

# 1 Einleitung

To err is human, to forgive divine,  
but to include errors into your design  
is statistical.

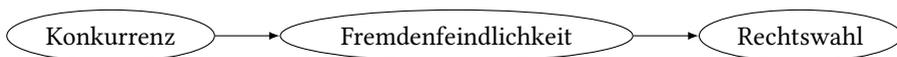
---

Leslie Kish

## 1.1 Wieso, weshalb, warum? Strukturgleichungsmodelle in der Politikwissenschaft

Empirische politikwissenschaftliche Theorien bestehen aus Annahmen über die Realität (Hypothesen), deren Gültigkeit zumindest prinzipiell überprüft werden kann, wie diese drei bekannten Beispiele zeigen: Arbeitslosigkeit verursacht politische Apathie (Jahoda, Lazarsfeld und Zeisel, 1975), Demokratien führen untereinander seltener Angriffskriege als Nicht-Demokratien (u. a. Doyle, 1983), Konsensdemokratien sind erfolgreicher als Mehrheitsdemokratien (Lijphart, 1999). Das bei weitem am häufigsten verwendete Verfahren zur Prüfung solcher Hypothesen ist die (multiple) lineare Regression, die eine *abhängige* Variable  $y$  (z. B. Zahl der Angriffskriege) zu einer oder mehreren *unabhängigen* Variablen  $x_1, x_2, x_3 \dots$  (Regimetyp, Vorhandensein einer gemeinsamen Grenze, Größe und Ausrüstung der jeweiligen Armeen etc.) in Beziehung setzt. Für jede unabhängige Variable wird ein Koeffizient geschätzt, der einen Eindruck davon vermittelt, wie stark und in welche Richtung sich Veränderungen in der jeweiligen unabhängigen Variablen auf die abhängige Variable auswirken.

Das Verfahren der linearen Regression ist robust, einfach anzuwenden und zu interpretieren. Es bildet den Grundstein für eine Reihe von weiteren Regressionsverfahren (etwa logistische oder Poisson-Regression), die z. B. der besonderen Verteilung einer gegebenen abhängigen Variablen Rechnung tragen. Etwaige Beziehungen *zwischen* den unabhängigen Variablen (beispielsweise könnten Demokratien *ceteris paribus*



**Abbildung 1.1:** Eine einfache Kausalkette

über kleinere Armeen verfügen) verursachen in der Regel keine Probleme,<sup>1</sup> werden vom Verfahren aber als gegeben betrachtet und nicht gesondert modelliert.

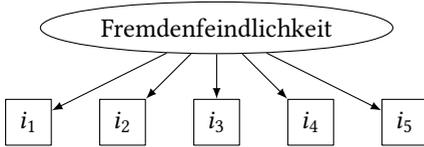
Letzteres ist dann von Nachteil, wenn die jeweilige Theorie eine ganze Reihe von miteinander verbundenen Hypothesen enthält, die gemeinsam getestet werden sollen, wie ein viertes Beispiel zeigt. So ist es erstens plausibel anzunehmen, dass Menschen, die auf dem Arbeitsmarkt mit Migranten konkurrieren, gegenüber dieser Bevölkerungsgruppe negative Einstellungen entwickeln. Zweitens dürften solche negativen Einstellungen ein wesentlicher Faktor für die Wahlentscheidung zugunsten fremdenfeindlicher Parteien sein. Ein und dieselbe Einstellung (Fremdenfeindlichkeit) wird also einmal als abhängige und einmal als unabhängige Variable betrachtet (vgl. Abbildung 1.1). Betrachtet man über diese drei Variablen hinaus beispielsweise noch die Ursachen einer Konkurrenz auf dem Arbeitsmarkt, die weiteren Folgen fremdenfeindlicher Einstellungen sowie die Folgen einer rechten Wahlentscheidung ergibt sich rasch ein ganzes „Netz“ (Falter, 1977b) von Kausalbeziehungen.

Aus dem Wunsch, ein solches Netz durch ein komplexeres mathematisches Modell abzubilden (was keineswegs immer notwendig oder sinnvoll ist), entstand das Verfahren der klassischen Pfadanalyse, das von dem amerikanischen Genetiker Sewall Wright in den 1920er und 1930er Jahren entwickelt und seit den 1960er Jahren für die Soziologie und andere Sozialwissenschaften nutzbar gemacht wurde (Kaplan, 2009, S. 1-6; Raftery, 2001).

Die Anwendung der Pfadanalyse setzt allerdings voraus, dass die Variablen, die die Knoten des kausalen Netzes bilden, direkt beobachtet werden können. In der Politikwissenschaft ist dies häufig *nicht* der Fall: Konzepte wie „Fremdenfeindlichkeit“ oder „Demokratie“ müssen erst mit Hilfe einer Messanweisung operationalisiert werden. Eine solche Messanweisung setzt eine nicht-beobachtbare Variable (auch: latente Variable, Konstrukt, Faktor) zu einer oder mehreren beobachtbaren Variablen (auch: manifeste Variable, Indikator) in Beziehung. Dabei ist es nicht notwendig, dass ein Indikator die latente Variable perfekt abbildet oder dass alle Indikatoren in gleicher Weise geeignet sind, eine latente Variable zu messen. Vielmehr ist die Modellierung der unvermeidlichen MESSFEHLER Bestandteil des Verfahrens (siehe Abschnitt 2.3).

Bereits zu Beginn des 20. Jahrhunderts wurden von Charles Spearman und Karl Pearson verschiedene Verfahren zur Analyse des Zusammenhangs zwischen manifesten und latenten Variablen entwickelt. Diese Verfahren werden unter dem Namen

<sup>1</sup>Vorausgesetzt wird lediglich, dass zwischen den unabhängigen Variablen keine (nahezu) *perfekten* linearen Beziehungen (KOLLINEARITÄT) bestehen.



**Abbildung 1.2:** Messung einer latenten Variablen durch mehrere Indikatoren

FAKTORENANALYSE zusammengefasst und fanden vor allem in der Psychologie weite Verbreitung. Sie werden dort primär zur Identifikation von Intelligenz- und anderen Persönlichkeitsfaktoren sowie zur Auswertung von Einstellungstests eingesetzt.

Solche komplexen Konzepte werden mit guten Gründen durch jeweils mehrere Indikatoren abgebildet. Beispielsweise dürfte es sehr schwierig sein, ein einzelnes Fragebogen-Item zu finden, das eine komplexe und zugleich relativ diffuse Einstellung wie Fremdenfeindlichkeit in perfekter Weise misst. Stattdessen wird man eine große Zahl von Items ( $i_1, i_2, i_3, \dots$ ) entwickeln, die sich auf diesen Gegenstandsbe- reich beziehen, und diese einer Gruppe von Befragungspersonen vorlegen. Erfah- rungsgemäß werden unterschiedliche Frageformulierungen in ähnlicher, aber nicht identischer Weise beantwortet, weil jedes einzelne Item bei den Befragten spezifische Reaktionen auslöst. Deshalb wird man anschließend versuchen, mit Hilfe faktorana- lytischer Verfahren aus den Antwortmustern einen gemeinsamen Faktor, nämlich die Einstellung „Fremdenfeindlichkeit“ zu extrahieren, die einen mehr oder minder großen Anteil des Antwortverhaltens erklären kann (Abbildung 1.2).<sup>2</sup>

Von einem politikwissenschaftlichen Standpunkt erscheint es naheliegend, Faktor- und Pfadanalyse in einem gemeinsamen mathematischen Modell zu kombinieren, weil ein solches Konstrukt die Struktur vieler politikwissenschaftlicher Theorien adäquat abbildet. Tatsächlich wurden solche Strukturgleichungsmodelle<sup>3</sup> jedoch erst Ende der 1960er/Anfang der 1970er Jahre entwickelt. Zum einen fehlten zuvor die mathematischen Grundlagen, zum anderen setzte der praktische Einsatz dieser neuen Modelle den Zugang zu leistungsfähigen Computern voraus, die in der Lage waren, die Modellparameter zu schätzen. Die ersten Programme, mit denen dies möglich war, liefen auf den großen Computeranlagen der Universitätsrechenzentren und waren nur durch Spezialisten zu bedienen (siehe dazu <http://www.ssicentral.com/lisrel/history.html>).

Erst die Verbreitung von leistungsstarken PCs in den 1990er Jahren führte zur Entwicklung benutzerfreundlicher(er) Programme. Dementsprechend sind Struktur- gleichungsmodelle für die Politikwissenschaft im Gegensatz zu anderen Disziplinen wie der Psychologie oder der empirischen Bildungsforschung immer noch eine relativ neue Methode. Zudem ist die Grundlagenforschung in diesem Bereich noch

<sup>2</sup>Die Konventionen für die Gestaltung von Pfaddiagrammen werden in Abschnitt 2.4.2 vorgestellt.

<sup>3</sup>Strukturgleichungsmodelle verbinden ein „strukturelles“ Modell – die Beziehungen zwischen den endogenen Variablen – mit einem oder mehreren Modellen zur Messung dieser Variablen.

keineswegs abgeschlossen: Nach wie vor werden die zugrundeliegenden mathematischen Modelle modifiziert und erweitert. Ziel dieses Buches kann es deshalb nur sein, einen ersten, anwendungsbezogenen Überblick über die Möglichkeiten des Verfahrens zu vermitteln. Der mathematische Apparat ist auf das absolute Minimum reduziert, auf Herleitungen und Beweise wird vollständig verzichtet.

Vorausgesetzt werden lediglich mathematische Grundkenntnisse aus der Sekundarstufe I. Eine kompakte, auf sozialwissenschaftliche Anwendungen bezogene Darstellung dieser Grundlagen bietet Hagle (1995). Gill (2006) vermittelt ebenfalls anwendungsbezogen zusätzliches Hintergrundwissen. Geeignete Einführungen in die sozialwissenschaftliche Statistik sind Gehring und Weins (2009) sowie Kühnel und Krebs (2012). Weitere Hinweise auf weiterführende Literatur finden sich in Kapitel 5.

In der Forschungspraxis werden heute alle Berechnungen von Computerprogrammen vorgenommen, ohne dass der Anwender im einzelnen verstehen muss (oder kann) was innerhalb des Programms vor sich geht. Dennoch können (und sollen!) die einfachen Rechenbeispiele im Text mit Papier und Bleistift nachvollzogen werden.<sup>4</sup> Nur auf diese Weise lässt sich ein tieferes Verständnis des Verfahrens erreichen, was letztlich zu einem souveräneren Umgang mit den Ergebnissen der computergestützten Analysen führt.

## 1.2 Aufbau des Buches

Das vorliegende Buch gliedert sich in vier große Teile: Auf diese Einleitung folgt Kapitel 2, das sich mit Grundlagen beschäftigt. Leser, die bereits über entsprechendes Wissen verfügen oder sich einen ersten Überblick verschaffen wollen, können dieses Kapitel zunächst überspringen. Dies gilt insbesondere für die mit einem \* markierten Abschnitte 2.1 und 2.5, die primär zum Nachschlagen gedacht sind. Kapitel 3 ist ganz den Anwendungsbeispielen gewidmet und zeigt, wie Strukturgleichungsmodelle in der Praxis der politikwissenschaftlichen Forschung nutzbringend eingesetzt werden können. Im Anschluss daran beschäftigt sich Kapitel 4 mit fortgeschrittenen Themen und spezielleren Fragestellungen. Auch hier steht aber die Anwendungsorientierung im Vordergrund. Zu allen Beispielen aus den Kapiteln 3 und 4 können die Datensätze sowie der nötige Programmcode für Stata, LISREL und Mplus von der website zum Buch heruntergeladen werden: [www.kai-arzheimer.com/beispiele-sem/](http://www.kai-arzheimer.com/beispiele-sem/).

---

<sup>4</sup>Mathematikprogramme sind eine komfortable Alternative zu Papier und Bleistift. Maxima ist kostenlos erhältlich und verfügt über eine intuitive graphische Oberfläche (<http://maxima.sourceforge.net/>).

| Name    | Kleinbuchstabe | Großbuchstabe |
|---------|----------------|---------------|
| Alpha   | $\alpha$       | A             |
| Beta    | $\beta$        | B             |
| Gamma   | $\gamma$       | $\Gamma$      |
| Delta   | $\delta$       | $\Delta$      |
| Epsilon | $\epsilon$     | E             |
| Zeta    | $\zeta$        | Z             |
| Eta     | $\eta$         | H             |
| Theta   | $\theta$       | $\Theta$      |
| Iota    | $\iota$        | I             |
| Kappa   | $\kappa$       | K             |
| Lambda  | $\lambda$      | $\Lambda$     |
| My      | $\mu$          | M             |
| Ny      | $\nu$          | N             |
| Xi      | $\xi$          | $\Xi$         |
| Omikron | $o$            | O             |
| Pi      | $\pi$          | $\Pi$         |
| Rho     | $\rho$         | P             |
| Sigma   | $\sigma$       | $\Sigma$      |
| Tau     | $\tau$         | T             |
| Ypsilon | $\upsilon$     | $\Upsilon$    |
| Phi     | $\phi$         | $\Phi$        |
| Chi     | $\chi$         | X             |
| Psi     | $\psi$         | $\Psi$        |
| Omega   | $\omega$       | $\Omega$      |

**Tabelle 1.1:** Das griechische Alphabet von A bis  $\Omega$

### 1.3 Konventionen

Skalare Variablen werden in diesem Buch durch lateinische Kleinbuchstaben in Kursivdruck symbolisiert:  $x, y, z, \dots$ . Als Platzhalter für skalare Parameter<sup>5</sup> werden griechische Kleinbuchstaben verwendet:  $\alpha, \beta, \gamma, \dots$  (siehe dazu Tabelle 1.1). Für Vektoren und Matrizen (siehe Abschnitt 2.1) werden fettgedruckte Klein- bzw. Großbuchstaben benutzt ( $\mathbf{a}, \mathbf{\Omega}$ ). Auch hier werden Variablen durch lateinische, zu schätzende Parameter hingegen durch griechische Buchstaben symbolisiert. Namen von Programmen werden in serifenloser Schrift gedruckt (LISREL), für Webadressen und Programmco-

<sup>5</sup>Für Variablen liegen empirische Werte vor. Parameter hingegen sind unbekannte Werte, die auf Grundlage der empirischen Daten geschätzt werden müssen. Siehe dazu Abschnitt 2.6.

de wird Schreibmaschinenschrift verwendet (<http://www.kai-arzheimer.com/>). Wichtige Fachbegriffe werden im Glossar am Anfang des Buches erläutert. Bei ihrer ersten Nennung erscheinen sie als Hinweis darauf in KAPITÄLCHEN. Als Dezimaltrennzeichen wird, wie im englischen Sprachraum üblich, ein Punkt verwendet, da die Mehrzahl der Statistikprogramme, die zur Schätzung von Strukturgleichungsmodellen in Frage kommen, ihre Ausgaben nach dieser Konvention formatieren:  $\frac{3}{4} = 0.75$ .

Größere Abschnitte von Code, die in ein Statistikprogramm einzugeben sind, werden vom Fließtext durch horizontale Linien abgesetzt. Die Zeilennummern am linken Rand dienen der besseren Orientierung und werden nicht mit eingegeben. Alle Eingaben sind in Schreibmaschinenschrift formatiert. Schlüsselwörter (Kommandos) der betreffenden Sprache sind zusätzlich durch eine Unterstreichung hervorgehoben. Kommentare, die dem besseren Verständnis dienen, erscheinen in Schreibmaschinenschrift und mit der Kommentarsyntax der jeweiligen Sprache.

## 1.4 Software und Internetquellen

Für viele Sozialwissenschaftler ist das Programm LISREL noch immer ein Synonym für die Anwendung von Strukturgleichungsmodellen. Tatsächlich waren die zu Beginn der 1970er Jahre von Karl Jöreskog zusammen mit Dag Sörbom entwickelten Programme LISREL und ACOVFSF bzw. COFAMM die ersten Werkzeuge, mit denen es überhaupt möglich war, Strukturgleichungsmodelle zu schätzen, ohne selbst programmieren zu müssen. Allerdings mussten die Modelle in einer Matrix-Notation definiert werden, was den Kreis der potentiellen Benutzer stark einschränkte. Erst Mitte der 1990er Jahre wurde LISREL um die alternative Schnittstelle SIMPLIS ergänzt, die es ermöglicht, ein Modell durch eine Reihe von Gleichungen zu definieren, die den Pfeilen eines Kausaldiagramms entsprechen. Wenig später erschienen als Reaktion auf modernere Konkurrenzprogramme Versionen von LISREL für das Windows-Betriebssystem, die es dem Benutzer gestatteten, mit Hilfe der Maus ein Kausalmodell zusammenzuklicken, dessen Aufbau interaktiv verändert und in andere Windows-Programme importiert werden konnte. Aktuell wird LISREL in der Version 9.1 vertrieben. Der Distributor bietet sowohl eine vollständige Testversion (gültig für 15 Tage) als auch eine Studentenversion zum kostenlosen Download an (<http://www.ssicentral.com/lisrel/downloads.html>). Letztere weist einige Einschränkungen auf (u. a. was die Zahl der beobachteten Variablen betrifft), reicht aber aus, um viele Beispiele nachvollziehen zu können. LISREL ist derzeit ausschließlich für Windows erhältlich, kann aber möglicherweise innerhalb von Emulations- und Virtualisierungsumgebungen wie Wine, Boot Camp oder XEN auch unter anderen Betriebssystemen genutzt werden.

Mitte der 1980er Jahre stellte Peter Bentler, der ähnlich wie Jöreskog führend an der Entwicklung des Feldes beteiligt war, mit EQS ein weiteres kommerzielles Programm zur Schätzung von Strukturgleichungsmodellen vor. Obwohl EQS auf einem etwas anderen Ansatz basiert als LISREL, sind beide Programme im wesentlichen äquivalent. Auch EQS verfügt über eine grafische Schnittstelle. Allerdings ist EQS zumindest im deutschen Sprachraum weit weniger verbreitet. Nach dem die Entwicklung von EQS etwas ins Stocken geraten war, ist 2013 eine neue Version (6.2) erschienen (<http://www.mvsoft.com/eqs60.htm>). Eine vergünstigte Studierendenversion ist im Online-Shop erhältlich, eine kostenlose Testversion muss per Email angefordert werden. Auch EQS ist derzeit ausschließlich für Windows erhältlich.<sup>6</sup>

Ende der 1980er Jahre erschien mit dem von James Arbuckle entwickelten AMOS ein weiteres kommerzielles Konkurrenzprogramm zu LISREL. Von den älteren Programmen unterschied sich AMOS neben einigen eher technischen Besonderheiten vor allem durch die konsequente Nutzung der grafischen Benutzeroberfläche. Bis heute ist AMOS ein sehr benutzerfreundliches Programm, das auch nach der Übernahme durch SPSS weiterentwickelt wird. Es bietet eine Vielzahl innovativer Optionen und erfreut sich im deutschsprachigen akademischen Bereich einiger Beliebtheit. Aktuell vertreibt SPSS/IBM die Version 22 des Programms (<http://www-03.ibm.com/software/products/de/spss-amos>), die ausschließlich unter Windows eingesetzt werden kann.<sup>7</sup> Eine bezüglich der maximalen Zahl von Variablen geringfügig eingeschränkte ältere Version von AMOS kann kostenlos aus dem Internet heruntergeladen werden (<http://www.amosdevelopment.com/download/index.htm>).<sup>8</sup> Die meisten im Text besprochenen Beispiele können mit dieser Version nachvollzogen werden.

Als bislang jüngstes kommerzielles Produkt erschien schließlich Ende 1998 die erste Version von Bengt Muthéns Mplus.<sup>9</sup> Mplus unterscheidet sich insofern relativ stark von den anderen drei Programmen, als es die Implementation eines sehr allgemeinen mathematischen Modells darstellt, das Strukturgleichungsmodelle als einen Spezialfall einschließt. Die grafische Oberfläche ist im Vergleich mit LISREL, EQS und vor allem AMOS rudimentär. Ähnlich wie bei SIMPLIS können die Modelle jedoch leicht durch eine Reihe von Gleichungen beschrieben werden. Eine kostenlose Variante der aktuellen Version 7.2 ist im Internet verfügbar (<http://www.statmodel.com/demo.shtml>). Diese ist bezüglich der Anzahl der Variablen jedoch stark eingeschränkt. Studierende können die Vollversion zu einem

---

<sup>6</sup>Für das neuere und stärker spezialisierte Programm EQSIRT existieren auch Versionen für Mac OS X und Linux.

<sup>7</sup>AMOS wurde zwischenzeitlich als Modul für das Programmpaket SPSS vertrieben, ist aber heute wieder ein selbständiges Programm, das unabhängig von SPSS genutzt werden kann.

<sup>8</sup>Bei dem mit der Übernahme durch SPSS verbundenen Versionsprung in den zweistelligen Bereich scheint es sich in erster Linie um eine Marketingmaßnahme zu handeln.

<sup>9</sup>Eine Vorgängerversion war Muthéns LISCOMP.

reduzierten Preis erhalten. Anders als im Falle der älteren Programme existieren von Mplus native Versionen für Windows, Mac OS X und zahlreiche Varianten von Linux. Byrne (2012) gibt einen ausführlichen Überblick über die Möglichkeiten von AMOS, EQS, LISREL und Mplus (auf dem Stand des Jahres 2011) und entwickelt Kriterien für eine Kaufentscheidung.

Neben diesen spezialisierten Programmen verfügen auch einige der großen Pakete über mehr oder weniger eingeschränkte Fähigkeiten zur Schätzung von Strukturgleichungsmodellen. SAS enthält seit langer Zeit mit PROC CALIS eine entsprechende Prozedur, mit der sich viele Standardmodelle schätzen lassen. Verglichen mit den oben vorgestellten Programmen wirkt PROC CALIS jedoch antiquiert.

Für Stata ist das von Sophia Rabe-Hesketh auf der Grundlage ihrer Arbeiten mit Anders Skrondal und Andrew Pickles entwickelte Erweiterungsmodul GLLAMM verfügbar, das kostenlos aus dem Internet heruntergeladen werden kann (<http://www.gllamm.org/>). Ähnlich wie Mplus basiert GLLAMM auf einem sehr allgemeinen mathematischen Modell und kann *auch* Strukturgleichungsmodelle schätzen. Leider unterscheiden sich Syntax, Ausgabe und Philosophie von GLLAMM stark von derjenigen anderer Programme, was den Zugang erschwert. Zudem arbeitet GLLAMM aufgrund seiner Implementation der Parameterschätzung selbst für viele relativ einfache Modelle so langsam, dass eine interaktive Nutzung praktisch ausgeschlossen ist. Da GLLAMM von Rabe-Hesketh sehr aktiv weiterentwickelt wird, wird sich dies in Zukunft jedoch möglicherweise ändern.

Seit der Version 12 enthält Stata selbst die Prozedur `sem`, mit der sich sehr viele Standardmodelle effizient schätzen lassen. Der große Vorteil liegt hier darin, dass `sem` vollständig in Statas System zur Verwaltung von Daten und Schätzergebnissen eingebunden ist. Da Stata in den politikwissenschaftlichen Instituten im deutschsprachigen Raum inzwischen weite Verbreitung gefunden hat, dürfte `sem` vielen Lesern einen einfachen Zugang zur Schätzung von Strukturgleichungsmodellen bieten. Mit der Version 13 wurden die Möglichkeiten von Stata noch einmal erheblich erweitert. Die neue Prozedur `gsem` – das `g` steht für „generalized“ – unterstützt nun bei fast identischer Syntax auch kategoriale beobachtete Variablen (siehe Abschnitt 4.1). Da es sich bei GLLAMM und `(g)sem` um Prozeduren handelt, die innerhalb von Stata aufgerufen werden, sind beide unter allen Betriebssystemen verfügbar, die nativ von Stata unterstützt werden (Windows, aktuelle Versionen von Mac OS X und sowie alle x-86-basierten Linux-Varianten).

Schließlich steht mit dem von John Fox entwickelten `sem`-Paket eine Erweiterung für die open-source Programmierumgebung R (<http://cran.r-project.org/>) zur Verfügung, die in ihrer Leistungsfähigkeit mit früheren Versionen von LISREL vergleichbar ist und Modellspezifikationen in einer SIMPLIS-artigen Syntax akzeptiert. `sem` befindet sich derzeit noch in einer frühen Entwicklungsphase, stellt aber, wenn es um die Modellierung von Standardproblemen geht, bereits jetzt eine echte Alternative zu den kommerziellen Programmen dar. Durch die Integration in R

stehen den Benutzern von sem eine Vielzahl von flexiblen und innovativen Analysemöglichkeiten zur Verfügung. Alternative R-Pakete zur Schätzung von typischen Strukturgleichungsmodellen sind OpenMx, semPLS, lava und lavaan. Speziell über die Möglichkeiten von sem und OpenMx informieren Fox u. a. (2012). Auf stärker spezialisierte Angebote verweist der CRAN Task View „Psychometrics“ (<http://cran.r-project.org/web/views/Psychometrics.html>). Wie das gesamte R-System sind auch sem, lava und lavaan und viele weitere Pakete kostenlos und im Quelltext für alle gängigen Betriebssysteme erhältlich.

Darüber hinaus existiert eine Vielzahl von Stand-alone Programmen, die – oft schon in den 1980er Jahren – im akademischen Bereich entwickelt wurden und frei oder nach einer kostenlosen Registrierung verfügbar sind. Zu den wichtigsten dieser Programme zählen Mx (<http://www.vcu.edu/mx/>), das in seiner Zielsetzung mit LISREL, EQS und AMOS konkurriert und in seiner Windows-Variante über eine (separat zu installierende) graphische Benutzeroberfläche verfügt, sowie TETRAD (<http://www.phil.cmu.edu/projects/tetrad>) und SmartPLS<sup>10</sup> (<http://www.smartpls.de/>), die spezielle Ansätze implementieren, die derzeit (noch) außerhalb des Mainstream liegen. Insbesondere Mx stellt aber inzwischen eine ernstzunehmende Alternative zu den kommerziellen Programmen dar.

Neben Programmen und Beispieldatensätzen sind im Internet darüber hinaus eine Vielzahl von teils hochkarätigen Informationsquellen verfügbar. Zu den wichtigsten Anlaufstellen zählt SEMNET, eine klassische Mailingliste, die seit 1993 besteht und derzeit weltweit über 2 300 Abonnenten zählt, darunter zahlreiche Wissenschaftler, die führend an der Entwicklung der Methode beteiligt waren und sind. Mehrere zehntausend Nachrichten, die über SEMNET verschickt wurden, wurden archiviert und können über ein Formular (<http://bama.ua.edu/archives/semnet.html>) nach Schlüsselwörtern durchsucht werden. Ansonsten ist SEMNET jedoch nicht mit den heute üblichen Foren oder Communities vergleichbar. Um an den Diskussionen teilnehmen zu können, schickt man aus einem beliebigen Mailprogramm eine Nachricht an die Mailadresse [listserv@bama.ua.edu](mailto:listserv@bama.ua.edu), die nur das Kommando SUBSCRIBE SEMNET enthält, und erhält dann im Gegenzug eine Mail, in der man aufgefordert wird, sein Abonnement zu bestätigen. Von diesem Moment an kann man Nachrichten an die Liste senden und erhält im Gegenzug eine Kopie aller Mails, die dort eingehen. Außerhalb der Ferienzeit sind dies einige 100 Nachrichten pro Monat.

---

<sup>10</sup>Ebenso wie das R-Paket semPLS basiert SmartPLS auf dem Partial Least Squares-Ansatz, der letztlich eine hintereinander geschaltete Serie von (OLS)-Regressionen implementiert, statt simultan die vollständigen Kovarianzstrukturen zu analysieren. Der PLS-Ansatz stellt geringere Anforderungen an die Daten. Seine Anwendung bietet sich insbesondere dann an, wenn die Zahl der Indikatoren sehr groß ist. Allerdings weist er auch spezifische Schwächen auf und wurde u. a. deshalb in den Sozialwissenschaften kaum adaptiert. Im folgenden wird er deshalb nicht weiter behandelt. Einen Überblick über die Besonderheiten des PLS-Ansatzes geben Haenlein und Kaplan (2004). Weiber und Mühlhaus (2014) haben das erste Lehrbuch vorgelegt, das sich in größerem Umfang mit dem PLS-Ansatz auseinandersetzt.

Es empfiehlt sich deshalb, für die Nutzung von SEMNET eine eigene Mailadresse mit viel Speicherplatz einzurichten (etwa bei Google oder Yahoo) oder zumindest für die Standardadresse geeignete Filterregeln einzurichten, um unliebsame Überraschungen zu vermeiden. Außerdem benötigt man zwingend ein E-Mailprogramm beziehungsweise einen entsprechenden Webclienten, der in der Lage ist, Nachrichten, die sich auf denselben Gegenstand beziehen (threads), zu erkennen und gruppiert darzustellen. Ansonsten ist es unmöglich, den Diskussionen zu folgen.

Eine weitere hilfreiche Ressource ist „Ed Rigdon’s SEM FAQ“ (<http://www2.gsu.edu/~mkteer/semfaq.html>), eine Liste von weiterführenden Links und Antworten auf Fragen, die sich im Zusammenhang mit Strukturgleichungsmodellen immer wieder stellen. Leider wurden viele der Seiten seit Mitte der 1990er Jahren nicht mehr aktualisiert. Für eine Auseinandersetzung mit klassischen Stolpersteinen sind sie aber dennoch nützlich. Von der Zielsetzung vergleichbar, aber aktueller ist die von David A. Kenny gepflegte website <http://davidakenny.net/cm/causal.htm>. Eine nach Unterthemen gegliederte Bibliographie, die hunderte von Artikeln, Kapiteln und Monographien enthält, stellt Jason T. Newsom zur Verfügung (<http://www.upa.pdx.edu/IOA/newsom/semrefs.htm>). Seit 2002 wurde dieses Projekt bedauerlicherweise nicht weitergeführt, verzeichnet aber immer noch einen Großteil der relevanten Literatur. Auf einer weiteren Seite desselben Autors (<http://www.upa.pdx.edu/IOA/newsom/semclass/default.htm>) finden sich außerdem Datensätze, Übungen und Seminarpräsentationen. Sehr umfassend und aktuell schließlich ist das rund 300 Seiten umfassende Skript von Wolfgang Langer zu dessen LISREL-Kurs, das unter <http://www.soziologie.uni-halle.de/langer/lisrel/index.html> abrufbar ist.

Abschließend sei schließlich noch auf die Web-Angebote der kommerziellen Programme verwiesen. Insbesondere die Distributoren von LISREL (<http://www.ssicentral.com/lisrel>) und Mplus (<http://www.statmodel.com/>) bieten eine Vielzahl von Datensätzen, Hintergrundinformationen, Literaturverweisen und Arbeitspapieren an, die sich als äußerst nützlich erweisen können.