

DISKUSSIONSBEITRÄGE
AUS DEM
FACHBEREICH
WIRTSCHAFTSWISSENSCHAFTEN
DER UNIVERSITÄT DUISBURG–ESSEN

Nr. 127
Juli 2003

Das Identifikationsproblem in der Ökonometrie

Peter von der Lippe

Das Identifikationsproblem in der Ökonometrie (Einführung)

von Peter von der Lippe

1	Charakter des Problems (Seite 1)
2	Beispiele: verschiedene Marktmodelle und ihre reduzierte Form (Seite 3)
3	Interdependenz und Identifikation (Seite 5)
4	Identifizierbarkeit durch a priori Kenntnisse über Parameter, Störgrößen und exogene Variablen (Seite 8)
5	Identifikationskriterien: Restriktionen bezüglich struktureller Parameter (Seite 10)
6	Über- und Unteridentifikation, Multikollinearität (Seite 14)
	Vier Anhänge (Seite 15)

1. Charakter des Problems

1.1. Die Vollständigkeit eines ökonometrischen Modells

$$(1.1) \quad \mathbf{B}\mathbf{y} + \mathbf{G}\mathbf{z} = \mathbf{u}$$

$$\mathbf{B}: G \times G, \mathbf{G}: K \times G, \mathbf{y} \text{ und } \mathbf{u}: G \times 1, \mathbf{z}: K \times 1$$

besagt, dass bei gegebenen Variablen-Vektoren \mathbf{u} und \mathbf{z} sowie Koeffizienten-Matrizen \mathbf{B} und \mathbf{G} die G endogenen Variablen \mathbf{y} des Vektors $\mathbf{y}' = [y_1, \dots, y_G]$ eindeutig bestimmt sind, d.h., dass Gl. 1 aus G unabhängigen Gleichungen besteht.

Ein vollständiges Modell ist zweitens **identifizierbar**, wenn bei gegebenen \mathbf{u} , \mathbf{z} , \mathbf{y} auch \mathbf{B} und \mathbf{G} eindeutig bestimmbar sind. Ist ein vollständiges Modell zugleich auch identifizierbar (was nicht notwendig aus der Vollständigkeit folgt), so ist drittens zu untersuchen, ob und wie es *geschätzt* werden kann.

In einem Modell können evtl. nur einzelne strukturelle Gleichungen oder auch nur einzelne Parameter identifizierbar sein. Ein rekursives Modell ist stets identifizierbar. Definitionsgleichungen bieten kein Identifikationsproblem, da alle Parameter a priori bekannt sind. Sie haben nur die Funktion, die Lösungsmenge zu beschränken und zählen mit bei den G Gleichungen des Modells. Zu unterscheiden ist auch Unter- und Überidentifikation. Unteridentifikation ist kein stochastisches Problem (hier ist überhaupt keine Schätzung möglich), wohl aber die Überidentifikation, bei der es ein Nebeneinander verschiedener Schätzmöglichkeiten gibt.

1.2. Es gibt verschiedene Definitionen der Identifizierbarkeit (exakte Identifikation oder Überidentifikation), die äquivalent sind¹:

1. die strukturellen *Gleichungen* müssen statistisch *verschieden* sein: eine Gleichung y_i ($i = 1, \dots, G$), in der *alle* Variablen des Modells vorkommen, ist stets unteridentifiziert, weil jede Linearkombination dieser Gleichung mit einer anderen Gleichung des Modells von der ursprünglichen Gleichung für y_i nicht unterscheidbar ist (zum Begriff der *Beobachtungsäquivalenz*, siehe Abschn. 3.3);
2. die Likelihoodfunktion als Funktion in den Strukturkoeffizienten $L = L(\mathbf{B}, \mathbf{G}) = L(\mathbf{A})$ hat ein eindeutiges Maximum (exakte- bzw. Überidentifikation) oder aber mehrere Maxima (Unteridentifikation), man beachte, dass \mathbf{A} hier die Blockmatrix $[\mathbf{B}; \mathbf{G}]$ ist; und

¹ Valvanis, Econometrics, S. 90 f.

3. die Parameter können aus der bedingten Verteilung $f(y_1, \dots, y_G | x_1, \dots, x_K, v_1, \dots, v_G)$ also der reduzierten Form bestimmt werden.

Die Identifizierbarkeit **nichtlinearer** Gleichungssysteme ist nicht in der gleichen Weise vor der Schätzung mit Identifikationskriterien (siehe unten) zu überprüfen, wie das bei linearen Systemen möglich ist. Es soll nun der dritte der genannten Begriffe (eindeutiger Schluß aus der reduzierten Form) näher erläutert werden.

1.3 (von der reduzierten Form zur strukturellen Form) Die Identifikation eines ökonomischen Modells als *Ganzes* verlangt, dass von der reduzierten Form eindeutig auf die strukturellen Form zurück geschlossen werden kann:

reduzierte Form		strukturelle Form
(1.2) $\mathbf{y}_t = -\mathbf{B}^{-1}\mathbf{G}\mathbf{z}_t + \mathbf{B}^{-1}\mathbf{u}_t = \mathbf{P}\mathbf{z}_t + \mathbf{B}^{-1}\mathbf{u}_t$	→	$\mathbf{B}\mathbf{y}_t + \mathbf{G}\mathbf{z}_t = \mathbf{u}_t$ oder $[\mathbf{B} \quad \mathbf{G}] \begin{bmatrix} \mathbf{y}_t \\ \mathbf{z}_t \end{bmatrix} = \mathbf{u}_t$

wobei $\mathbf{P} = -\mathbf{B}^{-1}\mathbf{G}$ aus GK Koeffizienten besteht und in $\mathbf{A} = [\mathbf{B}:\mathbf{G}]$ genau $G(G + K)$ -G Koeffizienten (bei Normalisierung) also $G(G + K - 1)$ Koeffizienten auftreten.

Mit GK Koeffizienten der reduzierten Form sind also $G(G - 1) + GK$ zu schätzen, d.h. es müssen mindestens $G(G - 1)$ Koeffizienten β bzw. γ (in den Koeffizientenmatrizen \mathbf{B} und \mathbf{G}) Null sein (*notwendige* Bedingung).

Die reduzierte Form läßt sich stets mit der Methode der kleinsten Quadrate bestimmen, denn $E(\mathbf{y} | \mathbf{z}) = \mathbf{P}\mathbf{z}$ sofern nicht die Regressoren \mathbf{z} kollinear sind und mit \mathbf{P} und \mathbf{S}_v (Varianz-Kovarianzmatrix) ist die bedingte Verteilung von \mathbf{y} bestimmt. Oder:

Eine Struktur (S_i) eines Modells (S_1, S_2, \dots) ist identifiziert, wenn keine andere Struktur desselben Modells dieselbe reduzierte Form (\mathbf{P}, \mathbf{S}_v) hat. Andernfalls sind mehrere Strukturen mit den Beobachtungen (d.h. der bedingten Verteilung von y_1, \dots, y_G) verträglich (sie sind *beobachtungäquivalent* [observational equivalence]).

1.4. Das Problem läßt sich **verallgemeinern**: ein Identifikationsproblem ist stets dann gegeben, wenn verschiedene Spezifikationen eines Modells zu den gleichen beobachtbaren Konsequenzen führen. Dann ist mit empirischen Methoden nicht entscheidbar, welche Struktur S_1, S_2, \dots des Modells M ($S_1, S_2, \dots \in M$) die Beobachtungen erzeugt haben könnte.

Beispiel: "Fehler in den Variablen" (Meßfehler) in y und x bei der Regression von y auf x . Die Regressionsfunktion ist dann ohne Spezifikationen über die Verteilung der Meßfehler nicht identifizierbar.

Im folgenden behandeln wir nur Beispiele für das Identifikationsproblem im Zusammenhang mit simultanen (also $\mathbf{B} \neq \mathbf{I}$) ökonomischen Modellen.

1.5. Auch bei **dynamischen Modellen** entstehen Identifikationsprobleme. Das autoregressive Modell $\mathbf{y}_t = \sum \beta_\tau L^\tau \mathbf{y}_t + \gamma \mathbf{z}_t + \mathbf{u}_t$ (wobei \mathbf{g} und \mathbf{z} K -reihige Vektoren und L der Lag-Operator ist), sind die Parameter b_t nur bei Restriktionen bezüglich der zulässigen Prozesse $\{\mathbf{u}_t\}$ identifizierbar. In allen Fällen, in denen erwartungstreue und konsistente Schätzungen von β möglich sind, ist auch die Identifizierbarkeit gesichert. Da jedes autoregressive Modell in ein distributed-lag-Modell $\mathbf{y}_t = \sum \beta_\tau L^\tau \mathbf{z}_t + \gamma \mathbf{z}_t + \mathbf{u}_t$ überführt werden kann und umgekehrt, können auch hier Identifikationsprobleme entstehen².

² vgl. hierzu: Schönfeld. Methoden der Ökonometrie, Bd. 2, S. 33, 60f, 87.

2. Beispiele: verschiedene Marktmodelle und ihre reduzierte Form

2.1. Untersuchung der reduzierten Formen

Ohne Beschränkung der Allgemeinheit kann bei der Nachfrage- und der Angebotsfunktion das absolute Glied γ_{10} und γ_{20} weggelassen werden (Verwendung von deviation scores [zentrierte Variablenwerte, so dass der Mittelwert 0 ist]!).

Es seien y_1 = Menge, y_2 = Preis und z_1, z_2, z_3, \dots seien verschiedene exogene Variablen

Modell	Nachfrage	Angebot
A	$y_1 + \beta_{12}y_2 = u_1$	$y_1 + \beta_{22}y_2 = u_2$
B	$y_1 + \beta_{12}y_2 + \gamma_{11}z_1 = u_1$	$y_1 + \beta_{22}y_2 = u_2$
C	$y_1 + \beta_{12}y_2 + \gamma_{11}z_1 = u_1$	$y_1 + \beta_{22}y_2 + \gamma_{22}z_2 = u_2$
D	$y_1 + \beta_{12}y_2 + \gamma_{11}z_1 + \gamma_{13}z_3 = u_1$	$y_1 + \beta_{22}y_2 + \gamma_{22}z_2 = u_2$

oder in Matrixschreibweise (Modell D, die anderen sind Spezialfälle hiervon)

$$(2.1) \quad \begin{bmatrix} 1 & \beta_{12} \\ 1 & \beta_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} \gamma_{11} & 0 & \gamma_{13} \\ 0 & \gamma_{22} & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{bzw. } \mathbf{B}\mathbf{y} + \mathbf{G}\mathbf{z} = \mathbf{u}$$

Die reduzierte Form errechnet sich unter Berücksichtigung von $|\mathbf{B}| = \beta_{22} - \beta_{12}$ und der In-

versen $-\mathbf{B}^{-1}\mathbf{G} = \frac{1}{\beta_{22} - \beta_{12}} \begin{bmatrix} -\beta_{22}\gamma_{11} & \beta_{12}\gamma_{22} - \beta_{22}\gamma_{13} \\ \gamma_{11} & -\gamma_{22} & \gamma_{13} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \end{bmatrix}$ sowie $\mathbf{B}^{-1} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ als

$$(2.2) \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{oder}$$

$\mathbf{y} = (-\mathbf{B}^{-1}\mathbf{G})\mathbf{z} + \mathbf{B}^{-1}\mathbf{u}$ (= reduzierte Form), so dass b_{ij} die Koeffizienten in \mathbf{B}^{-1} sind.

Für die oben unterschiedenen Modelle bedeutet dies dass das System nach Art der Übersicht 1 aufgebaut ist. so dass man die Strukturkoeffizienten gemäß Übersicht 2 berechnen kann.

Im Modell D muss danach gelten

$$\beta_{22} = \frac{-\pi_{11}}{\pi_{22}} = \frac{-\pi_{13}}{\pi_{23}},$$

was bei empirischen Koeffizienten der reduzierten Form aber keineswegs notwendig erfüllt sein muß, d. h. die **Angebotsfunktion im Modell D** ist **überidentifiziert**. Man wird in der

Regel mit $\frac{-\pi_{11}}{\pi_{22}}$ und $\frac{-\pi_{13}}{\pi_{23}}$ unterschiedliche Schätzer $\hat{\beta}_{22}$ für β_{22} erhalten.

Das **Modell A** ist demgegenüber **unteridentifiziert** und im Modell B ist die Angebots- aber nicht die Nachfragekurve identifizierbar.

Nur das Modell C ist als ganzes **exakt identifiziert**.

Übersicht 1: reduzierte Formen der Marktmodelle

Modell	$b_{i1}u_1$	$b_{i2}u_2$	$\pi_{i1}Z_1$	$\pi_{i2}Z_2$	$\pi_{i3}Z_3$
A	$y_1 = \frac{\beta_{22}u_1}{\beta_{22} - \beta_{12}}$	$\frac{-\beta_{12}u_2}{\beta_{22} - \beta_{12}}$			
	$y_2 = \frac{-u_1}{\beta_{22} - \beta_{12}}$	$\frac{u_2}{\beta_{22} - \beta_{12}}$			
B	$y_1 =$	wie Modell A	$\frac{-\beta_{22}\gamma_{11}Z_1}{\beta_{22} - \beta_{12}}$		
	$y_2 =$	wie Modell A	$\frac{\gamma_{11}Z_1}{\beta_{22} - \beta_{12}}$		
C	$y_1 =$	wie Modell B		$\frac{\beta_{12}\gamma_{22}Z_2}{\beta_{22} - \beta_{12}}$	
	$y_2 =$	wie Modell B		$\frac{-\gamma_{22}Z_2}{\beta_{22} - \beta_{12}}$	
D	$y_1 =$	wie Modell C			$\frac{-\beta_{22}\gamma_{13}Z_3}{\beta_{22} - \beta_{12}}$
	$y_2 =$	wie Modell C			$\frac{\gamma_{13}Z_3}{\beta_{22} - \beta_{12}}$

* die Größen b_{ij} sind die Koeffizienten in \mathbf{B}^{-1} gem. Gl. 2.2 (reduz. Form)

Übersicht 2: Berechnung der Koeffizienten der strukturellen Form (Strukturkoeffizienten)

im Modell				
	A	B	C	D
β_{12}	nicht möglich	nicht möglich	$-\pi_{12}/\pi_{22}$	wie bei Modell C
β_{22}	nicht möglich	$-\pi_{11}/\pi_{21}$	wie bei Modell B	es müssen zwei Gleichungen für β_{22} gelten*
γ_{11}	existiert nicht	nicht möglich	$-\pi_{11} - \pi_{21}\beta_{12}$	wie bei Modell C
γ_{22}	existiert nicht	existiert nicht	$-\pi_{12} - \pi_{22}\beta_{22}$	wie bei Modell C
γ_{13}	existiert nicht	existiert nicht	existiert nicht	$-\pi_{13} - \pi_{23}\beta_{12}$

* vgl. Text oben

2.2. Abzählregel (notwendiges, aber nicht hinreichendes Identifikationskriterium):

Die Anzahl N der in einer Gleichung ausgeschlossenen Variablen (endogen oder vorherbestimmt) muß $G-1$ sein (G = Anzahl der endogenen Variablen des Modells).

In den Modellen (obigen Beispielen unter Nr. 2.1) ist $G = 2$ und

Modell	K (vorherbestimmte Variable)	ausgeschlossen in der Nachfragefunktion	ausgeschlossen in der Angebotsfunktion	Bemerkung
A	keine ($K = 0$)	$G = 2, N = 0 < G - 1$ Null Variablen ausgeschlossen	$N = 0$ Variablen ausgeschlossen	unteridentifiziert
B	$K = 1$ (z_1)	$N = 0$ (keine Variable ausgeschlossen)	$N = 1$ (nämlich z_1 ausgeschlossen)	nur Angebot ist identifiziert
C	$K = 2$ (z_1, z_2)	$N = 1$ (z_2 ausgeschlossen)	$N = 1$ (z_1)	exakt identifiziert
D	$K = 3$ (z_1, z_2, z_3)	$N = 1$ (z_2)	$N = 2$ (z_1 und z_3)	nur Nachfragefunktion ist identifiziert

3. Interdependenz und Identifikation

3.1. OLS und Haavelmo-Bias³

Wendet man die Methode der kleinsten Quadrate auf die strukturellen Gleichungen (= ordinary least squares OLS) des Modells A an, so erhält man die folgenden Schätzwerte

$$(3.1) \quad \hat{\beta}_{12} = \hat{\beta}_{22} = \frac{\text{Cov}(y_1 y_2)}{\text{Var}(y_2)} = \frac{C(y_1 y_2)}{V(y_2)}.$$

Mithilfe der folgenden Gleichungen (vgl. Übers. 1) erhält man

$$(3.2) \quad y_1 = \frac{\beta_{12}u_2 - \beta_{22}u_1}{\beta_{12} - \beta_{22}} \quad \text{und} \quad y_2 = \frac{u_2 - u_1}{\beta_{12} - \beta_{22}}$$

so dass man für die Varianz $V(\cdot)$ und Kovarianz $C(\cdot)$ in Gl. 3.1 erhält

$$(3.3a) \quad V(y_2) = \frac{\sigma_2^2 - 2\sigma_{12} + \sigma_1^2}{(\beta_{12} - \beta_{22})^2} \quad \text{und}$$

$$(3.3b) \quad C(y_1 y_2) = \frac{\beta_{12}\sigma_2^2 - (\beta_{12} + \beta_{22})\sigma_{12} + \beta_{22}\sigma_1^2}{(\beta_{12} - \beta_{22})^2}$$

so dass sich für

$$(3.4) \quad \frac{C(y_1 y_2)}{V(y_2)} = \frac{\beta_{12}\sigma_2^2 - (\beta_{12} + \beta_{22})\sigma_{12} + \beta_{22}\sigma_1^2}{\sigma_2^2 - 2\sigma_{12} + \sigma_1^2}$$

³ Der hier dargestellte simultaneous equations bias (erstmalig gezeigt in T. Haavelmo. The Statistical Implications of a System of Simultaneous Equations, *Econometrica* 11 (1943), 1 - 12) wird üblicherweise nicht an diesem Angebots- und Nachfragenkurvenschema sondern an dem Zwei-Gleichungs-Modell (1) $C = C(Y)$ (Konsumfunktion, daher $Y \rightarrow C$) und (2) $Y = C + I$ (daher auch die umgekehrte Kausalität $C \rightarrow Y$) dargestellt. Im Anhang wird noch einmal deutlich gemacht, dass dieser Bias nicht mit Frage zu tun hat, ob ein Modell identifizierbar ist oder nicht, sondern nur mit der Existenz eines interdependenten Modells.

ergibt. Dabei ist im allgemeinen wegen der Interdependenz von y_1 und y_2 nicht zwischen β_{12} und β_{22} zu unterscheiden, weil das Modell A ein unteridentifiziertes Modell ist⁴, d.h. die Angebots- (β_{22}) und Nachfragekurve (β_{12}) können gar nicht identifiziert werden.

3.2. Interpretation der Unteridentifikation vom Modell A

Man kann jetzt annehmen

Annahmen		Konsequenz
1. Unkorreliertheit der Störgrößen und Konstanz der Nachfrage (alle Beobachtungen liegen auf einer gegebenen Nachfragekurve und sind entstanden durch Verschiebungen der Angebotskurve)	$\sigma_{12} = 0$ $\sigma_1 = 0$	$\frac{C(y_1 y_2)}{V(y_2)} = \frac{\beta_{12} \sigma_2^2}{\sigma_2^2} = \beta_{12}$
2. Unkorreliertheit der Störgrößen und Konstanz des Angebots (alle Beobachtungen liegen auf einer gegebenen Angebotskurve und sind entstanden durch Verschiebungen der Nachfragekurve)	$\sigma_{12} = 0$ $\sigma_2 = 0$	$\frac{C(y_1 y_2)}{V(y_2)} = \frac{\beta_{22} \sigma_1^2}{\sigma_1^2} = \beta_{22}$

Im Fall 1 wird durch die Regression von y_1 auf y_2 die Nachfrage- (β_{12} identifizierbar), im Fall 2 die Angebotskurve (β_{22} identifizierbar) geschätzt. Dies sind Beispiele einer Identifikation durch Annahmen über die Varianz-Kovarianzmatrix der Störgrößen u_1 und u_2 also über

$$\mathbf{S} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \text{ (vgl. auch Abschn. 4.2).}$$

3.3. Beobachtungsäquivalenz von Angebot und Nachfrage im Modell A

Im Modell A sind die Nachfragefunktion $\beta_1' \mathbf{y} = \mathbf{u}_1$ und die Angebotsfunktion $\beta_2' \mathbf{y} = \mathbf{u}_2$ vom statistischen Standpunkt gleich, d.h. gleiche Funktionstypen und gleiche Variablen y_1 , y_2 und

u_1 , u_2 in den Vektoren $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ und $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$. Es genügt auch nicht zu wissen, dass $\beta_{22} > 0$

und $\beta_{12} < 0$ sein sollen (also aus ökonomischen Gründen die Steigungen positiv bzw. negativ sein sollten), um die Gleichungen zu unterscheiden, denn *jede* Linearkombination, d.h. Verschiebung, der beiden Funktionen, etwa

$$k_1 \beta_1' \mathbf{y} + k_2 \beta_2' \mathbf{y} = (k_1 \beta_1' + k_2 \beta_2') \mathbf{y} = \mathbf{d}_1' \mathbf{y} = v_1 \text{ mit } k_1 + k_2 = 1 \text{ (wegen der Normierung } \beta_{11} = \beta_{21} = 1)$$

$$\text{und ganz analog } r_1 \beta_1' \mathbf{y} + r_2 \beta_2' \mathbf{y} = (r_1 \beta_1' + r_2 \beta_2') \mathbf{y} = \mathbf{d}_2' \mathbf{y} = v_2 \text{ mit } r_1 + r_2 = 1,$$

erzeugt wieder lineare Gleichungen in y_1 und y_2 mit der *gleichen* reduzierten Form, nämlich

$$(3.5) \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{D}^{-1} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \text{ mit}$$

$$\mathbf{D}^{-1} = \begin{bmatrix} \mathbf{d}_1' \\ \mathbf{d}_2' \end{bmatrix}^{-1} = \frac{1}{(\beta_{22} - \beta_{12})(k_1 - r_1)} \begin{bmatrix} r_1(\beta_{12} - \beta_{22}) + \beta_{22} & -k_1(\beta_{12} - \beta_{22}) - \beta_{22} \\ -1 & 1 \end{bmatrix}$$

Bei $v_1 = k_1 u_1 + k_2 u_2$ und $v_2 = r_1 u_1 + r_2 u_2$. Ausmultipliziert ergibt das die reduzierte Form von Abschn. 2.1. Mit den Parametervektoren \mathbf{b}_1 und \mathbf{b}_2 sowie den Störgrößen u_1 und u_2 entsteht die gleiche bedingte Wahrscheinlichkeitsverteilung, wie mit \mathbf{d}_1 und \mathbf{d}_2 sowie den Störgrößen v_1 und v_2 .

⁴ Man beachte, dass dies ein grundlegenderer Mangel ist, als dass nur der Parameter β nicht erwartungstreu zu schätzen ist, wie im Beispiel von Haavelmo (vgl. vorangegangene Fußnote).

3.4. Verallgemeinerung

Die ökonomische Struktur $\mathbf{B}\mathbf{y} + \mathbf{G}\mathbf{z} = \mathbf{u}$ oder $\mathbf{A}\mathbf{x} = \mathbf{u}$ mit $\mathbf{A} = [\mathbf{B}:\mathbf{G}]$ und $\mathbf{x}' = [\mathbf{y}':\mathbf{z}']$ ist einer Struktur $\mathbf{A}^*\mathbf{x} = \mathbf{u}$ beobachtungsäquivalent, wenn eine Transformationsmatrix \mathbf{T} dergestalt existiert, dass

- 1) $\mathbf{A}^* = \mathbf{T}\mathbf{A}$ alle a priori Restriktionen von \mathbf{A} erfüllt
- 2) $\mathbf{S}^* = \mathbf{T}\mathbf{S}\mathbf{T}'$ alle a priori Restriktionen von \mathbf{S} erfüllt und
- 3) \mathbf{T} nichtsingulär ist, also $|\mathbf{T}| \neq 0$.

Das Modell ist dann (als Ganzes) exakt identifiziert, wenn die einzige zulässige Transformationsmatrix $\mathbf{T} = \mathbf{I}$ ist.

Zwei beobachtungsäquivalente Strukturen \mathbf{A} und \mathbf{A}^* haben die gleiche Likelihood-Funktion und die gleiche reduzierte Form⁵:

$$\mathbf{A}^* \mathbf{x} = (\mathbf{T}\mathbf{A})\mathbf{x} = (\mathbf{T}\mathbf{B})\mathbf{y} + (\mathbf{T}\mathbf{G})\mathbf{z} = \mathbf{u}^* = \mathbf{T}\mathbf{u} \text{ hat die reduzierte Form}$$

$$(\mathbf{T}\mathbf{B})^{-1}(\mathbf{T}\mathbf{B})\mathbf{y} + (\mathbf{T}\mathbf{B})^{-1}(\mathbf{T}\mathbf{G})\mathbf{z} = (\mathbf{T}\mathbf{B})^{-1}\mathbf{u} \text{ also}$$

$$(3.6) \quad \mathbf{y} = -(\mathbf{B}^{-1}\mathbf{T}^{-1}\mathbf{T}\mathbf{G})\mathbf{z} + (\mathbf{B}^{-1}\mathbf{T}^{-1}\mathbf{T})\mathbf{u} = -\mathbf{B}^{-1}\mathbf{G}\mathbf{z} + \mathbf{B}^{-1}\mathbf{u} = ? \mathbf{z} + \mathbf{v}.$$

Beispiel:

In 3.3 wurde das Modell A transformiert mit $\mathbf{T} = \begin{bmatrix} k_1 & 1 - k_1 \\ r_1 & 1 - r_1 \end{bmatrix}$ und $|\mathbf{T}| = k_1 - r_1$; dann ist

$$\mathbf{B}^* = \mathbf{T}\mathbf{B} = \begin{bmatrix} 1 & k_1(\beta_{12} - \beta_{22}) + \beta_{22} \\ 1 & r_1(\beta_{12} - \beta_{22}) + \beta_{22} \end{bmatrix}$$

Man erkennt hieran auch, dass $|\mathbf{T}| = k_1 - r_1 \neq 0$ sein muß, denn für $k_1 = r_1$ ist \mathbf{B}^* vom Range 1,

also $|\mathbf{B}^*| = 0$. Mit \mathbf{T} wird auch $\hat{\mathbf{a}}$ nicht verändert. Wenn im Modell A gilt $\mathbf{S} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$,

dann ist $\mathbf{S}^* = \begin{bmatrix} \sigma_1^{*2} & \sigma_{12}^* \\ \sigma_{12}^* & \sigma_2^{*2} \end{bmatrix}$ mit

$$\sigma_{12}^* = k_1 r_1 \sigma_1^2 + (k_1 r_2 + k_2 r_1) \sigma_{12} + k_2 r_2 \sigma_2^2,$$

$$\sigma_1^* = k_1^2 \sigma_1^2 + 2k_1 k_2 \sigma_{12} + k_2^2 \sigma_2^2 \text{ und}$$

$$\sigma_2^* = r_1^2 \sigma_1^2 + 2r_1 r_2 \sigma_{12} + r_2^2 \sigma_2^2.$$

Eine Restriktion $\sigma_{12} = 0$ bliebe durch diese Transformation jedoch nicht in Form von $\sigma_{12}^* = 0$ erhalten⁶.

3.5. Identifizierbarkeit rekursiver Modelle⁷

Rekursive Modelle sind *stets identifizierbar*, weil die einzige zulässige Transformationsmatrix \mathbf{T} in $\mathbf{B}^* = \mathbf{T}\mathbf{B}$ die Matrix $\mathbf{T} = \mathbf{I}$ ist, wobei \mathbf{B}^* und \mathbf{B} Dreiecksmatrizen sind (kennzeichnend für den rekursiven Modelltyp). \mathbf{B}^* ist ebenfalls eine Dreiecksmatrix sein und das

⁵ Johnston, *Econometric methods*, S.353 f.

⁶ Zahlenbeispiele hierzu: Schneeweiß, *Ökonometrie*, S. 264f.

⁷ Schneeweiß, *Ökonometrie*, S. 286, Aufgaben 10 - 17.

Produkt zweier Dreiecksmatizen $\mathbf{T} = \mathbf{B}^* \mathbf{B}^{-1}$ ist ebenfalls eine Dreiecksmatrix denn. Mit der normierten Dreiecksmatrix \mathbf{T} ($t_{ii} = 1$) wird aber die Diagonalgestalt von $\hat{\mathbf{a}}$ zerstört, denn

$$\mathbf{S}^* = \mathbf{T} \mathbf{S} \mathbf{T}' = \begin{bmatrix} \sigma_1^2 & t_{21} \sigma_1^2 & \dots \\ t_{21} \sigma_1^2 & t_{21}^2 \sigma_1^2 + \sigma_2^2 \dots & \\ \vdots & \vdots & \vdots \end{bmatrix}, \text{ eine symmetrische Matrix mit den Elementen}$$

$$s_{ii} = \sum_{j=1}^i t_{ij}^2 \sigma_j^2 \text{ in der Hauptdiagonale (bei } t_{ii} = 1) \text{ und } s_{ij} = \sum_{k=1}^i t_{ik} t_{jk} \sigma_k^2 + t_{ij} \sigma_j^2 \text{ sonst.}$$

Ergebnis: Jede Transformationsmatrix $\mathbf{T} \neq \mathbf{I}$ zerstört entweder nur die Restriktionen der Matrix $\hat{\mathbf{a}}$ oder diese und diejenigen von \mathbf{B} , so dass keine Beobachtungsäquivalenz mehr vorliegt. Das rekursive Modell ist also stets identifizierbar.

4. Identifizierbarkeit durch a priori Kenntnisse über Parameter, Störgrößen und exogene Variablen

Das unteridentifizierte Modell A:

$$\begin{aligned} y_1 + \beta_{12} y_2 &= u_1 && \text{Nachfrage} \\ y_1 + \beta_{22} y_2 &= u_2 && \text{Angebot} \end{aligned}$$

kann durch a priori Annahmen oder Informationen über strukturelle Parameter, über die Störvariablen und über exogene Variablen, bzw. eine Kombination dieser Fälle, identifizierbar gemacht werden. Wir betrachten hier nur einfache *Beispiele* von Annahmen:

4.1. Annahmen über Strukturkoeffizienten:

Identifizierbarkeit der Nachfrage im Modell B von Abschnitt 2.1 also der Gleichung $y_1 + \beta_{12} + \gamma_{11} z = u_1$ durch die Restriktion $\beta_{11} + \gamma_{11} = 0$, also nach Normalisierung ($\beta_{11} = 1$) $1 + \gamma_{11} = 0$, $\gamma_{11} = -1$. Die (identifizierbare) Angebotskurve bleibe wie bisher angenommen. Die reduzierte Form ist dann

$$-\mathbf{B}^{-1} \mathbf{G} = \frac{1}{\beta_{22} - \beta_{12}} \begin{bmatrix} +\beta_{22} \\ -1 \end{bmatrix} = \begin{bmatrix} \pi_{11} \\ \pi_{21} \end{bmatrix}, \text{ so dass mit } \beta_{12} = \frac{1 - \pi_{11}}{\pi_{21}} \text{ die Nachfragefunktion durch die}$$

Restriktion $\gamma_{11} = -1$ identifizierbar geworden ist.

Übliche a priori Restriktionen sind ferner $\beta_{ij} = 0$ oder $\beta_{ij} / \beta_{ke} = a$, eine bekannte Konstante a oder $\beta_{ij} = \beta_{ke}$. Bloße Ungleichheiten (Ungleichungen) wie $\beta_{ij} > 0$ oder $\beta_{ke} \leq 0$ helfen jedoch nicht bei der Beseitigung von Unteridentifikation.

4.2. über die Varianz – Kovarianz – Matrix \mathbf{S}_u der Störgrößen u

Beispiel: Identifikation der Nachfrage im Modell B von Abschn. 2.1 durch die Annahme

$$\mathbf{S}_u = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \text{ oder noch spezieller } \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & k \sigma_1^2 \end{bmatrix}, \text{ denn für transformierte Störgrößen } u^* \text{ gilt}$$

bei $\mathbf{T} = \begin{bmatrix} t_1 & 1 - t_1 \\ t_2 & 1 - t_2 \end{bmatrix}$, dass $\Sigma_u^* = \mathbf{T} \Sigma_u \mathbf{T}'$ nicht diagonal ist (vgl. Abschn. 3.4). Man erhält die

neue Störgröße $u_1^* = t_1 u_1 + (1 - t_1) u_2$. Da u_1 und u_2 unabhängig sind, kann u_1^* wie gefordert nur dann von u_2 unabhängig sein, wenn $t_1 = 0$, d.h. mit der obigen Spezifikation \mathbf{S}_u (als Diagonalmatrix) wird auch die Nachfrage identifizierbar. Da die reduzierte Form stets identifizierbar ist, ist die Varianz-Kovarianz-Matrix der Störgrößen v_1, v_2 der reduzierten

Form S_v gegeben. Ferner ist $\Sigma_v = \mathbf{B}^{-1} \Sigma_u (\mathbf{B}^{-1})'$ und $\Sigma_v = \Sigma_y$ (der beobachteten endogenen Variablen y), so dass man mit $\sigma_2^2 = k \cdot \sigma_1^2$ zwei Gleichungen in β_{12} und β_{22} erhält

$$(4.1) \quad S_v = \frac{\sigma_1^2}{(\beta_{22} - \beta_{12})^2} \begin{bmatrix} \beta_{22}^2 - \beta_{12}^2 k & -\beta_{22} - \beta_{12} k \\ -\beta_{22} - \beta_{12} k & 1 + k \end{bmatrix}$$

Daraus ergibt sich (ebenso aus Abschn. 3.1)

$$(4.2) \quad -\beta_{22} - k\beta_{12} = (1+k) \frac{C(v_1 v_2)}{V(v_2)} = (1+k) \frac{C(y_1 y_2)}{V(y_2)}$$

$$(4.3) \quad \beta_{22}^2 - k\beta_{12}^2 = (1+k) \frac{V(v_1)}{C(v_1, v_2)} = (1+k) \frac{V(y_1)}{C(y_1 y_2)}.$$

Doppeldeutigkeit⁸ ist dadurch zu eliminieren, dass aus ökonomischen Gründen $\beta_{12} < 0$ und $\beta_{22} > 0$ sein muss und β_{22} identifizierbar ist.

4.3. über exogene Variablen⁹

Zerlegung der Störgrößen u_1 in eine systematische (durch z_1 erklärte) und eine zufällige Komponente: $u_1 = \gamma_{11} z_1 + v_1$. Damit entsteht aus Modell A das Modell B für die Nachfragefunktion.

Da z_1 exogen, kann es als instrumentelle Variable (Instrumentalvariable) benutzt werden, d.h. Multiplikation von $y_1 = -\beta_{22} y_2 + u_2$ mit z_1 und Bildung von Erwartungswerten liefert

$E(y_1 z_1) = -\beta_{22} E(y_2 z_1) + E(u_2 z_1)$, wobei $E(u_2 z_1)$, die Kovarianz der Störgröße u_2 mit der exogenen Variablen z_2 null ist. Man schätzt somit

$$(4.4) \quad -\hat{\beta}_{22} = \frac{C(y_1 z_1)}{C(y_2 z_1)},$$

d.h. die Angebotsfunktion ist identifizierbar. Analoges Vorgehen bei der Nachfragekurve führt zu

$$(4.5) \quad -C(y_1 z_1) = \hat{\beta}_{12} C(y_2, z_1) + \hat{\gamma}_{11} V(z_1),$$

so dass die Nachfragekurve *nicht* identifizierbar ist.

4.4. Trotz exakter Identifizierbarkeit einer Gleichung im Rahmen eines Modells kann diese evtl. **nicht schätzbar** sein. Mit den Instrumentalvariablen z_1 und z_2 erhält man für die Nachfrage im Modell C

$$(4.6) \quad \begin{bmatrix} C(y_2 z_1) & V(z_1) \\ C(y_2 z_2) & C(z_1 z_2) \end{bmatrix} \begin{bmatrix} \hat{\beta}_{12} \\ \hat{\gamma}_{11} \end{bmatrix} = - \begin{bmatrix} C(y_1 z_1) \\ C(y_1 z_2) \end{bmatrix},$$

was voraussetzt, dass die Kovarianzmatrix vom Range 2 ist. Sie ist es nicht, wenn $z_1 = k z_2$, so dass die Zeilen linear abhängig sind (dann ist auch die Angebotsfunktion des Modells C nicht schätzbar: Kollinearität), bzw. wenn $y_1 = k z_1$ (oder $y_2 = k z_1$), so dass die Spalten linear abhängig sind (in diesem Fall kann die andere Gleichung, d.h. die Angebotsfunktion jedoch trotzdem schätzbar sein).

⁸ Im Grunde reicht Gl. 4.2 da ja β_{22} bekannt ist (Angebotsfunktion identifizierbar).

⁹ Ausgangspunkt der folgenden Betrachtung ist wieder die Nachfragekurve im Modell A.

5. Identifikationskriterien¹⁰: Restriktionen bezüglich struktureller Parameter

5.1. Ohne Beschränkung der Allgemeinheit wird die Identifizierbarkeit einer (der ersten) Strukturgleichung eines *deterministischen* Modells $\mathbf{B}\mathbf{y} + \mathbf{G}\mathbf{z} = \mathbf{0}$ untersucht, dessen reduzierte Form

$$(5.1) \quad \mathbf{y} = -\mathbf{B}^{-1}\mathbf{G}\mathbf{z} = \mathbf{P}\mathbf{z}$$

lautet, wobei \mathbf{B} und \mathbf{G} Blockmatrizen mit folgenden Vektoren und Untermatrizen seien

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}'_1 \\ \dots \\ \mathbf{B}_{\cdot 1} \end{bmatrix} \text{ mit } \mathbf{b}'_1 \text{ vom Typ } 1 \times G \text{ und } \mathbf{B}_{\cdot 1} \text{ vom Typ } (G-1) \times G \text{ ist und } \mathbf{G} = \begin{bmatrix} \mathbf{g}'_1 \\ \dots \\ \mathbf{G}_{\cdot 1} \end{bmatrix} \text{ analog definiert}$$

ist mit K Spalten. Dann lautet die erste Gl. ¹¹ der strukturellen Form $\mathbf{b}'_1 \mathbf{y} = -\mathbf{g}'_1 \mathbf{z}$ und unter Verwendung von Gl. 5.1 $\mathbf{b}'_1 \mathbf{P} \mathbf{z} = -\mathbf{g}'_1 \mathbf{z}$.

Postmultiplikation mit der Diagonalmatrix $\mathbf{D} = \{d_{ii}\} = \left\{ \begin{bmatrix} 1 \\ \dots \\ z_i \end{bmatrix} \right\}$, so dass gilt $\mathbf{z}\mathbf{D} = \mathbf{I}$, liefert

$-\mathbf{g}'_1 = \mathbf{b}'_1 \mathbf{P}$ bzw. als Spalten- statt Zeilenvektor

$$(5.2) \quad -\mathbf{g}'_1 = \mathbf{g}'_1 \mathbf{B}_{\cdot 1},$$

ein System vom K (Anzahl der exogenen Variablen) Gleichungen mit $K + G - 1$ Unbekannten, sofern keiner der Koeffizienten $1, \beta_{12}, \beta_{13}, \dots, \beta_{1G}, \gamma_{11}, \gamma_{12}, \dots, \gamma_{1K}$, a priori Null gesetzt werden kann. Annahme: von den Koeffizienten der G endogenen und der K exogenen Variablen seien in den Vektoren \mathbf{b}_1 und \mathbf{g}_1 genau $G-g$ bzw. $K-k$ Null, also

$$\mathbf{b} = \begin{bmatrix} 1 & \beta_{12} & \dots & \beta_{1g} & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{b}'_1 & \mathbf{0}' \\ 1 \times k & 1 \times (G-g) \end{bmatrix} \text{ und analog}$$

$$\mathbf{g}' = \begin{bmatrix} 1 & \gamma_{12} & \dots & \gamma_{1k} & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{g}'_1 & \mathbf{0}' \\ 1 \times k & 1 \times (K-k) \end{bmatrix}$$

Dann ist Gl. 5.2 zu schreiben als

$$(5.2a) \quad \begin{bmatrix} -\mathbf{g}'_1 \\ \dots \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{g}'_{k,g} & \mathbf{g}'_{k,G-g} \\ \mathbf{g}'_{K-k,g} & \mathbf{g}'_{K-k,G-g} \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{0}_{G-g} \end{bmatrix}$$

$$\text{mit den Dimensionen } \begin{bmatrix} k \\ K-k \end{bmatrix} = \begin{bmatrix} k \times g & k \times (G-g) \\ (K-k) \times g & (K-k) \times (G-g) \end{bmatrix} \begin{bmatrix} g \\ G-g \end{bmatrix}.$$

So dass man zwei Systeme erhält:

$$(5.3) \quad -\mathbf{g}'_1 = \mathbf{g}'_{k,g} \mathbf{b}_1 \text{ und}$$

$$(5.4) \quad \mathbf{0}_{K-k} = \mathbf{g}'_{K-k,g} \mathbf{b}_1.$$

Aus dem homogenen System 5.4 von $K-k$ Gleichungen mit g Unbekannten läßt sich der Vektor \mathbf{b}_1 (mit g Elementen) und damit gem. Gl. 5.3 auch der Vektor \mathbf{g}_1 (mit k Elementen)

¹⁰ Für die Identifizierbarkeit einer Gleichung im Rahmen eines (Mehrgleichungs)Modells.

¹¹ Der Einfachheit halber betrachten wir o.B.d.A. nur die erste Gleichung des Systems.

berechnen, sofern der Rang (Anzahl der linear unabhängigen Spaltenvektoren) von $P_{g, K-k}$ genau $g-1$ oder kleiner ist (er kann nicht größer als $g-1$ sein, da es nur g Zeilen sind und bei vollem Rang für b_1 nur die triviale Lösung möglich wäre, folglich ist der Fall $rg(P) \geq g$ ausgeschlossen).

Das **Rangkriterium** der Identifizierbarkeit der Strukturgleichung für y_1 ist dann

$$(5.5) \quad rg(P_{g, K-k}) \leq g - 1$$

Wegen $rg(P_{g, K-k}) = \min(g, K-k)$ sind bestimmte Kombinationen mit dem notwendigen, aber nicht hinreichenden Abzählkriterium (order condition vgl. Abschn. 2.2) nämlich

$$(K - k) + (G - g) \geq G - 1 \text{ also } K - k \geq g - 1$$

von vornherein ausgeschlossen, so dass gilt:

Abzählkriterium	Der Rang $rg(P_{g, K-k})$ ist		
	$< g - 1$	$= g - 1$	$> g - 1$
$K - k < g - 1$ (unteridentifiziert)	unteridentifiziert	unmöglich	unmöglich
$= g - 1$ (exakt identifiziert)	unteridentifiziert*	exakt identifiziert	unmöglich
$> g - 1$ (überidentifiziert)	unteridentifiziert*	überidentifiziert da lineare Abhängigkeit	unmöglich

In den Fällen * würde man sich mit dem Abzählkriterium irren. Dabei ist jeweils der Rang durch lineare Abhängigkeit zwischen Zeilen oder Spalten reduziert.

Im Spezialfall $g = 1$ und $k = K$ (keine simultanen Gleichungen, Fall der multiplen Regression von y_1 auf z_1, z_2, \dots, z_k , so dass das Identifikationsproblem gar nicht erst entsteht, erhält man aus Gl. 5.2 weil die reduzierte Form mit der strukturellen Form identisch ist und man OLS einfach auf die strukturellen Form anwenden kann: $\gamma_{1j} = -\pi_{1j}$ für $j = 1, 2, \dots, k$ (bzw. K , da $k = K$)

5.2. Andere Fassung des Rangkriteriums:

Eine Verhaltensgleichung ist identifizierbar, wenn wenigstens eine $G - 1$ reihige Determinante konstruierbar ist aus den Koeffizienten der Variablen, die aus der fraglichen Verhaltensgleichung ausgeschlossen sind (und in den übrigen $G - 1$ Verhaltensgleichungen enthalten sind), die nicht null ist.

Beispiel: Bei der Angebotskurve¹² im Modell B im Abschnitt 2.1 ist Identifizierbarkeit geben

wenn $\gamma_{11} \neq 0$ weil $\begin{vmatrix} 1 & 0 \\ 0 & \gamma_{11} \end{vmatrix} \neq 0$.

5.3. Für die stochastische Version gilt nach Prämultiplikation der reduzierten Form mit β_1' :

$$\beta_1' y = \beta_1' z + \beta_1' B^{-1} u$$

Die erste Zeile des Produkts $\beta_1' B^{-1}$ lautet $1 \ 0 \ 0 \ \dots \ 0$, so dass mit $\beta_1' B^{-1} u$ die Störgröße (Skalar!) u_1 der ersten Gleichung entsteht, die auch $u_1 = \beta_1' y + \beta_1' z$ ist. Daraus folgt $-\beta_1' z = \beta_1' y - u_1$ und damit wie oben Gl. 5.2.

¹² Beachte y_1 stellt zwei verschiedene Variablen dar (angebotene- und nachgefragte Menge).

5.4. Es empfiehlt sich, das Rangkriterium auch aus der strukturellen Form herzuleiten¹³. Vereinbart man für $\mathbf{B}\mathbf{y} + \mathbf{G}\mathbf{z} = \mathbf{0}$ mit $\mathbf{B} = \begin{bmatrix} \mathbf{0}'_{1 \times (G-g)} & \mathbf{B}'_1 \\ \mathbf{B}_0 & \dots \end{bmatrix}$ wobei \mathbf{B}_0 vom Typ $(G-1) \times (G-g)$ und \mathbf{y} ein $G \times 1$ Spaltenvektor ist und $\mathbf{G} = \begin{bmatrix} \mathbf{0}'_{1 \times (K-k)} & ?'_{1 \times k} \\ \mathbf{G}_0 & \dots \end{bmatrix}$ wobei \mathbf{G}_0 vom Typ $(G-1) \times (K-k)$ und \mathbf{z} ein $K \times 1$ Spaltenvektor ist. Dann muss zur exakten Identifizierbarkeit mit $\mathbf{A}_0 = [\mathbf{B}_0 : \mathbf{G}_0]$ die folgende Bedingung erfüllt sein

$$(5.6) \quad \text{rg}(\mathbf{A}_0) = G - 1$$

Beweis: Der Rang von \mathbf{A}_0 bleibt unverändert, wenn man \mathbf{A}_0 um die entsprechenden Nullvektoren¹⁴ zu einer $G \times (G-g + K-k)$ Matrix $\mathbf{A} = \begin{bmatrix} \mathbf{0}'_{1 \times (G-g)} & \mathbf{0}'_{1 \times (K-k)} \\ \mathbf{B}_0 & \mathbf{G}_0 \end{bmatrix}$. Das gilt auch, wenn man \mathbf{A} mit \mathbf{B}^{-1} prämultipliziert (da $\text{rg}(\mathbf{B}) = G > G-1$, allgemein: der Rang einer Matrix ist gegenüber der Multiplikation mit einer regulären Matrix invariant), wobei in \mathbf{B} die ersten $G-g$ Reihen durch die Koeffizienten der in Gl. 1 des Modells nicht enthaltenen endogenen Variablen besetzt seien, so entsteht wegen $\mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$, $-\mathbf{B}^{-1}\mathbf{G} = \mathbf{P}$ die Blockmatrix \mathbf{P}

$$\mathbf{P} = \begin{bmatrix} \overbrace{\mathbf{I}_{(G-g) \times (G-g)}}^{G-g} & \overbrace{-?_{(G-g) \times (K-k)}}^{K-k} \\ \mathbf{0}_{g \times (G-g)} & -?_{g \times (K-k)} \end{bmatrix}$$

wobei $?_{(G-g) \times (K-k)}$ eine Matrix ist, deren Gestalt im weiteren ohne Belang ist. Die Determinante dieser Matrix ist offenbar $|\mathbf{I} \cdot |-\mathbf{P}_{g \times (K-k)} + \mathbf{0}_{g \times (K-k)}| = 1 \cdot 0 = 0$, so dass die Blockmatrix \mathbf{P} nicht vom vollen Rang G ist, sondern vom Rang $G - 1$, denn

$$\text{rg}(\mathbf{P}) = \underbrace{\text{rg}(\mathbf{I})}_{G-g} + \underbrace{\text{rg}(?_{g \times (K-k)})}_{g-1} = \text{rg}(\mathbf{A}_0) = G - 1, \text{ da } \mathbf{P} \text{ eine (Block-)Dreiecksmatrix ist}^{15}.$$

$\text{rg}(\mathbf{A}_0) < G-1$ kann eintreten wenn:

1. \mathbf{A}_0 zu wenige Spalten ($G-g + K-k$) hat, so dass $G-g + K-k < G-1$ und damit $K-k < g-1$ (dieser Fall bestätigt lediglich das Abzählkriterium), und
2. Wenn eine Zeile oder Spalte von \mathbf{A}_0 identisch Null ist.
Gegeben sei das Modell ($G = 3$)

$$(1) \quad y_1 + \beta_{12}y_2 + \frac{1}{\beta_{21}}z_1 = u_1$$

$$(2) \quad \beta_{21}y_1 + y_2 + z_1 = u_2$$

$$(3) \quad \beta_{32}y_2 + y_3 + \gamma_{32} \cdot z_2 = u_3$$

Dann gilt für die erste Gleichung $\mathbf{A}_0 = \begin{bmatrix} 0 & 0 \\ 1 & \gamma_{32} \end{bmatrix}$ also $\text{rg}(\mathbf{A}_0) = 1$.

Zur Herleitung von \mathbf{A}_0 beachte: (Gl. 2) $0 \cdot y_3 + 0 \cdot z_2$ und (Gl.3) $1 \cdot y_3 + \gamma_{32} \cdot z_2$

¹³ Wonnacott u. Wonnacott, *Econometrics*, S. 348ff, ähnlich Johnston, *Econometric Methods*, S. 359f, aber mit völlig anderer Symbolik!

¹⁴ Zeilenvektoren, die die erste Zeile von \mathbf{A} bilden.

¹⁵ Gantmacher, Bd. 1, S. 44.

- 3) Wenn lineare Abhängigkeit von Zeilen oder Spalten in \mathbf{A}_0 besteht, was für Gl. 3 in diesem Modell der Fall ist, denn dann ist (die Zeilen und Spalten beziehen sich jetzt auf y_1 und z_1)

$$\mathbf{A}_0 = \begin{bmatrix} 1 & 1/\beta_{21} \\ \beta_{21} & 1 \end{bmatrix} \text{ also auch hier wieder } \text{rg}(\mathbf{A}_0) = 1 < G-1, \text{ weil } |\mathbf{A}_0| = 1 - 1 = 0.$$

5.5. Man kann die Identifikationsregeln auch wie folgt schreiben: gegeben sei ein Modell mit G endogenen und K exogenen Variablen. Die i -te Gleichung ist dann

Bedingung	unteridentifiziert underidentified	exakt identif. just identified	überidentifiz. overidentified
notwendig: Anzahl der in der i -ten Gl. nicht enthaltenen Variablen*	$< G - 1$	$= G - 1$	$> G - 1$
notwendig und hinreichend: Anzahl der nichtverschwindenden $G-1$ reihigen Determinanten gebildet aus den Koeffizienten der in der i -ten Gl. nicht enthält. Variablen*	0	1	> 1

* egal ob exogen oder endogen

Für das Kriterium mit den $G-1$ reihigen Determinanten vgl. Anhang Nr. 2, Seite 17 unten.

5.6. Transformationsmatrizen für a priori Restriktionen (des Typs der **zero restrictions**): Die Matrix $\mathbf{A}_0 = [\mathbf{B}_0; \mathbf{G}_0]$ erhält man für die i -te Strukturgleichung durch $\mathbf{A} \cdot \mathbf{f}_i$, wobei \mathbf{f}_i ein Spaltenvektor¹⁶ mit $G+K$ Zeilen ist. \mathbf{f}_i enthält das Element 1 für eine zero restriction, und Elemente 0 sonst. *Beispiel:* Die Nachfragefunktion im Modell D von Abschn. 2.1 ist

$$\mathbf{A} = \begin{bmatrix} 1 & \beta_{12} & \gamma_{11} & 0 & \gamma_{13} \\ 1 & \beta_{22} & 0 & \gamma_{22} & 0 \end{bmatrix}, \mathbf{f}_1' = [0 \quad 0 \quad 0 \quad 1 \quad 0], \mathbf{A}\mathbf{f}_1 = \begin{bmatrix} 0 \\ \gamma_{22} \end{bmatrix} \text{ und für die Angebotsfunktion gilt}$$

$$\mathbf{f}_2' = \begin{bmatrix} 00100 \\ 00001 \end{bmatrix}, \mathbf{A}\mathbf{f}_2 = \begin{bmatrix} \gamma_{11} & \gamma_{13} \\ 0 & 0 \end{bmatrix}$$

Das Kriterium der Identifizierbarkeit der i -ten Gleichung ist dann $\text{rg}(\mathbf{A}\mathbf{f}_i) = \text{rg}(\mathbf{A}_0) = G - 1$. Ein Modell ist dann überidentifiziert, wenn es wenigstens eine Gleichung i enthält, für die gilt $\text{rg}(\mathbf{A}\mathbf{f}_i) = G$.

5.7. Nichtlinearität kann Identifizierbarkeit (eines nicht identifizierbaren linearen Modells) bewirken. *Beispiel:* Modell A werde wie folgt modifiziert¹⁷

$$y_1 + \beta_{12}y_2 = u_1$$

$$y_1 + \beta_{22}y_2 + \beta_{23}y_2^2 = u_2.$$

Dann ist $\mathbf{A}\mathbf{f}_1 = \begin{bmatrix} 0 \\ \beta_{23} \end{bmatrix}$ vom Rang 1 und damit ist β_{12} identifizierbar, wenn $\beta_{23} \neq 0$; dagegen

ist $\text{rg}(\mathbf{A}\mathbf{f}_2) = 0$, so dass die Angebotsfunktion nicht identifiziert bleibt. Jede Linearkombination des Angebots mit der Nachfrage bleibt auch einer Angebotsfunktion beobachtungsäquivalent, nicht aber umgekehrt. Schätzung mit der Methode der kleinsten Quadrate auf die

¹⁶ allgemein eine Matrix mit einer Spalte für jede Restriktion in der i -ten Gleichung.

¹⁷ ähnlich Malinvaud, Statistical Methods..., S. 557. Wie in Modell B die Angebotsfunktion durch Hinzunahme von z_1 identifizierbar gemacht wurde, so hier die Nachfragefunktion durch Hinzunahme von $(y_2)^2$.

Strukturgleichungen (also OLS¹⁸) ist wegen Abschn. 3.1 gleichwohl nicht zulässig, obgleich damit $\hat{\beta}_{12} \neq \hat{\beta}_{22}$ wären¹⁹.

6. Über- und Unteridentifikation, Multikollinearität

6.1. Interpretation der Überidentifikation der zweiten Gl. (Gl. 6.2) im Modell

$$(6.1) \quad y_1 + \beta_{12}y_2 + \gamma_{11}z_1 + \gamma_{12}z_2 = u_1 \text{ (Nachfrage) [unteridentifiziert] und}$$

$$(6.2) \quad \beta_{21}y_1 + y_2 = u_2 \text{ (Angebot) [überidentifiziert].}$$

Man erhält dann nämlich zwei Schätzungen für β_{21} , nämlich

$$(6.3a) \quad \hat{\beta}_{21}^{(1)} = \frac{-\pi_{21}}{\pi_{11}} \text{ und} \quad (6.3b) \quad \hat{\beta}_{21}^{(2)} = \frac{-\pi_{22}}{\pi_{12}}$$

Mit der indirekten Methode der kleinsten Quadrate (ILS) erhält man die Normalgleichungen (Stichprobenumfang T, Beobachtungen $\tau = 1, 2, \dots, T$)²⁰

$$(6.4) \quad \mathbf{Z}'\mathbf{Z}\mathbf{p}_i = \mathbf{Z}'\mathbf{y}_i$$

$$\text{mit } \mathbf{Z}' = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1T,1} \\ z_{21} & z_{22} & \cdots & z_{2T,1} \end{bmatrix}, \mathbf{p}_i = \begin{bmatrix} \pi_{i1} \\ \pi_{i2} \end{bmatrix} \text{ und } \mathbf{y}_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix}, i = 1, 2,$$

Woraus sich dann auch die genannten 4 Koeffizienten \mathbf{p} ($i = 1, 2$) wie folgt errechnen lassen.

$$(6.5a) \quad -\hat{\beta}_{21}^{(1)} = \frac{\pi_{21}}{\pi_{11}} = \frac{V(y_2)}{V(y_1)} \cdot \frac{r_{y_2z_1} - r_{y_2z_2}r_{z_1z_2}}{r_{y_1z_1} - r_{y_1z_2}r_{z_1z_2}}$$

$$(6.5b) \quad -\hat{\beta}_{21}^{(2)} = \frac{\pi_{22}}{\pi_{12}} = \frac{V(y_2)}{V(y_1)} \cdot \frac{r_{y_2z_2} - r_{y_2z_1}r_{z_1z_2}}{r_{y_1z_2} - r_{y_1z_1}r_{z_1z_2}},$$

also mit Gleichungen, die sich bei unabhängigen Regressoren ($r_{z_1z_2} = 0$, also verschwindender Korrelation zwischen z_1 und z_2) entsprechend vereinfachen. Die beiden Schätzungen stimmen nur dann überein (d.h. die zweite Gl. des Modells, also Gl. 6.2. ist identifizierbar)²¹, wenn dann für die Verhältnisse der partiellen Korrelationskoeffizienten gilt

$$(6.6) \quad \frac{r_{y_1z_2 \cdot z_1}}{r_{y_2z_2 \cdot z_1}} = \frac{r_{y_1z_1 \cdot z_2}}{r_{y_2z_1 \cdot z_2}}.$$

Der Koeffizient $r_{y_1z_2 \cdot z_1}$ ist die Korrelation zwischen y_1 und z_2 "bei Konstanz" von z_1 (wenn z_1 "auspartialisiert" ist). Die Mehrdeutigkeit von β_{21} entsteht also dadurch, dass auch die in Gl. 6.2 ausgelassenen Variablen y_1 und y_2 beeinflussen, und zwar im unterschiedlichen Ausmaß.

¹⁸ ordinary least squares.

¹⁹ Das Theorem von Abschn. 5.5 ist von F.M. Fisher für Modelle des in Abschn. 5.6 vorgestellten Typs verallgemeinert worden, sofern mindestens eine exogene Variable z_i existiert.

²⁰ Man erkennt an der folgenden Gleichung unschwer das Muster der "Normalgleichungen" bei der multiplen Regression.

²¹ In der Vorlesung "Multivariate Analyse" werden im Kap. 3 (= Pfadanalyse) zahlreiche Betrachtungen dieser Art vorgeführt, bei denen sich die Überidentifikation dahingehend ausdrückt, dass dann bestimmte Gleichungen für die (empirischen) Korrelationskoeffizienten gelten müssten (wenn das [Kausal-] Modell gelten soll).

Wäre $r_{y_1 z_2} = r_{y_1 z_1}$ und $r_{y_2 z_2} = r_{y_2 z_1}$, also bestimmte (einfache) Korrelationen gleich, so wäre damit auch Gl. 6.6 erfüllt, d.h.²² es wäre dann

$$\hat{\beta}_{21}^{(1)} = \hat{\beta}_{21}^{(2)} = \frac{\sum y_2 z_2}{\sum y_1 z_2} = \frac{\sum y_2 z_1}{\sum y_1 z_1}.$$

Bei nicht unterscheidbarem Einfluss von z_1 einerseits und z_2 andererseits auf y_1 und y_2 würde sich jedoch das gesamte Modell zum Modell B von Abschn. 2.1 reduzieren, wo (nur) das Angebot identifizierbar ist²³. Man kann weiter zeigen²⁴, dass $\beta_{21}^{(1)}$ und $\beta_{21}^{(2)}$ konsistente, aber nicht erwartungstreue Schätzwerte sind.

6.2. Unteridentifikation und Kollinearität: Für die unteridentifizierte Nachfragekurve im Modell B (Abschn. 2.1) erhält man aus der reduzierten Form nur *eine* Gleichung für b_{12} und γ_{11} nämlich $\gamma_{11} = -\pi_{11} - \pi_{21}\beta_{12}$. Die gleiche Situation tritt ein, wenn man für das Modell C, wo die Nachfrage identifiziert ist, annimmt $z_2 = \delta \cdot y_2$ (lineare Abhängigkeit, Fall der offenen Kollinearität). Man erhält dann nämlich die reduzierte Form mit den beiden Gleichungen

$$\pi_{11} = \frac{\beta_{22}\gamma_{11} + \gamma_{11}\gamma_{22}\delta}{\beta_{12} - \beta_{22} - \gamma_{22}\delta} = \frac{\gamma_{11}(\beta_{22} + \gamma_{22}\delta)}{\beta_{12} - \beta_{22} - \gamma_{22}\delta} \quad \text{und} \quad \pi_{21} = \frac{-\gamma_{11}}{\beta_{12} - \beta_{22} - \gamma_{22}\delta},$$

die sich reduzieren zu nur einer Gleichung, nämlich $\gamma_{11} = -\pi_{11} - \pi_{21}\beta_{12}$ (also der gleichen Gleichung, die man bei Unteridentifikation erhält, siehe oben), als die *einzige* Bestimmungsgleichung für b_{12} und γ_{11} .

Wie man hieran sieht, kann eine lineare Beziehung zwischen b_{ig} und γ_{ik} , die durch unendlich viele Wertetupel befriedigt werden kann (d.h. Unteridentifikation von b_{ig} und γ_{ik}), durch Kollinearität von y_g und z_m (wobei $y_{im} = 0$) entstehen²⁵.

Anhang

1. Beispiel für zwei beobachtungsäquivalente Strukturen eines Modells:

$$\begin{aligned} \text{Modell} \quad & y_1 + \beta_{12}y_2 + \gamma_{11}z = u_1 \\ & y_2 + \beta_{22}y_2 + \gamma_{21}z = u_2 \end{aligned}$$

Mit den beiden folgenden Strukturen wird das Modell gleichermaßen spezifiziert:

Struktur 1 $y_1 - 2y_2 - z = u_1$ und $y_1 + y_2 - 2z = u_2$ und damit $\mathbf{B}_1 = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$, $\mathbf{G}_1 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$ zur weiteren Konkretisierung sei angenommen, dass $T = 4$ Beobachtungen bezüglich der Störgrößen vorliegen: $\mathbf{U}' = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & -2 & 2 & -2 \end{bmatrix}$.

²² bei Variablen y und z als standardisierte Variablen; Summation über $\tau = 1, 2, \dots, T$.

²³ es gilt das dort Gesagte, siehe oben.

²⁴ Valvanis, a.a.O., S. 99f.

²⁵ Weitere Interpretationen vgl. Valvanis a.a.O. (mit Hilfe der Gestalt der Likelihood-Funktion). Danach sind Unteridentifikation und Multikollinearität nur zwei Erscheinungsformen des gleichen statistischen Problems, der "Konfluenz". In beiden Fällen, Unteridentifikation und Kollinearität liegt ein Mangel an unabhängiger Variation der Variablen des Modells vor.

Struktur 2 $y_1 - y_2 - \frac{4}{3}z = w_1$ und $y_1 + \frac{2}{5}y_2 - \frac{9}{5}z = w_2$ somit $\mathbf{B}_2 = \begin{bmatrix} 1 & -1 \\ 1 & 2/5 \end{bmatrix}$, $\mathbf{G}_2 = \begin{bmatrix} -4/3 \\ -9/5 \end{bmatrix}$,

konkrete Werte für w : $\mathbf{W}' = \begin{bmatrix} 4/3 & 0 & 0 & -4/3 \\ 9/5 & -7/5 & 7/5 & -9/5 \end{bmatrix}$, also wieder $T=4$.

Es lohnt sich die im folgenden vorgeführten Berechnungen selbst durchzuführen und nachzuprüfen. Man kann dann leicht erkennen, dass die beiden Strukturen in der Tat zu den gleichen Beobachtungen führen. Man erhält nämlich die gleiche reduzierte Form und auch die gleichen bedingten gemeinsamen Verteilungen $f(y_1, y_2 | z)$. Berechnungen für

Struktur 1: $\mathbf{S}_u = \frac{1}{T} \mathbf{U}' \mathbf{U} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$, $\mathbf{B}_1^{-1} = \begin{bmatrix} 1/3 & 2/3 \\ -1/3 & 1/3 \end{bmatrix}$, $\mathbf{?}_1 = -\mathbf{B}_1^{-1} \mathbf{G}_1 = \begin{bmatrix} 5/3 \\ 1/3 \end{bmatrix}$, so dass man

die folgende reduzierte Form erhält

$y_1 = \frac{5}{3}z + \frac{1}{3}u_1 + \frac{2}{3}u_2$	$y_2 = \frac{1}{3}z - \frac{1}{3}u_1 + \frac{1}{3}u_2$
--	--

Man sieht ferner leicht, dass gilt $\dot{\mathbf{U}} = \mathbf{B}_1^{-1} \mathbf{U}' = \begin{bmatrix} 5/3 & -1 & 1 & -5/3 \\ 1/3 & -1 & 1 & -1/3 \end{bmatrix}$ und $\mathbf{S}_{\dot{u}} = \mathbf{B}_1^{-1} \mathbf{S}_u (\mathbf{B}_1^{-1})' =$

$\begin{bmatrix} 17/9 & 7/9 \\ 7/9 & 5/9 \end{bmatrix}$ für die Varianz-Kovarianzmatrix der Störgrößen der reduzierten Form. Es ist

nun leicht zusehen, welche Werte man für die bedingte (bei gegebenem z) Verteilung von y_1 und y_2 erhält mit unseren angenommenen vier Werten für die Störgrößen:

u_1, u_2	$y_1 = (5/3)z +$	$y_2 = (1/3)z +$	Wahrscheinlichkeit
$u_1 = 1 \quad u_2 = 2$	$5/3$	$1/3$	$1/4$
$u_1 = 1 \quad u_2 = -2$	-1	-1	$1/4$
$u_1 = -1 \quad u_2 = 2$	$+1$	$+1$	$1/4$
$u_1 = -1 \quad u_2 = -2$	$-5/3$	$-1/3$	$1/4$

Die reduzierte Form ist $\mathbf{Y}' = \mathbf{?}_1 \mathbf{Z}' + \dot{\mathbf{U}}'$ und es ist interessant zu sehen, dass man für Struktur 2 genau die gleiche Form erhält und insbesondere $\mathbf{?}_2 = \mathbf{?}_1$ ist und auch $\dot{\mathbf{U}}$ auftritt, so dass man die gleichen "Daten" (das gleiche Streudiagramm) erhält.

Struktur 2: $\mathbf{S}_w = \frac{1}{T} \mathbf{W}' \mathbf{W} = \begin{bmatrix} 8/9 & 6/5 \\ 6/5 & 13/5 \end{bmatrix}$, $\mathbf{B}_2^{-1} = \begin{bmatrix} 2/7 & 5/7 \\ -5/7 & 5/7 \end{bmatrix}$, $\mathbf{?}_2 = -\mathbf{B}_2^{-1} \mathbf{G}_2 = \begin{bmatrix} 5/3 \\ 1/3 \end{bmatrix} = \mathbf{?}_1$

und damit

$y_1 = \frac{5}{3}z + \frac{2}{7}w_1 + \frac{5}{7}w_2$	$y_2 = \frac{1}{3}z - \frac{5}{7}w_1 + \frac{5}{7}w_2$
--	--

Ferner ist $\dot{\mathbf{W}} = -\mathbf{B}_2^{-1} \mathbf{W}' = \dot{\mathbf{U}}$, so dass auch gilt

$$\mathbf{Y}' = \mathbf{?}_1 \mathbf{Z}' + \dot{\mathbf{U}}' = \mathbf{?}_2 \mathbf{Z}' + \dot{\mathbf{W}}'$$

so dass keine Möglichkeit existiert um von dieser *einen* reduzierten Form auf *zwei* verschiedene Strukturen zu schließen, etwa auf $\mathbf{B}_1 = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \neq \mathbf{B}_2 = \begin{bmatrix} 1 & -1 \\ 1 & 2/5 \end{bmatrix}$.

2. Die Matrix $P_{K-k,g}$ in den Modellen des Abschn. 2.1 (Marktmodelle)

Modell	Nachfrage (Gleichung 1)	Angebot (Gleichung 2)
B (K = 1)	$K-k = 0$, so dass $?_{K-k,g}$ leer ist; Rang $0 < g-1$ also <i>nicht</i> identifizierbar	$?_{K-k,g} = \begin{bmatrix} -\beta_{22}\gamma_{11} & \gamma_{11} \\ \beta_{22} - \beta_{12} & \beta_{22} - \beta_{12} \end{bmatrix}$
C (K = 2)	$?_{K-k,g} = \begin{bmatrix} -\beta_{12}\gamma_{22} & -\gamma_{22} \\ \beta_{22} - \beta_{12} & \beta_{22} - \beta_{12} \end{bmatrix}$ *)	$?_{K-k,g} = \begin{bmatrix} -\beta_{22}\gamma_{11} & \gamma_{11} \\ \beta_{22} - \beta_{12} & \beta_{22} - \beta_{12} \end{bmatrix}$ *)
D (K = 3)	wie im Modell C	$K-k = 3-1 = 2 > g-1 = 1$, Matrix P ist jetzt $\begin{bmatrix} -\beta_{22}\gamma_{11} & \gamma_{11} \\ \beta_{22} - \beta_{12} & \beta_{22} - \beta_{12} \\ -\beta_{22}\gamma_{13} & \gamma_{13} \\ \beta_{22} - \beta_{12} & \beta_{22} - \beta_{12} \end{bmatrix}$ wie man leicht sieht ist die Determinante dieser Matrix 0, so dass der Rang 1 ($< K-k$) ist und die Angebotsfunktion überidentifiziert ist.

* Rang = 1, also exakt identifiziert.

3. Weitere Bemerkungen zur "Haavelmo Bias": die Demonstration der Zusammenhänge am Beispiel interdependenter Modelle nach Art von Abschn. 2.1 (Marktmodelle), in dem auch auf die "simultaneous equation bias" hingewiesen wurde sollte nicht zu dem Fehlschluß verleiten, dass dieses Schätzproblem (bei Interdependenz) und das Identifikationsproblem das gleiche Problem seien²⁶. Es mag deshalb nützlich sein, noch einmal kurz auf das von Haavelmo aufgeworfene Problem einzugehen²⁷.

Gegeben sei das (allerdings unter identifizierte) Modell²⁸

(1) $Y = aX + e_1$

(2) $X = bY + e_2$

oder ein Marktmodell mit $p = b_1q + u$ (Nachfrage) und $p = b_2q + v$ (Angebot), was äquivalent ist, wenn $p = y$, $q = x$, $b_1 = a$ und $b_2 = 1/b$. Gl. 1 und 2 aufgelöst nach den endogenen Variablen ergibt

(1a) $Y = \frac{ae_2 + e_1}{1 - ab}$ und (2a) $X = \frac{be_1 + e_2}{1 - ab}$

Daraus erhält man folgende Varianzen und Kovarianzen (wenn $\sigma_i^2 = V(e_i)$, $\sigma_{12} = C(e_1e_2)$, und $E(e_i) = 0$, $i = 1, 2$)

$V(X) = (b^2\sigma_1^2 + 2b\sigma_{12} + \sigma_2^2)/(1 - ab)^2$

$V(Y) = (a^2\sigma_2^2 + 2a\sigma_{12} + \sigma_1^2)/(1 - ab)^2$

$C(XY) = [a\sigma_2^2 + b\sigma_1^2 + (1 + ab)\sigma_{12}]/(1 - ab)^2$.

²⁶ Beim Identifikationsproblem geht es nicht um ein Schätzproblem. Im Falle der Unteridentifikation eines Modells kann man mit keinem Trick die Parameter schätzen und bei Überidentifikation hat man eine Wahl zwischen verschiedenen möglichen Schätzungen der gleichen Parameter.

²⁷ Wir demonstrieren es allerdings nicht mit einem makroökonomischen Modell (Konsumfunktion und einer Funktion mit von ΔC abhängigen Investitionen) und Störgrößen x und y sondern mit einem Marktmodell.

²⁸ Wir sehen davon ab, dass man für die Zufallsvariable e_1 und e_2 eigentlich große Buchstaben verwenden sollte, also E_1 und E_2 . Das gilt natürlich auch in anderen Zusammenhängen für die Störgrößen u und v . Die Symbole u und v sind Schreibweisen, die sich jedoch einmal in dieser Art eingebürgert haben.

Man sieht sofort, dass die OLS Schätzer für a und b von Gl. 1 und 2 in der Regel nicht übereinstimmen mit a und b , es sei denn man macht spezielle Annahmen bezüglich der Varianzen der Störgrößen und deren Kovarianz σ_{12} . Man erhält nämlich

$$(3a) \quad \hat{a} = C(XY)/V(X) \text{ und}$$

$$(3b) \quad \hat{b} = C(XY)/V(Y).$$

Selbst $\sigma_{12} = 0$ bedeutet nicht, dass

$$(4) \quad \hat{a} = (a\sigma_2^2 + b\sigma_1^2)/(b^2\sigma_1^2 + \sigma_2^2)$$

gleich a ist. Setzt man $\hat{a} = a$ so erhält man die Bedingung $ab = 1$ was gleich bedeutend ist mit $r_{xy} = 1$. Der "single equation approach" (OLS) liefert also nur in einem sehr speziellen Fall unverzerrte Schätzwerte für a (und, analog zu zeigen, für b).

Zur Begründung kann man auch wie folgt vorgehen. Multipliziert man Gl. 1 mit X oder Gl. 2a mit e_1 und bildet man Erwartungswerte, so erhält man mit

$$(5a) \quad E(Xe_1) = \frac{b\sigma_1^2 + \sigma_{12}}{1 - ab}$$

einen Ausdruck, der im allgemeinen $\neq 0$ ist es sei denn $\sigma_1^2 = \sigma_{12} = 0$. Das gleiche gilt für

$$(5b) \quad E(Ye_2) = \frac{a\sigma_2^2 + \sigma_{12}}{1 - ab}$$

was ebenfalls von Null verschieden ist, es sei denn $\sigma_{12} = \sigma_2^2 = 0$. Das Verschwinden einer Varianz (etwa $\sigma_1^2 = 0$) wird üblicherweise so interpretiert: man kann eine Funktion (etwa die Nachfragekurve) schätzen, wenn diese konstant ist ($(\sigma_1^2 = \sigma_u^2 = 0)$) und die beobachteten Punkte des Streudiagramms allein durch Verschiebungen der anderen Funktionen (der Angebotskurve, so dass $\sigma_v^2 > 0$) zustande gekommen sind. Dass die Störgröße e_1 bzw. e_2 vom Regressor X bzw. Y *nicht* unabhängig ist bewirkt, dass die Schätzung jeder einzelnen Gleichung mit der gewöhnlichen Methode der kleinsten Quadrate (OLS) nicht zulässig ist. Die Schätzung von a mit \hat{a} (Gl. 3a) bzw. b mit \hat{b} (Gl. 3b) ist damit nicht unverzerrt. Denn aus Gl. 3a und Gl. 5a folgt

$$E(\hat{a}) = a + \frac{(b\sigma_1^2 + \sigma_{12})(1 - ab)}{b^2\sigma_1^2 + 2b\sigma_{12} + \sigma_2^2} = a + \frac{E(Xe_1)}{V(X)}$$

was deutlich zeigt, dass die Ursache der "Haavelmo Bias" in $E(Xe_1) \neq 0$ zu sehen ist. Entsprechend erhält man

$$E(\hat{b}) = b + \frac{E(Ye_2)}{V(Y)}.$$

Man beachte, dass zwar a und b nicht erwartungstreu geschätzt werden wohl aber der bedingte Erwartungswert mit $E(Y|X=x) = \hat{y} = ax$ und $E(X|y) = \hat{x} = by$ denn es gilt $E(xe_1) = xE(e_1) = x \cdot 0 = 0$ und entsprechend $E(ye_2) = 0$

Das folgende Modell ist leicht erkennbar *exakt* identifiziert und gleichwohl würde bei OLS der Haavelmo Bias entstehen²⁹

²⁹ Auch hier könnte es vielleicht besser sein die Störgröße U statt u zu nennen.

$$(6) \quad C = a + bY + u$$

$$(7) \quad Y = C + I.$$

die isolierte Schätzung von Gl. 6 mit OLS, also $\hat{b}_{OLS} = \frac{C(CY)}{V(Y)}$ ist i. d. R. positiv verzerrt

$E(\hat{b}_{OLS}) > b$, so wie \hat{a} negativ verzerrt ist. Die reduzierte Form ist in diesem Modell

$$C = \frac{a}{1-b} + \frac{b}{1-b}I + \frac{1}{1-b}u = \alpha + \beta I + u^*$$

$$Y = \frac{a}{1-b} + \frac{1}{1-b}I + \frac{1}{1-b}u = \alpha + \gamma I + u^*$$

Man erkennt, dass man b schätzen kann mit $\hat{b} = \hat{\beta} / \hat{\gamma}$ und auch für a einen und nur einen Schätzwert erhält, so dass das Modell exakt identifiziert ist. Außerdem sieht man leicht, dass

$$E(Yu) = \frac{1}{1-b} V(U) = \frac{1}{1-b} \sigma_u^2 \neq 0 \text{ ist, so dass man nicht einfach } a \text{ und } b \text{ mit einer isolierten}$$

Schätzung von Gl. 6 bestimmen sollte, denn

$$(8) \quad \hat{b}_{OLS} = b + \frac{(1-b)\sigma_u^2}{\sigma_I^2 + \sigma_u^2} = b + B$$

wobei wegen $0 < b < 1$ (b ist ja die marginale Konsumquote) der Bias offenbar positiv ist ($B > 0$). Die Schätzung von b aus der reduzierten Form (also mit der indirekten Methode der kleinsten Quadrate, (ILS), d. h.

$$(9) \quad \hat{b}_{ILS} = \frac{\hat{\beta}}{\hat{\gamma}} = \frac{C(CI)}{C(YI)}$$

ist dagegen unverzerrt (unbiased).

Man beachte, dass wegen Gl. 7 notwendig gilt $C(YI) = C(CI) + V(I)$, so dass auch stets gilt $\gamma - \beta = 1$ (nicht nur $\hat{\gamma} - \hat{\beta} = 1$).

Wendet man I als Instrumentalvariable in Gl. 6 an, so erhält man den gleichen Schätzwert \hat{b} wie bei ILS, da nämlich $C(Iu) = 0$, was übrigens auch der Grund ist für die Erwartungstreue des Schätzens der Gl. 9, also $E(\hat{b}_{ILS}) = b$.

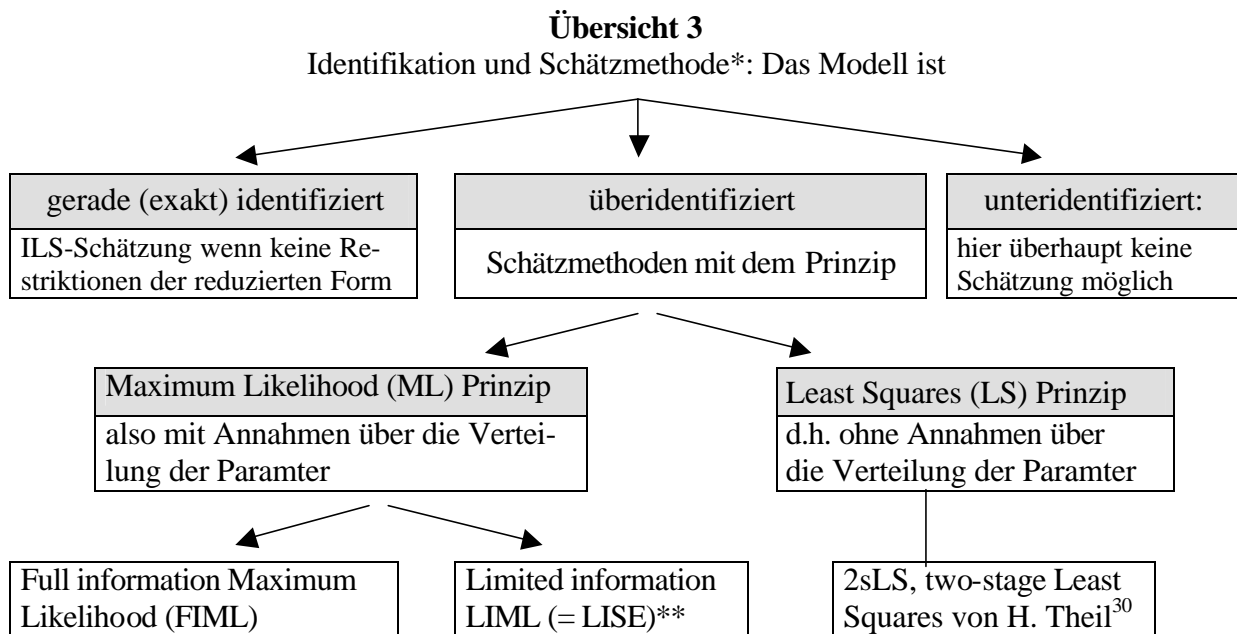
4. Schätzmethoden und directional least squares (DLS)

Man kann den ILS Schätzer nicht nur mit der Methode der Instrumentalvariablen interpretieren, sondern auch als Anwendung von directional least squares (DLS).

Bei der gewöhnlichen Methode der kleinsten Quadrate minimiert man senkrechte Abstände (u_i in Abb. 1), bei der orthogonalen Regression lotrechte (Lot auf die gesuchte Regressionsgerade) Abstände p_i und bei DLS die Abstände d_i , (mit 45° Linien wie in Abb. 1) die mit u_i wie folgt zusammenhängen

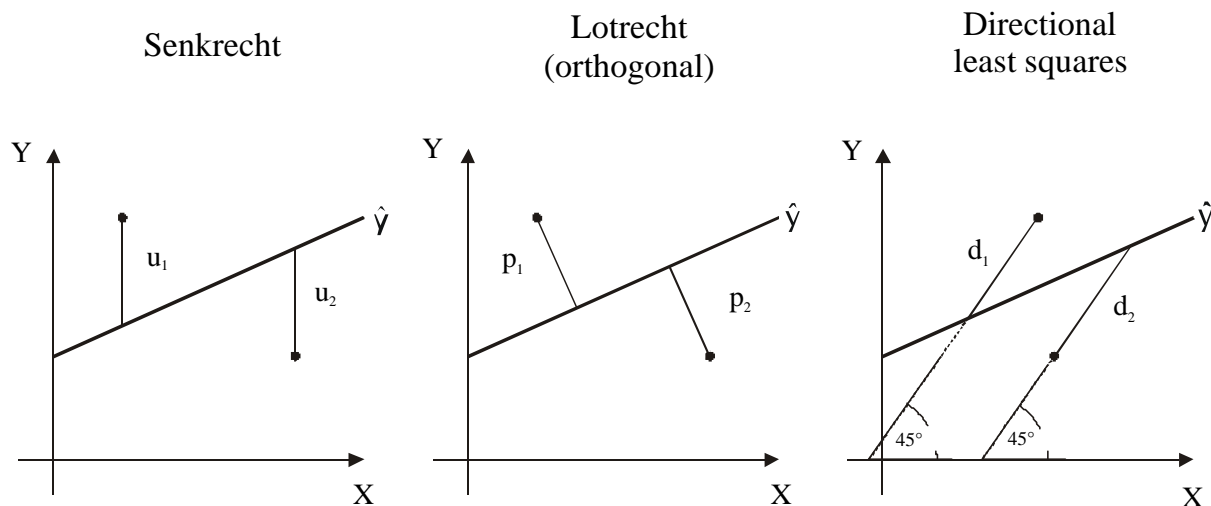
$$p_i = \frac{\sqrt{2}}{1-b} u_i \text{ da } 1/\sqrt{2} = \sin(45^\circ), \text{ bei der Quadrierung gilt dann } \sum p_i^2 = 2 \sum (u_i^*)^2$$

wobei u_i^* die Störgröße in der reduzierten Form ist.



* Nur die wichtigsten Schätzmethoden sind erwähnt. Full information method = mit Berücksichtigung aller der reduzierten Form auferlegten (für sie geltenden) identifizierenden Restriktionen und Festlegungen
 ** limited information single equation method von Anderson und Rubin

Abbildung 1: Unterschiedliche Arten, die Summe quadrierter Abstände zu minimieren



³⁰ Zu einer Interpretation dieser Methode mit der Pfadanalyse vgl. Vorlesung "Multivariate Analyse" im Hauptstudium Statistik.