BI HANDELSHØYSKOLEN

Nydalsveien 37, 0484 Oslo
Postadresse 0442 Oslo

Telefon 06600
Telefaks 21 04 80 00
Studieinfo. tlf. 810 00 500

www.bi.no
info@bi.no
Org. nr. 971 22 8865

TEMA 8: Modellseleksjon

1. Motivasjon
2. Utelating av relevante forklarings-variabler
3. Inkludering av irrelevante forklarings-variabler
4. Feil funksjonell form

BI HANDELSHØYSKOLEN
Nydalsveien 37, 0484 Oslo
Postadresse 0442 Oslo
Telefon 06600
Telefaks 21 04 80 00
Studieinfo. tlf. 810 00 500
www.bi.no
info@bi.no
Org. nr. 971 22 8865

(1) Motivasjon

(2) Utelating av relevante forklaringsvariabler

→ Kan føre til skjeve og inkonsistente estimater

ikke-forventningsrette "ukorrekte"

Eksempel (s. 221 →).

Anta at følgende modell er "sann" i den forstand at de klassiske forutsetningene er oppfylt:

$$Y = B_1 + B_2 X_2 + B_3 X_3 + u$$

Anta videre at vi estimerer:

$$Y = A_1 + A_2 X_2 + v$$ (vi har utelatt X_3)

Hva betyr dette? Hvis X_2 og X_3 er korrelerte og $B_3 \neq 0$, så vil:

→ $E(a_1) \neq B_1$ og $E(a_2) \neq B_2$

estimatet til A_1 — estimatet til A_2

→ a_1 og a_2 vil være inkonsistente

(*)

(*): Estimatene er "ukorrekte"

BI HANDELSHØYSKOLEN
Nydalsveien 37, 0484 Oslo
Postadresse 0442 Oslo
Telefon 06600
Telefaks 21 04 80 00
Studieinfo. tlf. 810 00 500
www.bi.no
info@bi.no
Org. nr. 971 22 8865

Hvis enten X_2 og X_3 er ukorrelerte eller $\underbrace{B_3 = 0}$ så vil:

X_3 har ingen innvirkning

$\rightarrow$ $\underbrace{E(a_1) = B_1}$ og $\underbrace{E(a_2) = B_2}$

forventningsrettet

$\rightarrow$ a_1 og a_2 vil være konsistente

(**)

(**): estimatene vil være "korrekte"

Hva med tilfellet hvor vi utelater mange forklaringsvariabler?

$\rightarrow$ Generelt så viser det seg at det er vanskelig å skissere effekter som gjelder veldig ofte

$\rightarrow$ Som oftest, hvis de klassiske forutsetningene er oppfylt, så vil vi bevare forventningsrettet og konsistens

Konklusjonen: Sjekk om de klassiske forutsetningene er oppfylt:

BI HANDELSHØYSKOLEN
Nydalsveien 37, 0484 Oslo
Postadresse 0442 Oslo
Telefon 06600
Telefaks 21 04 80 00
Studieinfo. tlf. 810 00 500
www.bi.no
info@bi.no
Org. nr. 971 22 8865

Autokorr. / korrjon

- Parameterstabilitet
- Forklaringsvariablene og u er ukorrelerte
- Feilleddene er ukorrelerte

Heteroskedastisitet

- "Presisjonen" er konstant: $E(u^2) = \sigma^2$

(3) Inkludering av irrelevante variabler

Eksempel (ca. s. 223). La følgende modell oppfylle de klassiske forutsetningene:

$$(1) \quad Y = B_1 + B_2 X_2 + u$$

$$B_1 \neq 0 \qquad B_2 \neq 0$$

Anta videre at vi estimerer

$$(2) \quad Y = A_1 + A_2 X_2 + A_3 X_3 + u$$

Det antas X_3 ikke har noen effekt på Y.

Hva skjer hvis vi estimerer (2) i stedet for (1)?

→ Estimatene av A_1, A_2 og A_3 er "korrekte"

forventningsrettet og konsistens

BI HANDELSHØYSKOLEN
Nydalsveien 37, 0484 Oslo
Postadresse 0442 Oslo
Telefon 06600
Telefaks 21 04 80 00
Studieinfo. tlf. 810 00 500
www.bi.no
info@bi.no
Org. nr. 971 22 8865

→ Vanlige hypotesetestingsstrategier er gyldige

→ MEN: Estimatene til A_1, A_2 og A_3 er mer upresise (estimatene er ineffisiente) enn til fallet hvor vi utelater X_3

(4) Feil funksjonell form

I korthet: Det at én eller flere av de klassiske forutsetningene ikke er oppfylt; og dette skjer veldig ofte fordi vi har utelatt variabler

RESET-testen

Betrakt

$$Y = B_1 + B_2 X_2 + \dots + B_k X_k + u$$

Hvis KF3 er oppfylt, dvs. at $E(u \mid X) = 0$ for hver av X'ene, så kan det vises at ingen funksjon av X'ene burde være signifikante hvis vi legger dem til

BI HANDELSHØYSKOLEN
Nydalsveien 37, 0484 Oslo
Postadresse 0442 Oslo
Telefon 06600
Telefaks 21 04 80 00
Studieinfo. tlf. 810 00 500
www.bi.no
info@bi.no
Org. nr. 971 22 8865

Eksempel: La

$$Y = B_1 + B_2 X + u$$

oppfylle de klassiske forutsetningene. Det kan vises at B_3 i

$$Y = B_1 + B_2 X + B_3 X^2 + u$$

kunne være insignifikant

Stegene i RESET-testen

1. La

$$(*) \quad Y = b_1 + b_2 X_2 + \dots + b_k X_k + e$$

være modellen vi ønsker å teste

2. Estimér en ny modell

$$(**) \quad Y = c_1 + c_2 X_2 + \dots + c_k X_k + d_1 \hat{Y}^2 + d_2 \hat{Y}^3 + \dots + d_L \hat{Y}^{(L+1)} + v$$

hvor $\hat{Y}$ er forventningsverdien til (*)

3. Teste om $\hat{Y}^2, \hat{Y}^3, \dots, \hat{Y}^{L+1}$ er signifikante vha. en F-test:

RSS_r: Modell (*)

RSS_{ur}: Modell (**)

L: ant. restriksjoner i H_0

$$F = \frac{(RSS_r - RSS_{ur})/L}{RSS_{ur}/(n-k-L)}$$

BI HANDELSHØYSKOLEN
Nydalsveien 37, 0484 Oslo
Postadresse 0442 Oslo
Telefon 06600
Telefaks 21 04 80 00
Studieinfo. tlf. 810 00 500
www.bi.no
info@bi.no
Org. nr. 971 22 8865

som er $F(L, n-K-L)$ fordelt hvis d'ene er null, hvis den funksjonelle formen i (*) er korrekt

4. Konklusjon: Hvis F høyere enn kritisk verdi, så forkaster vi H_0

Eksempel: Huspriser 2 i USA

Estimér:

$$pris = \beta_1 + \beta_2 BOA + u$$

→ "Automatisk" RESET-test i EViews

1. Trykk på "View" knappen
2. "Stability Tests" → "Ramsey RESET Test"
3. Velg antall ledd, dvs. L

Hvis 1 ledd:

$F = 4{,}7413$ p-verdi $= 0{,}0322$

Tolkning: Det er en viss støtte for feilspesifikasjon siden P-verdien er såpass lav