

Teil 1: Überblicksarbeiten

Strukturgleichungsmodelle der zweiten Generation: Eine Einführung

W. Schneider

Einleitung

Schenkt man den Ausführungen von Fornell (1982) Glauben, so greifen empirische Untersuchungen im Bereich der Sozialwissenschaften zusehends stärker auf eine neue, zweite Generation von multivariaten Analyseverfahren zurück. Komplexe Zusammenhangsstrukturen in Gegenstandsbereichen der Pädagogik, Psychologie, Politologie und Soziologie werden demnach immer häufiger anhand von Prozeduren analysiert, die (im Unterschied zu den Verfahren der sog. 'ersten Generation') den Forscher zur Modellierung abstrakter bzw. theoretischer Konstrukte sowie zur expliziten Berücksichtigung von Meßfehlern und damit zur Kombination von a-priori-Wissen und empirischen Daten befähigen.

Fornell bezieht sich auf eine Reihe von komplexen multivariaten Prozeduren, die in der Mehrzahl erst im vergangenen Jahrzehnt entwickelt wurden und allgemein unter dem Oberbegriff 'Kausalmodelle' bzw. 'Kovarianzstruktur-Analysen' subsummiert werden können. Eine gewisse terminologische Konfusion entstand dadurch, daß diese Ansätze verschiedentlich auch als Strukturgleichungsmodelle, simultane Gleichungssysteme oder Pfadanalysen mit latenten Variablen eingeführt worden sind.

Es soll hier nur kurz auf die Probleme eingegangen werden, die beispielsweise dem wissenschaftstheoretisch geschulten Leser bei der Konfrontation mit dem Begriff 'Kausalmodell' entstehen mögen. Bagozzi's (1980; Kap. 1) detaillierte Abhandlung der Kausalitätskonzepte von

Aristoteles über Hume, Mill bis hin zu Russels und Popper legt es nahe, den Terminus in einem Kontext zu vermeiden, bei dem es sicherlich nicht um die Überprüfung deterministischer Kausalgesetze oder -prinzipien, sondern allenfalls um die Annahme zeitlicher Sukzessionsgesetze gehen kann. Andererseits ist etwa Bentler (1980) der Auffassung, daß insofern kein Grund zur Kritik besteht, als dem 'Ursache'-Begriff hier keinerlei philosophische Bedeutung über die einer adhoc-Kennzeichnung für einen hypothetischen, unbeobachteten Prozeß hinaus zukommt; Begriffe wie 'Prozeß'- bzw. 'System'-Modell könnten den Terminus 'Kausal'-Modell also problemlos ersetzen.

Gewichtiger scheint das Problem zu sein, daß trotz der offenbar zunehmenden Popularität von Strukturmodellen mit latenten Variablen bislang kaum umfassende Einführungstexte zu diesen Techniken vorliegen (vgl. Bentler, 1980). Während diese Verfahren in den Lehrbüchern zur multivariaten Statistik bestenfalls am Rande erwähnt werden, setzt das Verständnis der Spezialliteratur in der Regel ausgeprägte statistische Vorkenntnisse, in jedem Fall aber Versiertheit im Umgang mit Matrixalgebra voraus. Möglicherweise ist die von Bielby und Hauser (1977) beklagte enorme zeitliche Verzögerung zwischen der Exposition der neuen Strukturgleichungsmodelle und ihrer angemessenen empirischen Anwendung auf die Sprachbarriere zwischen Methodologen/Statistikern und inhaltlich arbeitenden Sozialwissenschaftlern zurückzuführen. Das Dilemma wird offenkundig, wenn man bedenkt, daß die korrekte Handhabung von Strukturgleichungsmodellen einerseits zweifellos fortgeschrittene statistische Kenntnisse voraussetzt, andererseits viele im Prinzip nur unzureichend über die Methode informierten inhaltlich arbeitenden Forscher aufgrund der durchweg positiven Rezeption von Kausalmodellen in den Sozialwissenschaften dazu verleitet werden, letztere als 'Allheilmittel' bei der Auswer-

tung der eigenen Arbeiten einzusetzen (vgl. Horn und McArdle, 1980).

Ein Teilziel bei der Zusammenstellung des vorliegenden Bandes besteht darin, dem Leser eine Reihe von Anwendungsbeispielen im Bereich sozialwissenschaftlicher Forschung zu präsentieren, die effektive Einsatzmöglichkeiten von Strukturgleichungsmodellen dokumentieren. Im Hinblick auf den Vorkenntnisaspekt muß allerdings konstatiert werden, daß (vielleicht mit Ausnahme des Beitrages von Bentler) die stärker methodologisch orientierten Beiträge überwiegend technisch formuliert sind und somit fortgeschrittene statistische Kenntnisse voraussetzen. Aus Platzgründen kann hier keine detaillierte Einführung in die Analyse linearer Kausalmodelle gegeben werden. Zu diesem Zweck wird auf den Vorläuferband in dieser Reihe (Hodapp, 1984, insb. Kap. 5 u. 10) verwiesen. Um dem mit der Materie weniger vertrauten Leser eine ungefähre Vorstellung von der grundsätzlichen Thematik zu vermitteln, soll im folgenden eine knappe Darstellung der wesentlichen Merkmale von Strukturgleichungsmodellen sowie eine Grobskizzierung der Ähnlichkeiten bzw. Unterschiede in der Vorgehensweise der wohl am weitesten verbreiteten Techniken gegeben werden (ausführliche Abhandlungen finden sich weiterhin bei Bentler, 1980 oder Long, 1983).

In einem ersten Schritt wird die von Jöreskog und Sörbom (1978, 1981) entwickelte Pfadanalyse mit latenten Variablen LISREL (Analysis of Linear Structural Relationships) dargestellt. Im nächsten Schritt wird kritisch geprüft, wo die Vor- und Nachteile, Unterschiede und Gemeinsamkeiten von LISREL und PLS (Partial Least Squares) von Wold liegen. Auf die Technik EQS (Equations-based language) geht der Beitrag von Bentler (i.d.B.) näher ein. Der Vergleich bezieht sich dabei sowohl auf die theoretischen Zielsetzungen der einzelnen Prozeduren wie auch auf Daten, die in empirischen bzw. Simulationsstudien gewonnen wurden. Eine besondere Stellung nimmt

schließlich die Frage ein, wie in den einzelnen Techniken das Problem der Hypothesen- bzw. Modelltestung gelöst wird, die für die Bewertung der Prozeduren von großer Relevanz scheint.

Das allgemeine LISREL-Modell

Kausalmodelle mit latenten Variablen lassen sich prinzipiell dadurch charakterisieren, daß in ihnen mit der Faktorenanalyse und der Pfadanalyse zwei multivariate Analyseverfahren kombiniert sind. So setzt sich LISREL aus zwei Teilen zusammen: während das Meßmodell die Beziehung zwischen den beobachteten Variablen (Indikatoren) und den latenten Größen (hypothetischen Konstrukten) definiert, dient das Strukturgleichungsmodell dazu, die Abhängigkeiten zwischen den latenten Variablen festzulegen.¹⁾

Die folgende algebraische Formulierung des allgemeinen LISREL-Modells bezieht sich auf die neueste Version (LISREL VI) (vgl. Jöreskog und Sörbom, 1984). Auf Unterschiede zwischen dieser Version und der von mehreren Autoren dieses Bandes benutzten älteren Version (LISREL IV) wird im Text verwiesen. Die Spezifikation des Modells für mehrere Versuchspersonen, Gruppen- und Testzeitpunkte findet sich bei Jöreskog und Sörbom (i.d.B.).

Das LISREL-Modell berücksichtigt Zufallsvektoren $\eta' = (\eta_1, \eta_2, \dots, \eta_m)$ und $\xi' = (\xi_1, \xi_2, \dots, \xi_n)$ von latenten abhängigen bzw. endogenen und unabhängigen bzw. exogenen Variablen, wobei das Strukturgleichungsmodell durch die Gleichung

$$\eta = B\eta + \Gamma\xi + \zeta \quad (1)$$

1) Latente Variablen sind hier lediglich als Produkte einer Modell-Spezifikation anzusehen, die niemals über das hinausgehen können, was an Information in den beobachteten Indikatoren vorliegt (vgl. Sörbom, 1982).

definiert ist: B ($m \times m$) und Γ ($m \times n$) stellen Koeffizienten-Matrizen dar, während $\tilde{\zeta} = (\zeta_1, \zeta_2, \dots, \zeta_m)$ als Zufallsvektor der Residuen bzw. Gleichungsfehler aufzufassen ist. Die Elemente von B repräsentieren direkte kausale Effekte von latenten endogenen Variablen auf andere latente endogene Variablen (η - Variablen). Demgegenüber stellen die Elemente von Γ direkte kausale Effekte der latenten exogenen Variablen (ξ - Variablen) auf η - Variablen dar. Das eigentliche Kausalmodell wird also auf der Ebene der latenten Variablen spezifiziert: exogene latente Variablen sowie bestimmte endogene latente Variablen "erklären" andere endogene latente Größen, wobei der Spezifikations- bzw. Gleichungsfehler ζ die Güte der Prädiktion angibt (eine relativ einfach strukturierte graphische Darstellung für ein Kausalmodell mit latenten Variablen findet sich bei Bentler (i.d.B., Abb. 5); die V 's bezeichnen hier manifeste, die F 's latente Variablen). Es wird angenommen, daß $\tilde{\zeta}$ und $\tilde{\xi}$ unkorreliert sind und $I - B$ nichtsingulär ist.

Das Meßmodell gibt die Beziehung der beobachteten Variablen $y' = (y_1, y_2, \dots, y_p)$ und $x' = (x_1, x_2, \dots, x_q)$ zu den korrespondierenden η - und ξ - Variablen an:

$$\tilde{y} = \Lambda_y \tilde{\eta} + \tilde{\epsilon} \quad (2)$$

$$\tilde{x} = \Lambda_x \tilde{\xi} + \tilde{\delta} \quad (3)$$

Die Zufallsvektoren $\tilde{\epsilon}$ und $\tilde{\delta}$ repräsentieren Meßfehler in den \tilde{y} - bzw. \tilde{x} - Vektoren, und die Matrizen Λ_y ($p \times m$) und Λ_x ($q \times n$) sind als Regressionsmatrizen von \tilde{y} auf $\tilde{\eta}$ bzw. von \tilde{x} auf $\tilde{\xi}$ aufzufassen. Es wird angenommen, daß die Meßfehler nicht mit $\tilde{\eta}$, $\tilde{\xi}$ und $\tilde{\zeta}$ korrelieren, während sie untereinander durchaus korrelieren dürfen.

Gleichung (1) unterscheidet sich von der Formulierung in LISREL IV dadurch, daß die Matrix B nun Nullen in der Diagonalen aufweist und die direkten Effekte jeder η - Variablen auf andere η - Variablen mit dem korrekten Vorzeichen vorliegen (bei der LISREL IV - Version mußten die Vorzeichen in der Matrix B nachträglich umgekehrt werden).

Die Modellspezifikation erfordert eine sorgfältige Einstufung der Parameter. Als Grundlagen für die jeweilige Festlegung der Modellparameter dienen dabei das verfügbare theoretische Vorwissen, logische Kriterien, empirische Evidenz oder Implikationen des experimentellen Designs (vgl. Bagozzi, 1980). Es besteht prinzipiell die Möglichkeit, Modellparameter auf bestimmte Werte festzulegen (sie zu "fixen"), sie dadurch einzuschränken, daß sie zwar als unbekannt angenommen werden, aber identisch mit anderen Parametern sein sollen ("constrained parameters") oder aber sie durch das Programm schätzen zu lassen ("free parameters"). Die hauptsächlichen Problempunkte des sich an die Modellspezifikation anschließenden Berechnungsverfahrens betreffen die Identifikation, die Schätzung und die Testung des Kausalmodells.

(a) Das Identifikationsproblem: Nach Jöreskog (1977, 1978) hängt die Identifizierbarkeit eines Modells davon ab, ob die Schätzparameter durch Σ (die geschätzte Populations-Kovarianzmatrix) determiniert sind. Modellparameter sind - vereinfacht ausgedrückt - dann als identifiziert zu betrachten, wenn sie identische Werte in verschiedenen Kausalmodellen annehmen, die identische beobachtete Daten generieren (vgl. Bentler, 1980). Trifft dies für alle Parameter eines Modells zu, wird von globaler Identifikation gesprochen. Lokale Identifikation wird dann angenommen, wenn lediglich einige Parameter identifiziert werden können.

Da keine generellen notwendigen und hinreichenden Bedingungen für die Identifikation des allgemeinen LISREL-Modells verfügbar sind (vgl. Bagozzi, 1980; Jöreskog, 1981 b), muß die Identifiziertheit des Modells in jedem Einzelfall vom Forscher demonstriert werden. Mit Hinblick darauf, daß die Lösung des Identifikationsproblems aufgrund seiner Komplexität für viele Programm-Benutzer zu schwierig sein dürfte, verweist Jöreskog (1981b) darauf, daß der Identifikationsstatus des jeweiligen Kausal-

modells durch das LISREL-Programm geprüft wird (allerdings ist die programminterne Kontrolle nicht absolut zuverlässig). Obwohl LISREL nichtidentifizierte Modelle formal schätzen kann, wird dringend empfohlen, solche Modelle durch Einführung geeigneter Zusatzbedingungen (z.B. Parameter-Restriktionen) identifizierbar zu machen.

(b) Die Modellschätzung: Das Ziel bei der Generierung von Schätzwerten für alle unbekanntes Modellparameter besteht darin, die Populations-Kovarianzmatrix Σ an die beobachtete Kovarianzmatrix \tilde{S} anzupassen. LISREL IV greift dabei ausschließlich auf die Maximum-Likelihood-Schätzung zurück, bei der für die Bestimmung eines Parameters alle Informationen über die weiteren Modellparameter simultan genutzt werden, während bei LISREL VI zusätzlich die ungewichtete Kleinstquadratschätzung zur Verfügung steht. Bei der Maximum-Likelihood-Schätzung wird dabei eine aus der logarithmierten Likelihoodfunktion abgeleitete Kriteriumsfunktion F minimiert und damit die Dichtefunktion der Indikatoren maximiert. Die resultierenden Parameterschätzungen sind unter der Annahme effizient, daß sämtliche beobachteten Variablen des Modells multinormal verteilt sind.

(c) Die Testung von Modellen: Die Güte eines Modells läßt sich daran ablesen, wie "genau" es die vorgegebenen Daten (Kovarianz- oder Korrelationsmatrizen) reproduzieren kann. Für den Test der Modellangemessenheit wird die Likelihood-Quotienten-Technik herangezogen, wobei eine χ^2 -verteilte Prüfgröße (goodness-of-fit-Statistik) berechnet wird, deren Freiheitsgrade der Differenz zwischen der Anzahl der bekannten Datenelemente und unbekanntes Parametern entsprechen. Getestet wird hierbei gegen die allgemeine Hypothese, daß Σ irgendeine beliebige positiv-definite Matrix ist. Ein im Verhältnis zu den Freiheitsgraden hoher χ^2 -Wert deutet darauf hin, daß das spezifizierte Modell keine angemessene Reproduktion der Daten ermöglicht, die Nullhypothese also zurückzuweisen ist.

Bei LISREL VI kann die Anpassungsgüte von Modellen weiterhin über die quadrierte multiple Korrelation der Einzelvariablen, den Determinationskoeffizienten für die Kombination aller Variablen, den "Goodness-of-fit" (GFI)-Index und das "root mean square"-Residuum (RMR) beurteilt werden.

Der Determinationskoeffizient gibt an, wie gut alle beobachteten Variablen zusammengenommen als Meßinstrumente für alle latenten Variablen (wiederum zusammen betrachtet) fungieren. Diese Koeffizienten nehmen Werte zwischen Null und Eins an, wobei hohe Werte auf gute Modelle hindeuten.

Der GFI-Index erfaßt die Anpassungsgüte des Gesamtmodells. Er ist ein Maß für die relative Menge von Momenten, die insgesamt im Modell berücksichtigt werden, und sollte Werte zwischen Null und Eins annehmen (theoretisch sind jedoch auch negative Werte möglich). Im Unterschied zum χ^2 -Maß ist der GFI-Index von der Stichprobengröße unabhängig und auch recht robust gegenüber Abweichungen von der Multinormalverteilung.

Die Anpassungsgüte des Gesamtmodells kann weiterhin über den RMR-Index erfaßt werden, der als Maß für den Durchschnitt der Residual-Momente fungiert. Er kann nur in Relation zur Größe der beobachteten Varianzen und Kovarianzen in \hat{S} interpretiert werden und eignet sich besonders gut für den Fall, daß die Anpassungsgüte von zwei oder mehreren Modellen für einen identischen Datensatz verglichen werden soll.

Für den Fall einer Modellfalsifikation stehen verschiedene Möglichkeiten offen, um das inadäquate Modell zu modifizieren. Es kann z.B. auf diejenigen Parameter verzichtet werden, deren Schätzwerte im Vergleich zu ihrem Standardfehler äußerst gering ausfallen (vgl. Bentler, 1980). Weiterhin lassen sich über Inspektion der ersten Ableitungen der Anpassungsfunktion für die 'gefixten' (d. h. als bekannt gesetzten Parameter) sowie über eine Analyse der Residualmatrix ($\hat{S} - \hat{\Sigma}$) Wege aufzeigen, wie das jeweilige Modell gezielt verbessert werden kann.

LISREL VI bietet weiterhin die Möglichkeit, die lokale Anpassungsgüte detailliert über die Inspektion der normierten (standardisierten) Residual-Momente zu überprüfen. Ein im Programm verfügbares Q-Plot dieser Residuen bietet eine sehr brauchbare Zusammenfassung der Befunde.

Es kann dann auf einen Spezifikationsfehler geschlossen werden, wenn Residuen Werte größer als zwei annehmen; die korrespondierenden Indices i und j geben üblicherweise einen Hinweis darauf, wo der Fehler lokalisiert ist.

Schließlich sind in LISREL VI sog. "Modifikations-Indices" verfügbar, die das Verhältnis zwischen der quadrierten Ableitung erster Ordnung und der Ableitung zweiter Ordnung wiedergeben. Es läßt sich zeigen, daß der Modifikations-Index mit dem erwarteten Abfall im χ^2 -Wert für den Fall identisch ist, daß ein eingeschränkter ("constrained") Parameter freigesetzt wird. Er läßt sich deshalb über die χ^2 -Verteilung mit einem Freiheitsgrad beurteilen. Der mit dem höchsten Modifikations-Index korrespondierende festgelegte ("fixed") Parameter ist wahrscheinlich derjenige, der freigesetzt werden muß, um die Anpassungsgüte maximal zu verbessern.

Vorzüge und Problempunkte des LISREL-Ansatzes

Die bisher dargestellten Charakteristika von LISREL haben sicherlich einige der grundsätzlichen Vorteile dieses Analyse-Systems wie etwa seine Flexibilität und Generalität erkennen lassen. Entgegen früheren Annahmen (Goldberger, 1973) kann die Methode nicht nur im Bereich der Feld- bzw. nicht-experimentellen Forschung gewinnbringend eingesetzt werden; es hat sich vielmehr in neueren Arbeiten (vgl. Bagozzi, 1977, 1980; Möbus, i.d.B.; Sörbom und Jöreskog, i.d.B.) zeigen lassen, daß LISREL auch in der experimentellen Forschung spezifische Vorteile bietet.

Besondere Vorzüge offenbart LISREL bei der Analyse von Längsschnittdaten. Sieht man einmal davon ab, daß bei diesem Datentyp Modellspezifikationen generell einfacher sein mögen als bei Querschnittsdaten, die das a-priori-Wissen des Forschers stärker beanspruchen, treten hier einige Probleme auf, auf die LISREL wohl am flexibelsten reagieren kann (vgl. Schaie und Hertzog, 1982). Komplexe Designs zur Analyse von Entwicklungsveränderungen umfassen oftmals mehrere Meßzeitpunkte für unterschiedliche Personenstichproben (Kohorten). Mit LISREL lassen sich simultane Mehrgruppenvergleiche durchführen, aus denen ersichtlich wird, ob ein einziges hypothetisches Längs-

schnittmodell für unterschiedliche Gruppen gilt. Die Überprüfung der Generalisierbarkeit eines Modells für mehrere Gruppen bzw. Kohorten kann dabei unterschiedlich rigoros vorgenommen werden: der strengste Test verlangt die Invarianz aller Schätzparameter von Meß- und Strukturmodell für alle erfaßten Gruppen, während ein "weicherer" Test etwa so aussehen könnte, daß in allen Gruppen die Annahme eines gleichartigen Pfadmusters für Meß- wie Strukturmodell geprüft wird, wobei die numerischen Ausprägungen der Schätzparameter variieren dürfen.

Es ist weiterhin möglich, Mittelwertsveränderungen in den latenten Variablen über die Zeit hinweg zu erfassen, wenn Daten zu mehreren Gruppen vorliegen. Unterschiede in den Mittelwertsstrukturen mehrerer Gruppen sind jedoch nicht absolut zu sehen, sondern als Abweichungen zu den Veränderungen in den latenten Variablen einer beliebigen Gruppe (Kohorte), deren Mittelwerte in den latenten Variablen auf Null festgelegt wurden.

Bei Längsschnittdaten ist die Annahme unkorrelierter Meßfehler prinzipiell fragwürdig, weil beispielsweise Indikatoren des ersten Meßzeitpunkts mit entsprechenden Variablen der zweiten Welle auch dann noch korrelieren können, wenn man die auf die Stabilität der Konstrukte zurückzuführende Korrelation herauspartialisiert (vgl. Möbus, im Druck; Sörbom, 1982). Diese etwa als Retest- bzw. als Erinnerungseffekt interpretierbare Restkorrelation kann in LISREL explizit dadurch berücksichtigt werden, daß man entsprechende Meßfehler korrelieren läßt.

Probleme von LISREL betreffen die Frage der Hypothesentestung. Angesichts der Güte-Standards in der experimentellen Forschung (vgl. etwa Hager und Westermann, 1983) ist es unerläßlich, daß klare Kriterien für die Modellbewertung vorliegen. Die zur Testung von LISREL-Modellen herangezogene χ^2 -Prüfgröße hat den Nachteil, daß die Stichprobengröße direkt proportional in ihre Berechnung eingeht. Für den Fall großer Stichproben ergibt sich mit

hoher Wahrscheinlichkeit ein numerisch großer χ^2 -Wert, der die Verwerfung des theoretischen Modells nahelegt, obwohl die numerischen Differenzen zwischen der Ausgangskorrelationsmatrix \tilde{S} und der geschätzten Matrix $\tilde{\Sigma}$ minimal sein können. Computersimulationsstudien (vgl. Bearden, Sharma und Teel, 1982; Boomsma, 1982) haben weiterhin gezeigt, daß die χ^2 -Statistik auch bei kleinen Stichproben und komplexeren Modellen zu große Werte annimmt und damit das hypothetische Modell oftmals ungerechtfertigt zurückgewiesen wird. Die von Jöreskog und Sörbom (1978, 1982) vorgeschlagene Möglichkeit, statt der Analyse der absoluten χ^2 -Werte hierarchisch geschachtelte Modelle im Hinblick auf ihre Datenanpassung zu vergleichen und damit lediglich die Differenzen in den jeweiligen χ^2 -Werten (die bei großem N wiederum annähernd χ^2 -verteilt sind) zur Bewertung heranzuziehen, läßt sich allerdings nur im Sinne einer explorativen Datenanalyse rechtfertigen (s. Möbus, im Druck). Da die geschachtelten Modelle meist am gleichen Datensatz getestet werden, kann hier nicht von einer statistischen Hypothesenprüfung im strengen Sinne gesprochen werden.

Alternative Prüfgrößen wie der von Bentler und Bonett (1980) vorgeschlagene "inkrementelle Fit-Index" (vgl. auch die Anwendung bei Sieber, i.d.B.) oder der normalisierte "Residual-Index" (der die Differenzen zwischen den Input-Kovarianzen und denen des geschätzten Modells vergleicht) reagieren ebenfalls sensibel auf Veränderungen der Stichprobengröße (vgl. Bearden et al., 1982), geben aber für mittelgroße Stichproben ($N \geq 100$) deutliche Hinweise darauf, daß ein Modell nicht an die Daten angepaßt ist.

Abzuwarten bleibt, ob der kürzlich unternommene Versuch von Satorra und Saris (1982a) geeignet ist, das Dilemma zu lösen: die Autoren entwickelten ein Verfahren zur Berechnung der Teststärke des χ^2 -Prüftests, über das sowohl die Teststärke als auch die Effekte von unterschiedlichen Reliabilitäten der Indikatoren sowie unterschied-

lichen Stichprobengrößen auf die Teststärke relativ genau bestimmt werden können (vgl. Satorra und Saris, 1982b).

Zusammenfassend bleibt festzuhalten, daß das Problem der Hypothesentestung anhand von LISREL noch nicht befriedigend gelöst ist.

Zum Vergleich von LISREL und PLS

Beide Ansätze sind darin ähnlich, daß sie Struktur- und Meßmodelle (bei PLS innere und äußere Modelle genannt) spezifizieren. Die Beziehungen zwischen Indikatoren und latenten Variablen können bei PLS jedoch zusätzlich zu der bei LISREL standardmäßigen "Auswärtsrichtung" (die Pfeile führen von den Konstrukten zu den beobachteten Variablen) andersherum definiert werden: "einwärts" gerichtete Beziehungen deuten an, daß die Indikatoren als "Ursachen" der latenten Variablen aufzufassen sind. Stellt sich eine latente Variable später als äußerst informativ heraus, wird empfohlen, die Indikatoren in der Folge als auswärts gerichtet oder "reflektiv" zu spezifizieren (vgl. Lohmöller, 1981a; Wold, 1982).

Der entscheidende Unterschied zwischen PLS und LISREL ist im Schätzalgorithmus und damit in der stochastischen Spezifikation des Modells zu sehen (vgl. Knepel, 1981; Lohmöller, 1981b, 1982). Während bei LISREL ein einziges übergeordnetes Optimierungskriterium (die Kriteriums-funktion F , s.o.) verwendet wird, legt PLS ein Kleinstquadratkriterium zugrunde: es werden die Residualvarianzen der einzelnen Regressionsgleichungen im Meß- und Strukturmodell iterativ minimiert. Dies kann auch systematische Unterschiede zwischen LISREL- und PLS-Lösungen erklären helfen (vgl. Lohmöller, 1982), die darin bestehen, daß die Residualvarianzen bei PLS kleiner ausfallen als bei LISREL und die Ladungen der latenten Variablen bei PLS folglich in der Regel größer sind. Dennoch soll-

ten die numerischen Unterschiede in den Parameter-Schätzungen für ein-und-dieselbe Variable zwischen PLS und LISREL nicht substantiell sein (Jöreskog und Wold, 1982) und sind es (für ausreichend große Stichproben) auch nicht, wie die Simulationsstudien von Areskoug (1982) bzw. Fornell und Bookstein (1982) belegen.

Theoretisch besteht ein wesentlicher Vorteil des LISREL-Ansatzes darin, daß für den Fall großer Stichproben und des Zutreffens der Normalitätsannahme die Parameterschätzungen effizient und optimal sind, während die Genauigkeit der PLS-Schätzungen aufgrund der weniger restriktiven Annahmen (z.B. keine Annahmen über die Variablen-Verteilung) geringer ausfällt und weiterhin durch die Konzeption der latenten Variablen als standardisierte Linearkombination der ihnen zugeordneten Indikatoren beeinträchtigt ist (s. Areskoug, 1982). Andererseits haben Computersimulationen (Hui und Wold, 1982) ergeben, daß PLS-Schätzungen für den Fall großer Stichproben sowie vieler Indikatoren pro Konstrukt den "wahren" Parameterwerten sehr nahe kommen.

Die hauptsächlichlichen Unterschiede zwischen PLS und LISREL lassen sich zusammenfassend so charakterisieren, daß PLS daten- und LISREL modellorientiert, PLS eher explorativ und LISREL eher konfirmatorisch arbeitet. LISREL sollte dann eingesetzt werden, wenn das spezifizierte Modell erprobt und "korrekt" ist sowie die erfaßten Merkmale intervallskaliert und multivariat normalverteilt sind. Wenn Zweifel am Vorliegen dieser Bedingungen bestehen, dürfte PLS das zunächst angemessenere Verfahren sein. Der Benutzer verschafft sich damit den Vorteil, keine Verteilungsannahmen treffen zu müssen, sich keine Identifikationsprobleme und auch keine negativen Varianzen (Heywood-Fälle) einzuhandeln (dies aufgrund der spezifischen Konzeption der latenten Variablen), nahezu immer konvergierende Lösungen zu erhalten und auch aufgrund der bedeutend niedrigeren Rechenzeit von PLS auch sehr komplexe Modelle schätzen zu können. Er nimmt dabei in Kauf, daß die geschätzten Modelle nicht im interferenzstatistischen Sinne testbar sind, da keine Verteilungsannahmen getroffen werden. Es sind jedoch durchaus Modellbewertungen

möglich, die über Plausibilitätstests, Kreuzvalidierungstechniken sowie Analysen der Anpassungsgüte erfolgen (vgl. Knepel, 1981; Lohmöller, 1981a).

Jöreskog und Wold (1982) ist zuzustimmen, wenn sie LISREL und PLS als komplementäre Verfahren betrachten. Die computertechnische Realisierung dieses Gedankens findet sich bei Lohmöllers (1981a) LVPLS-Programm, das die Übersetzung von PLS- in LISREL-Modelle explizit vorsieht.