Gauss-Markoff, Trendbereinigung, wahre und missspezifizierte Modelle, und Modellwahl

Kapitel 6

Ökonometrie I
Michael Hauser

Inhalt

- Gauss-Markoff
- Frisch-Waugh
- Detrending
- Missspezifikation
  - relevante Variable fehlt
  - Einschluss von irrelevanten Variablen
  - Endogenität in Systemen
- Modellspezifikation, Informationskriterien

Inhalt

- Das Gauss-Markoff Theorem fasst die statistischen Eigenschaften des OLS Schätzers zusammen.
- Trendbereinigung wird verwendet, wenn 2 Variable y und x mit einem Trend behaftet sind, dh über die Zeit steigen, und ihr Zusammenhang modelliert werden soll.

$$y = \beta_0 + \beta_1 x + u$$

  Meist erhält man bei Regressionen mit trendbehafteten Reihen hohe R^2, die den kontemporären Zusammenhang überbewerten. Um dies zu vermeiden regressiert man die trendbereinigten Variablen aufeinander.

Inhalt

- Ein brauchbares Modell soll alle relevanten Variablen enthalten. Umgekehrt, sollen keine irrelevanten Variablen vorkommen.
  Wir untersuchen im Abschnitt Missspezifikation, welche Eigenschaften der Schätzer verliert, wenn entweder zu viele, oder zu wenige Variable im Modell eingeschlossen werden.
- Oft schlagen verschiedene Theorien unterschiedliche Spezifikationsvarianten vor. Wir geben einfache Regeln an, wie man ein brauchbares Modell in kurzer Zeit finden kann.

Gauss-Markoff

Das Gauss-Markoff Theorem fasst die statistischen Eigenschaften des OLS Schätzers zusammen.

Gauss-Markoff

Wir gehen von dem linearen multiple Regressionsmodell mit n Beobachtungen aus, das wir als korrekt spezifiziert ansehen.

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$$

oder in Matrixschreibweise

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$$

$\mathbf{y}_{n\times 1}$, $\mathbf{X}_{n\times(k+1)}$, $\beta_{(k+1)\times 1}$, $\mathbf{u}_{n\times 1}$

Wir nehmen an,

- die Fehler sind iid $(0, \sigma^2)$
- $\mathbf{x}$'s sind nicht-stochastisch, und (daher) unabhängig von den u's.
- die $\mathbf{x}$'s sind linear unabhängig. Dh $\text{Rank}(\mathbf{X}'\mathbf{X}) = \text{Rank}(\mathbf{X}) = k + 1$. Damit existiert $(\mathbf{X}'\mathbf{X})^{-1}$.

Gauss-Markoff

Unter diesen Annahmen erhält man den

- besten (Minimum-Varianz),
- unverzerrten
- linearen Schätzer

BLUE (best, linear, unbiased estimator) von β, wenn man die Fehlerquadratsumme minimiert.

$$Q = \mathbf{u}'\mathbf{u} = (\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta)$$

Dieses Theorem wird **Gauss-Markoff Theorem** genannt.

Gauss-Markoff

Bem:

- Da sich der OLS Schätzer durch die Minimierung von Q ergibt, besitzt er die BLUE Eigenschaften.
- Für stochastische Regressoren gilt das Theorem (nur) asymptotisch ($n \to \infty$).
- Die Eigenschaft, dass der Schätzer linear ist, bezieht sich darauf, dass alle geschätzten Koeffizienten b_i als gewichtete *Summe* der Beobachtungen y_i dargestellt werden können. (Also nicht direkt auf die Linearität des Modells.)

Bem: Stochastische Regressoren liegen vor, wenn die Werte der unabhängigen Variablen selbst aus eine Stichprobe stammen.

Partitionierte Regression I

Partitionierte Regression

Partitioniert bedeutet, dass wir die erklärenden Variablen in 2 Blöcke teilen. Die einfachste Variante hat einen "Block" x_1 und einen "Block" x_2.

$$y = \beta_1 x_1 + \beta_2 x_2 + u$$

β_1 können wir - wie wir wissen - ceteris paribus interpretieren. Dh wir halten x_2 konstant und verändern $x_1 \rightarrow (x_1 + 1)$, dann steigt im Durchschnitt y um β_1.

Frisch-Waugh Theorem

Ein Verallgemeinerung der ceteris paribus Überlegung erhält man, wenn man nur die Regression y auf x_1 betrachtet *nachdem* der Effekt der Variablen x_2 aus beiden Variablen y *und* x_1 extrahiert wurde.

- Schritt 1:
  Man regressiert dazu die Variablen

  $$y \text{ auf } x_2 \qquad \text{und} \qquad x_1 \text{ auf } x_2.$$

  $$y = \gamma_y x_2 + v_y \qquad \text{und} \qquad x_1 = \gamma_1 x_2 + v_1$$

  Der Effekt von
  - y, der nicht von x_2 erklärt werden kann, befindet sich in $v_y = y - (y|x_2)$.
  - x_1, der nicht von x_2 erklärt werden kann, in $v_1 = x_1 - (x_1|x_2)$.

Frisch-Waugh Theorem

- Schritt 2:
  Regressieren wir nun die beiden Fehler v_y und v_1 aufeinander

$$v_y = \delta v_1 + w$$

so finden wir

$$\delta = \beta_1$$

Frisch-Waugh-Theorem:
Der Parameter δ aus dem 2-Schritt Verfahren ist **gleich** dem Parameter β_1 in

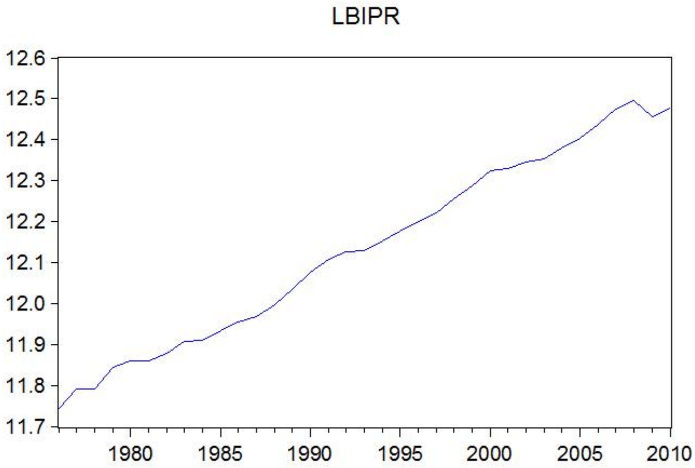
$$y = \beta_1 x_1 + \beta_2 x_2 + u$$

Bem: Statt $v_y = \delta v_1 + w$ kann man auch $[y - (y|x_2)] = \delta[x_1 - (x_1|x_2)] + w$, schreiben.

Trendbereinigung

Trendbereinigung und Konjukturzyklus

Eine trendbehaftete Reihe: $\log(BIP^R)$



$\log(BIP^R)$ $(LBIPR)$ ... BIP real, Ö

Schätzen eines linearen Trends für $\log(BIP^R)$

Wir wollen nun die Reihe $\log(BIP_t^R)$ um ihren Trend bereinigen. Dazu regressieren wir $\log(BIP^R)$ auf die Zeit bzw den Trend.

$$\log(BIP_t^R) = \alpha_0 + \alpha_1 t + u_t$$

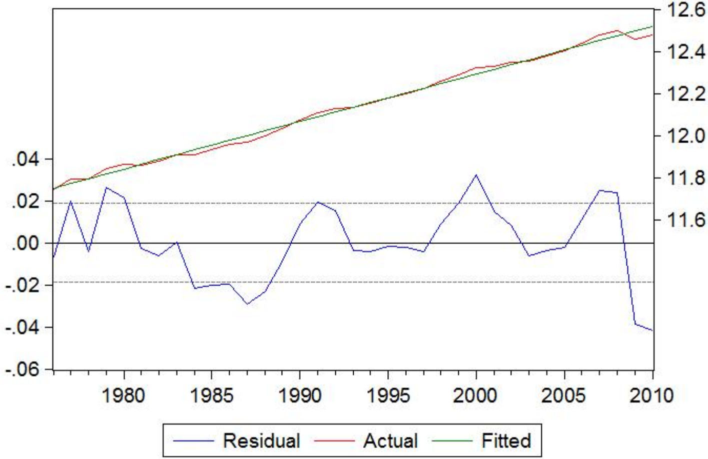
$$\begin{array}{rll} \widehat{\log(BIP^R)} = & \underset{(0.014)}{11.16} + \underset{(0.001)}{0.023}\, t & \qquad R^2 = 0.994 \\ & & \qquad RSS = 0.0115 \end{array}$$

Eine Trend-Variable erzeugt man in EViews durch

```
series t = @trend
```

t hat hier die Werte $0, 1, 2, 3, \ldots$.

Linearer Trend: $\log(BIP_t^R) = \alpha_0 + \alpha_1\, t + u_t$



$\log(BIP^R) \ldots$ BIP real, Ö

Trendbereinigte Reihe

Die Residuen der Regression sind die um einen linearen Trend bereinigte Reihe von $\log(BIP^R)$, die **trendbereinigte Reihe** $\log(BIP_t^R)^{tb}$.
Sie enthält keinen Trend, aber alle anderen nicht-trendbehafteten Komponenten.

$$\log(BIP_t^R)^{tb} = \widehat{u}_t = \log(BIP_t^R) - (\widehat{\alpha}_0 + \widehat{\alpha}_1\, t)$$

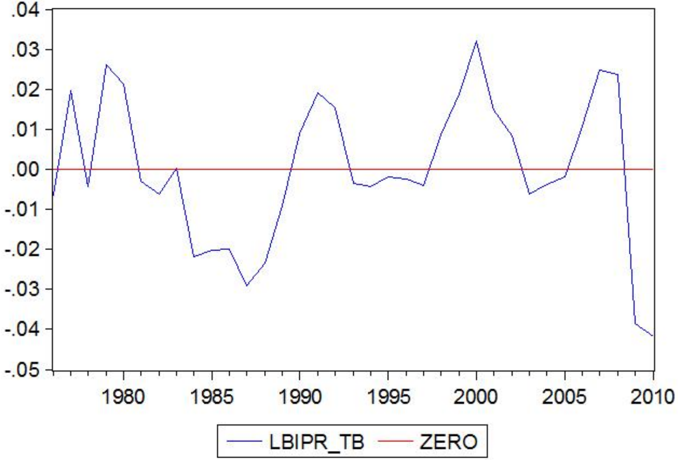
Es können die Perioden mit

- positiven Werten, dh die über dem Trend liegen, als *Hochkonjunktur*, und mit
- negativen Werten, dh die unter dem Trend liegen, als *Rezession, Depression*,
- und die dazwischen liegenden als *Aufschwung*-, und *Abschwungphasen*

beschrieben werden.

Bem: Ein Konjukturzyklus dauert mindestens 1 1/2 Jahre.

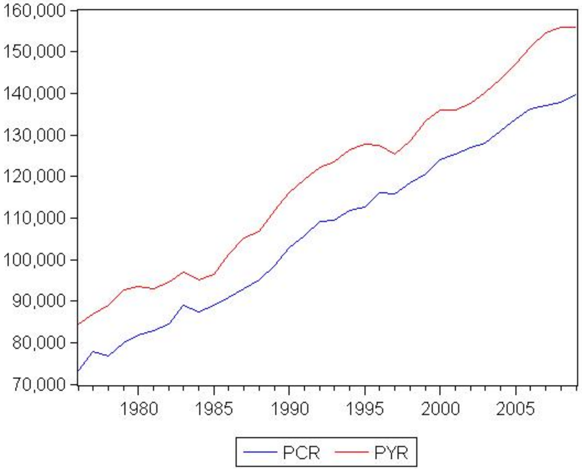
Der österreichische "Konjunkturzyklus", $\log(BIP_t^R)^{tb}$



Trendbereinigung

Regression mit trendbehafteten Variablen

Zwei trendbehaftete Reihen: C_t, Y_t^d



C_t (*PCR*) privater Konsum real, Y_t^d (*PYR*) disponibles Eink real

Modellieren von C_t und Y_t^d als lineare Trends

Die Schätzungen der Trend-Modelle ergeben

$$\widehat{C} = \underset{(526)}{72928} + \underset{(27.4)}{2073}\, t \qquad R^2 = 0.994, \quad RSS = 7.86\,07$$

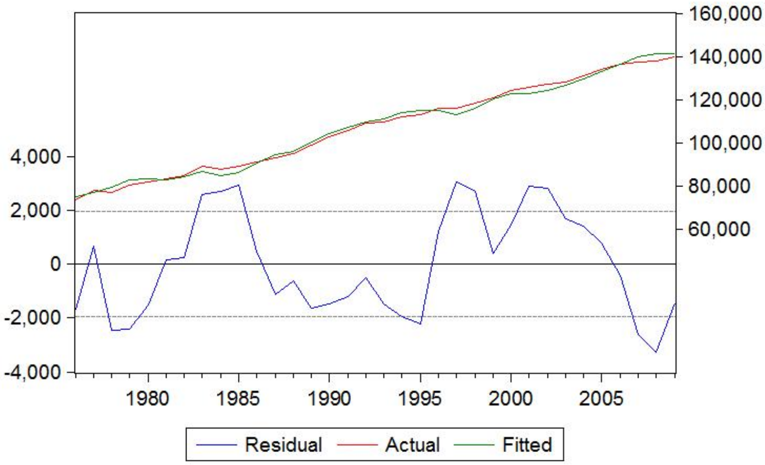
$$\widehat{Y^d} = \underset{(928)}{82602} + \underset{(48.4)}{2220}\, t \qquad R^2 = 0.985, \quad RSS = 2.45\,08$$

Beide R^2 sind sehr hoch, 0.994 und 0.985.

Wir untersuchen nun den Einfluss des Trends in beiden Variablen auf die Schätzung der Konsumfunktion

$$C_t = \beta_0 + \beta_1 Y_t^d + u_t$$

Konsumfunktion $C_t = \beta_0 + \beta_1 Y_t^d + u_t$



Konsumfunktion $C_t = \beta_0 + \beta_1 Y_t^d + u_t$

Es ist zu erwarten, dass der Zusammenhang zwischen C und Y^d sehr stark ist, da sich beide zu einem großen Ausmass durch den Trend beschreiben lassen.

$$\widehat{C}_t = \underset{(1844)}{-3191} + \underset{(0.015)}{0.925}\, Y_t^d \qquad R^2 = 0.991 \qquad RSS = 1.21\,08$$

Das R^2 ist - wie erwartet - sehr hoch. Es ist 0.991.

Konsumfunktion in den trendbereinigten Reihen

Nun überlegen wir uns, wie der stark der Zusammenhang zwischen C und Y^d ohne den Einfluss der Trends ist. Dazu verwenden wir das Frisch-Waugh Theorem an, und bereinigen C und Y^d um den Trend, und regressieren die trendbereinigten Variablen aufeinander.

Die trendbereinigten Variablen v_c und v_y für C und Y^d ergeben sich als

$$C = \alpha_0 + \alpha_1 t + v_c \qquad \text{sodass} \qquad v_c = C - (\alpha_0 + \alpha_1 t) = C_{tb}$$

$$Y^d = \gamma_0 + \gamma_1 t + v_y \qquad \text{sodass} \qquad v_y = Y^d - (\gamma_0 + \gamma_1 t) = Y^d_{tb}$$

Konsumfunktion in den trendbereinigten Reihen

Die Konsumfunktion in den trendbereinigten Variablen lautet

$$C_{tb} = \delta_0 + \delta_1 Y_{tb}^d + w$$

Die Schätzung ergibt

$$\begin{array}{llll} \widehat{C}_{tb} = & -1.65 \cdot 10^{-11} & + 0.371\, Y_{tb}^d & \qquad R^2 = 0.428 \\ & \quad (203) & \quad (0.076) & \qquad RSS = 4.49\,07 \end{array}$$

Sie zeigt ein vergleichsweises kleines R^2 von 0.428.
(Der Interzept ist nicht notwendig, da wir eigentlich um Trend *und Konstante* bereinigt haben.)

Das hohe R^2 der Regression der ursprünglichen Variablen von 0.991 *entsteht durch den gemeinsamen Trend und täuscht.*

Konsumfunktion in den trendbereinigten Reihen

δ_1 kann man c.p. als den Effekt auf 'die Abweichung von C_t von seinem langfristigen Trend' interpretieren, wenn sich die 'Abweichung von Y_t^d von seinem langfristigen Trend' um eine Einheit erhöht.

Frisch-Waugh und Trend

Konsumgleichung mit Trend

Zum Vergleich führen wir den Trend *explizit* in der Konsumgleichung ein

$$C = \beta_0 + \beta_1 Y^d + \beta_2 t + u$$

Wir erhalten für Österreich 1976 - 2009

$$\begin{array}{llll} \widehat{C} = & 42307 + 0.371\, Y^d + 1250\, t & \qquad R^2 & = 0.997 \\ & (6366) \quad (0.077) \qquad (172.1) & \qquad RSS & = 4.49\,07 \end{array}$$

und finden, dass der Koeffizient vor Y^d identisch mit dem von oben aus der Regression in den trendbereinigten Variablen ist.

Vergleich mit Regression trendbereinigter Variabler

- Anstelle der Bereinigung beider Variablen um den Trend, schätzt man direkt

$$C_t = \beta_0 + \beta_1 Y_t^d + \beta_2 t + u_t$$

- Den Erklärungsbeitrag des Trends ermittelt man durch den Vergleich des
  - 'R^2 aus dem Modell mit Trend' und dem
  - 'R^2 aus der Regression der trendbereinigten Variablen'.

$$0.997 - 0.428 = 0.571$$

Der Trend erklärt 57.1%, während das Einkommen nur 42.8% der Varianz des Konsums erklärt.

Exkurs: Bereinigung um eine Konstante

Gehen wir von einem Modell mit Interzept aus

$$y = \beta_0 + \beta_1 x + u$$

Wir bereinigen beide Variable, y und x um die Konstante.

$$y = \gamma_y + v_y \qquad \text{und} \qquad x = \gamma_x + v_x$$

So gilt

$$v_y = \beta_1 v_x + w$$

Die Regression der *mittelwert-bereinigten* Daten ohne Interzept ergibt die gleichen Koeffizienten wie die der Regression mit Interzept für die Originaldaten.

Partitionierte Regression II

Regressionsmodell

Kehren wir zum Modell

$$y = \beta_1 x_1 + \beta_2 x_2 + u$$

zurück. Der OLS Schätzer in Matrixform ist

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Multipliziert man mit der $(\mathbf{X}'\mathbf{X})$ von links, erhält man die **Normalgleichungen** in Matrixform.

$$(\mathbf{X}'\mathbf{X})\,\mathbf{b} = \mathbf{X}'\mathbf{y}$$

Frisch-Waugh

In partitionierter Schreibweise

$$\begin{bmatrix} \mathbf{x}_1'\mathbf{x}_1 & \mathbf{x}_1'\mathbf{x}_2 \\ \mathbf{x}_2'\mathbf{x}_1 & \mathbf{x}_2'\mathbf{x}_2 \end{bmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1'\mathbf{y} \\ \mathbf{x}_2'\mathbf{y} \end{pmatrix}$$

Wir multiplizieren aus und erhalten für b_1

$$b_1 = (\mathbf{x}_1'\mathbf{x}_1)^{-1}\mathbf{x}_1'\mathbf{y} - (\mathbf{x}_1'\mathbf{x}_1)^{-1}\mathbf{x}_1'\mathbf{x}_2 b_2 = (\mathbf{x}_1'\mathbf{x}_1)^{-1}\mathbf{x}_1'(\mathbf{y} - \mathbf{x}_2 b_2)$$

b_1 *hängt von den Daten x_2 und b_2 ab*. (Vgl Übung 3.)

Der Erwartungswert von b_1 hängt von x_2 und β_2 ab.

$$\mathrm{E}(b_1) = (\mathbf{x}_1'\mathbf{x}_1)^{-1}\mathbf{x}_1'(\mu_y - \mathbf{x}_2\beta_2)$$

Ann: $\mathrm{E}(y) = \mu_y$, und $\mathbf{x}$ sei nicht-stochastisch.

Seltener Fall: orthogonale Regressoren

Ist $\mathbf{x}_1'\mathbf{x}_2 = 0$, so *verschwindet* die Abhängigkeit von x_2. Die Regressoren sind dann orthogonale Regressoren, bzw x_1 und x_2 sind unkorreliert.

- In diesem Fall kann man statt einer multiplen Regression 2 separate bivariate rechnen

$$y = \beta_1 x_1 + v_1 \qquad \text{und} \qquad y = \beta_2 x_2 + v_2$$

  und erhält die gleichen Koeffizienten b_1 und b_2 wie in

$$y = \beta_1 x_1 + \beta_2 x_2 + u$$

Meist sind die Regressoren **nicht** orthogonal.

Missspezifikation

Wir vergleichen ein *'wahres'* Modell mit einem

- Modell in dem eine Variable *fehlt* (omitted variable bias),
  - ein Kommentar von JM Keynes dazu befindet sich im Anhang -
  und mit einem
- Modell in dem eine Variable *zu viel* ist.

Welche Auswirkungen haben Fehlspezifikationen auf die geschätzten Koeffizienten, und welche auf die Prognosevarianz?

Weiters geben wir ein Beispiel für den simultaneous equation bias.

Missspezifikation: relevante Variable fehlt

Angenommen das *wahre* Modell [W] ist

$$y = \beta_1 x_1 + \beta_2 x_2 + u$$

wir schätzen hingegen das *falsch* spezifizierte [F]

$$y = \beta_1 x_1 + u$$

Eine relevante Variable (x_2) fehlt.

Der OLS Schätzer von β_1 aus dem wahren Modell, b_1^W, ist oben (Folie 33) gegeben und hängt von x_2 und b_2 ab.

Missspezifikation: relevante Variable fehlt

Der OLS Schätzer von β_1 aus dem falschen Modell ist

$$b_1^F = (\mathbf{x_1}'\mathbf{x_1})^{-1}\mathbf{x_1}'\mathbf{y}$$

Daher ist b_1^F aus dem falschen Modell verzerrt

$$\mathrm{E}(b_1^W) \neq \mathrm{E}(b_1^F)$$

(außer x_1 und x_2 sind orthogonal/unkorreliert, oder $\beta_2 = 0$).

Die Nicht-Berücksichtigung relevanter Variabler führt zu verzerrten Schätzern (hier von β_1). Die Verzerrung heißt **omitted variable bias.**

In der Folge ist auch die Prognose verzerrt.

Missspezifikation: Einschluss irrelevanter Variabler

Angenommen das *wahre* Modell [W] ist

$$y = \beta_1 x_1 + u$$

wir schätzen hingegen das *falsch* spezifizierte [F]

$$y = \beta_1 x_1 + \beta_2 x_2 + u$$

Eine irrelevante Variable (x_2) ist miteingeschlossen.

- Da $\beta_2 = 0$, können wir unverzerrt schätzen.
- Als Schätzer erhalten wir immer einen Wert für $b_2 \neq 0$. Ein t-Test (zB mit $\alpha = 0.05$) zeigt jedoch nur in 95% der Fälle an, dass der Koeffizient b_2 “nicht signifikant von null verschieden” ist.

Missspezifikation: Einschluss irrelevanter Variabler

- Führt man mit dem *zu großen Modell* eine Prognose durch, erhält man eine *größere Prognosevarianz (-unsicherheit) als notwendig*.
  (Da b_2 im Durchschnitt zwar 0 ist, aber in der Stichprobe $\neq 0$ ist.)

Als Konsequenz könnte man für die Prognose alle Variable aus einem Modell entfernen, von denen man nicht überzeugt ist, dass sie einen relevanten Beitrag leisten.

Missspezifikation: Endogenität

Das Problem der Endogenität tritt in Systemen von Gleichungen auf. Die resultierende Verzerrung heißt **simultaneous equation bias**.

Bsp:
In diesem sehr einfachen 2-Gleichungs-Makromodell des realen Sektors

$$C = \beta Y + u \qquad \text{und} \qquad Y = C + I$$

$C \ldots$ priv Konsum, $Y \ldots$ Einkommen, $I \ldots$ Investitionen
seien die Investitionen exogen.

Da $Y = C + I$, und u der nicht-erklärte Teil von C ist, sind Y und u - entgegen den MLR-Annahmen - *korreliert*.

$$Y = C + I = C(u) + I = Y(u)$$

Missspezifikation: Endogenität

Der Ein-Gleichungs-OLS Schätzer der Konsumgleichung gibt (dh ohne Berücksichtigung von $Y = C + I$) für b

$$b = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{C}$$

Setzen wir hier für C die Beziehung $C = \beta Y + u$ ein.

$$b = (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'(\mathbf{Y}\beta + \mathbf{u}) = \beta + (\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{u}$$

Einsetzen von $Y = C + I$ gibt

$$\mathbf{Y}'\mathbf{u} = (\mathbf{C} + \mathbf{I})'\mathbf{u} = (\mathbf{C}(u) + \mathbf{I})'\mathbf{u} \qquad \text{und} \qquad \mathrm{E}(\mathbf{Y}'\mathbf{u}) \neq 0$$

Daher verschwindet der zweite Term nicht, $\mathrm{E}[(\mathbf{Y}'\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{u}] \neq 0$, und *der Schätzer ist verzerrt*.

Eine technische Lösung dafür wird in Ö2 diskutiert.

Missspezifikation und extrinsic closure

Das Endogenitätsproblem ist eine ganz generelles der Modellierung und betrifft die extrinsic closure ihres ökonomischen Modells.

Eine Verletzung der extrinsic closure im Ein-Gleichungsmodell ist das Weglassen von relevanten erklärenden Variablen auf der rechten Seite einer Gleichung.

Das Endogenitätsproblem hingegen bezieht sich auf das Ignorieren von relevanten systemischen Zusammenhängen. Die vermeintlich exogenen Variablen auf der rechten Seite werden aber simultan mit der abhängigen Variablen bestimmt.

Modellselektion

Informationskriterien, Variablenauswahl

Angenommen sie haben eine große Anzahl von potenziell erklärenden Variablen zur Auswahl, $x_1, \ldots, x_N$, und wissen nicht welche davon ins Modell aufgenommen werden sollen.

ZB lautet das wahre Modell

$$y = \beta_0 + \beta_3 x_3 + \beta_5 x_5 + \beta_{27} x_{27} + u$$

Eine übliche Praxis ist das Modell zu wählen, bei dem ein Informationskriterium **minimal** wird. Im Folgenden finden sich zwei häufig verwendete.

Informationskriterien: AIC und SBC

- **Akaike's Informationskriterium**, **AIC**

$$AIC(k) = \log(\hat{\mathbf{u}}'\hat{\mathbf{u}}/n) + 2k/n$$

- **Schwarz-Bayes' Informationskriterium**, **SBC**:

$$SBC(k) = \log(\hat{\mathbf{u}}'\hat{\mathbf{u}}/n) + \log(n)k/n$$

$\hat{\mathbf{u}}$ sind jeweils die Residuen aus einem Modell mit k erklärenden Variablen.
$\hat{\mathbf{u}}'\hat{\mathbf{u}}/n \approx \widehat{\sigma}_u^2$, die Residualvarianz (nicht um FGe korrigiert).

EViews berechnet AIC und SBC über die log-likelihood, $\ell\ell$:

$$\log(\hat{\mathbf{u}}'\hat{\mathbf{u}}/n) \approx -2\ell\ell(k)/n$$

Informationskriterien

Informationskriterien messen einen *Trade-off* zwischen

- dem *Fit* (Anpassungsgüte)

$$\log(\hat{\mathbf{u}}'\hat{\mathbf{u}}/n) \approx \log(\widehat{\sigma}_u^2)$$

  und

- der *Komplexität des Modells* (Anzahl der erklärenden Variablen, k), indem ein *Pönale*, das mit k steigt, von

$$2k \qquad \text{bzw} \qquad \log(n)k$$

berücksichtigt wird.

Der Trade-off im AIC

Graphik: k $\times$ Komponenten des AIC

AIC, SBC : $C_t = \beta_0 + \beta_1 Y_t^d + \beta_3 t + u_t$, `dats_01.wf1`

Dependent Variable: PCR
Method: Least Squares
Date: 01/11/12   Time: 15:05
Sample: 1977 2009
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	42100.65	6489.575	6.487427	0.0000
PYR	0.373791	0.078600	4.755624	0.0000
T	1241.069	176.7342	7.022235	0.0000

R-squared	0.996547	Mean dependent var	108165.1
Adjusted R-squared	0.996317	S.D. dependent var	20123.54
S.E. of regression	1221.311	Akaike info criterion	17.13974
Sum squared resid	44747996	Schwarz criterion	17.27579
Log likelihood	-279.8058	Hannan-Quinn criter.	17.18552
F-statistic	4328.867	Durbin-Watson stat	0.929998
Prob(F-statistic)	0.000000		

AIC Bsp, Modellwahl

Angenommen wir schätzen 3 Konsumfunktionen für 1977-2009.

$$M1: \widehat{C}_t = \underset{(6490)}{42100} + \underset{(0.079)}{0.374}\, Y_t^d + \underset{(176.7)}{1241}\, t \qquad R^2 = 0.997 \quad AIC = 17.1397$$

$$M2: \widehat{C}_t = \underset{(1266)}{423} + \underset{(0.082)}{0.331}\, Y_t^d + \underset{(0.089)}{0.640}\, C_{t-1} \qquad R^2 = 0.997 \quad AIC = 17.1037$$

$$M3: \widehat{C}_t = \underset{(9443)}{22733} + \underset{(0.078)}{0.291}\, Y_t^d + \underset{(0.140)}{0.370}\, C_{t-1} + \underset{(275.0)}{654.7}\, t \qquad R^2 = 0.997 \quad AIC = 16.986$$

Wir wählen das Modell mit dem kleinsten AIC-Wert. Das ist Modell 3.

AIC Bsp

Modell 3 könnten wir mittels F-Test auch gegen Modell 2 testen. Modell 3 wäre das unrestricted und Modell 2 das restricted ($\beta_3 = 0$) Modell. Man sagt auch: die beiden Modelle/Hypothesen sind **nested**.

Modell 1 können wir mit dem F-Test allerdings nicht gegen Modell 2 (oder umgekehrt) testen. Sie sind **nicht nested**.

Informationskriterien bieten eine einfache Möglichkeit auch zwischen nicht genesteten Modellen zu wählen.

Referenzen

Hackl 3.1.2, Anhang 3.B, 6.4, 2.3

Wooldridge 3.5, 3.3, 10.5 (AIC und SBC werden nicht behandelt)

Keynes, JM (1939): Professor Tinbergen's Method, *The Economic Journal*, 49(195), 558-77.

Übungen

1 Berechnen sie die trendbereinigte Reihe für das logarithmierte österreische reale BIP aus `okuns_law.wf1`. Erstellen sie eine Tabelle aus der die Auf- und Abschwungsphasen, wie auch die Hochkonjunktur- und Rezessionsphasen ablesbar sind.
  Vergleichen sie die Ergebnisse mit den Phasen des US BIP. Siehe dazu zB Wikipedia unter 'business cycles'.

Übungen

2 Gegeben sind Reihen für den realen Konsum, C, und das disponible Einkommen für Ö, Y^d, `dats_01.wf1`. Berechnen sie die trendbereinigte Reihe

(a) des Konsums, C_{tb}, und

(b) des Einkommens, Y^d_{tb}.

Schätzen sie das Modell

(c) $C = \beta_0 + \beta_1 Y^d + \beta_2 t + u$

(d) $C_{tb} = \gamma_0 + \gamma_1 Y^d_{tb} + v$

Vergleichen sie die Schätzer für β_1 und γ_1.

EViews: Abspeichern der Residuen (= trendbereinigte Reihen)
Modell schätzen. Im `Estimation` Fenster $\rightarrow$ `Proc` Menü $\rightarrow$ `Make Residuals`
Namen eingeben

Übungen

3. Angenommen das wahre Modell ist
   $C_t = \beta_0 + \beta_1 Y_t^d + \beta_2 C_{t-1} + u_t$
   Sie schätzen aber nur
   $C_t = \beta_0 + \beta_1 Y_t^d + u_t$
   Schätzen sie beide Modelle und interpretieren sie die Differenz der beiden Schätzer für β_1 als omitted variable bias.
   Verwenden sie `dats_01.wf1`.

Übungen

4 Wählen sie das Datenfile `hprice1.wf1`. Darin befinden sich uA die Variablen `price` ... Verkaufspreis, `assess` ... Schätzwert, `lotsize` ... Grundstücksfläche, `sqrft` ... m^2, `bdrms` ... Anzahl Schlafzimmer

Zur Diskussion steht das Modell

$$\log(price) = \beta_0 + \beta_1 \log(assess) + \beta_2 \log(lotsize) + \beta_3 \log(sqrft) + \beta_4 bdrms + u$$

Geben sie mit Hilfe des AIC ein geeignete Variablenauswahl an. Probieren sie verschiedene Spezifikationen.

Übungen

5 Wählen sie das Datenfile `dats_01.wf1`. Berechnen sie die Regression

```
pcr c pyr (wln(-1)/pcd(-1)) pcr(-1)
```

Welche Variablenkombination schlägt das AIC vor? Probieren sie verschiedene Spezifikationen.

Anhang: Keynes zum omitted variable bias

Keynes (1939), p.560:
“Am I right in thinking that the method of multiple correlation analysis essentially depends on the economist having furnished, not merely a list of the significant causes, which is correct so far as it goes, but a complete list? For example, suppose three factors are taken into account, it is not enough that these should be in fact verae causae; there must be no other significant factor. If there is a further factor, not taken account of, then the method is not able to discover the relative quantitative importance of the first three. If so, this means that the method is only applicable where the economist is able to provide beforehand a correct and indubitably complete analysis of the significant factors. The method is one neither of discovery nor criticism. It is a means of giving quantitative precision to what, in qualitative terms, we know already as the result of a complete theoretical analysis.“