_p} = \\ &= -\frac{1}{2} \text{tr} \Omega^{-1} \cdot \frac{\partial \Omega}{\partial \alpha_p} - \mu' \frac{\partial \mu}{\partial \alpha_p} = \\ &= -\frac{\text{tr} \Omega^{-1}}{2} \cdot H_p - \mu' \left(-\frac{1}{2} \right) \Omega^{-\frac{3}{2}} \cdot H_p \cdot ((I - \rho W)y - X\beta) = \\ &= -\frac{\text{tr} \Omega^{-1}}{2} \cdot H_p + \frac{1}{2} \mu' \Omega^{-\frac{3}{2}} \cdot H_p \cdot ((I - \rho W)y - X\beta).\end{aligned}$$

Um die Elemente der Informationsmatrix zu erhalten, müssen zunächst die Bedingungen erster Ordnung nochmals abgeleitet werden. Von den erhaltenen Ausdrücken ist dann der Erwartungswert zu nehmen.

Als Beispiel sei dies an $I_{\beta\rho}$ vorgeführt:

$$\begin{aligned}\frac{\partial LL}{\partial \beta \partial \rho} &= \frac{\partial}{\partial \rho} \left(\mu' \Omega^{-\frac{1}{2}} X \right) = \\ &= \frac{\partial}{\partial \rho} \left((I - \rho W)y - X\beta \right)' \Omega^{-\frac{1}{2}} \Omega^{-\frac{1}{2}} X = \\ &= \frac{\partial}{\partial \rho} \left(((I - \rho W)y)' \Omega^{-1} X \right) = \\ &= X \Omega^{-1} (-Wy) = -X \Omega^{-1} Wy.\end{aligned}$$

$$\begin{aligned}
I_{\beta\rho} &= -E\left(\frac{\partial LL}{\partial\beta\partial\rho}\right) = E(X'\Omega^{-1}Wy) = \\
&= X'\Omega^{-1}W \cdot E(y) = X'\Omega^{-1}W(I - \rho W)^{-1}X\beta.
\end{aligned}$$

q. e. d.

A5

$$\begin{aligned}
\Omega_\varepsilon &= E(\varepsilon\varepsilon') = E\left[(W_2u + v)(W_2u + v)'\right] = \\
&= E\left(W_2uu'W_2' + W_2uv' + vu'W_2 + vv'\right) = \\
&= M \cdot E(uu')W_2' + 0 + 0 + E(vv') = \\
&= \sigma_u^2 W_2 W_2' + \sigma_v^2 I_R.
\end{aligned}$$

q. e. d.

$$y = Z\xi + \varepsilon$$

$$\Rightarrow H'y = H'y - H'Z\xi$$

$$\begin{aligned}
VC(f) &= E(ff') = E\left(\frac{1}{R^2} \left[(H'\varepsilon)(H'\varepsilon)' \right]\right) = \\
&= \frac{1}{R^2} E(H'\varepsilon\varepsilon'H) = \frac{1}{R^2} H'\Omega_\varepsilon H.
\end{aligned}$$

$$VC(\hat{f}) = \frac{1}{R^2} H'\hat{\Omega}_\varepsilon H.$$

$$q(\xi) = \frac{1}{R} \underbrace{f' \left(\hat{VC}(f) \right)^{-1} f}_{\{\}} = \frac{1}{R} (H'y - H'Z\xi)' \{\}^{-1} (H'y - H'Z\xi);$$

$$\frac{\partial q}{\partial \xi} = -y'H\{\}^{-1}H'Z - y'H\{\}^{-1}H'Z + \xi'(Z'H\{\}^{-1}H'Z + Z'H\{\}^{-1}H'Z)' = 0$$

$$\Rightarrow 2\xi'(Z'H\{\}^{-1}H'Z) - 2y'H\{\}^{-1}H'Z = 0$$

$$\Rightarrow \xi' = (y'H\{\}^{-1}H'Z)(Z'H\{\}^{-1}H'Z)^{-1}$$

$$\xi_G = \left(Z'H \left(H'\hat{\Omega}_\varepsilon H \right)^{-1} H'Z \right)^{-1} Z'H \left(H'\hat{\Omega}_\varepsilon H \right)^{-1} H'y$$

$$\begin{aligned}
VC(\xi_G) &= \\
&\left(Z'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'Z \right)^{-1} Z'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'yy'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'Z \left(Z'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'Z \right)^{-1} = \\
&\left(Z'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'Z \right)^{-1} Z'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'Z \left(Z'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'Z \right)^{-1} = \\
&\left(Z'H \left(H' \hat{\Omega}_\epsilon H \right)^{-1} H'Z \right)^{-1} = (Z'\hat{D}Z)^{-1} .
\end{aligned}$$

q. e. d.

Literaturverzeichnis

- [1] ANSELIN, L. (1988): *Spatial Econometrics, Methods and Models*. Dordrecht: Kluwer.
- [2] ANSELIN, L. und R. FLORAX (Hrsg., 1995): *New Directions in Spatial Econometrics*. Berlin: Springer.
- [3] ANSELIN, L., A.K. BERA, R.FLORAX UND M.J.YOON (1996): „Simple Diagnostic Tests for Spatial Dependence“. *Regional Science and Urban Economics*, 26, 77-104.
- [4] ARORA,S. und M.BROWN (1977): „Alternative Approaches to Spatial Autocorrelation: An Improvement over Current Practice“. *International Regional Science Review*, 2, 67-78.
- [5] BESAG, J. (1974): „Spatial Interaction and the Statistical Analysis of Lattice Systems“. *Journal of the Royal Statistical Society*, ser. B., 36, 192-236.
- [6] BESAG, J. (1977): „Errors-in-Variables for Gaussian Lattice Schemes“. *Journal of the Royal Statistical Society*, ser. B, 39, 73-78.
- [7] BURRIDGE, P. (1980): „On the Cliff-Ord Test for Spatial Correlation“. *Journal of the Royal Statistical Society*, ser. B, 42, 107-108.
- [8] CASE, A.C. (1991): „Spatial Patterns in Household Demands“. *Econometrica*, 59, 953-965.
- [9] CASETTI, E. (1972): „Generating Models by the Spatial Expansion Method“. *Geographical Analysis*, 4, 81-91.
- [10] CASETTI, E. und J. POON (1995): „Econometric Models and Spatial Parametric Instability: Relevant Concepts and an Instability Index“. In: [2]ANSELIN, 301-321.

- [11] CLIFF, A.D. und J.K.ORD (1973): *Spatial Autocorrelation*. London: Pion.
- [12] CLIFF, A.D. und J.K.ORD (1981): *Spatial Processes*. London: Pion.
- [13] CLIFF, A.D., P.HAGGETT, J.K.ORD, K.A.BASSETT und R.DAVIES (1975): *Elements of Spatial Structure*. Cambridge: Cambridge Univ.
- [14] DUBIN, R.A. (1988): „Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms“. *Review of Economics and Statistics*, 70, 466-474.
- [15] GOURIEROUX, C. und A.MONFORT (1995): *Statistics and Econometric Models*. Cambridge: Cambridge Univ.
- [16] HAINING, R.P. (1995): „Data Problems in Spatial Econometric Modelling“. In: [2]ANSELIN, 156-171.
- [17] HOOPER, P. und G. HEWINGS (1981): „Some Properties of Space-Time Processes“. *Geographical Analysis*, 13, 203-223.
- [18] ISARD, W. (1969): *General Theory*. Cambridge: MIT.
- [19] KAU, J., C.Lee und C.Sirmans (1986): *Urban Econometrics: Model Developments and Empirical Results*. Greenwich, Conn.: JAI Press.
- [20] KELEJIAN, H.H. und I.R. PRUCHA (1995): „A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model“. Noch nicht veröffentlicht.
- [21] KELEJIAN, H.H. und D.P. ROBINSON (1993): „A Suggested Method of Estimation for Spatial Interdependent Models With Autocorrelated Errors, and an Application to a County Expenditure Model“. *Papers in Regional Science*, 3, 297-312.
- [22] KELEJIAN, H.H. und D.P. ROBINSON (1995): „Spatial Correlation: A Suggested Alternative to the Autoregressive Model“. In: [2]ANSELIN, 75-95.
- [23] KLOTZ, S. (1996): „Ökonometrische Modelle mit räumlichen Abhängigkeiten: Eine Übersicht“. *Unveröffentlichtes Manuskript der Universität Konstanz*.
- [24] MORAN, P.A.P. (1950): „Notes on Continuous Stochastic Processes“. *Econometrica*, 37, 17-23.
- [25] ORD, J.K. (1975): „Estimation Methods for Models of Spatial Interaction“. *Journal of the American Statistical Association*, 70, 120-126.
- [26] PAELINCK, J.H.P. und L.H.KLAASSEN (1979): *Spatial Econometrics*. Farnborough: Saxon House.
- [27] WHITTLE, P. (1954): „On Stationary Processes in the Plane“. *Biometrika*, 41, 434-449.