ECON4135

SW chapter9: Internal and external validity

World Statistics Day 20.10.2010

The International Statistical Institute

Declaration on Professional Ethics in Statistics

History of ISI

History of Statistics Norway

Einar Lie og Hege Roll-Hansen (2001) Faktisk Talt; Statistikkens historie I Norge. Universitetsforlaget

Schweder, T. (2003). Statistikkens historie i Norge—faktisk uten statistikere? Tidskrift for Samfunnsforskning, 44, 309–318.

Look up the Stock and Watson **student resources**

Validity

- A finding is valid if it is close to the **truth**
- Truth, of a **defined population** in a **particular setting**, is unknown.
- The quantified **uncertainty** in the finding should not be misleading for valid findings

- Validity can only be discussed with respect to method - are there **threats to validity** in the way data are gathered or analyzed?
- **Internal validity**: the findings are close to the truth for the population studied in the actual setting
- **External validity**: the findings can be extrapolated and hold for other populations or other settings
    - Without internal validity no external validity
- Findings could be
    - Numerical descriptions
        - Median income, income inequality
        - Degre of social generocity in a country
        - ...
    - Structure of relations
        - Which variables should be included in a regression
        - On what scales
    - **Quantified causal effect of a policy intervention**
        - Estimated effect on test scores of improving the student teacher ratio
    - Predicting the value of an unobserved variable
        - Unemployment rate next year

Threats to internal validity

- Omitted variables
- Overfitting
- Wrong functional form
- Measurement errors in regressors
- Biased sampling
- Simultaneous causality
- Fishing

Measurement errors regressors

The measured (with errors) regressor $\tilde{X} = X + w, \operatorname{cov}(X, w) = 0$ has $\operatorname{var}(\tilde{X}) = \sigma_X^2 + \sigma_w^2, \operatorname{cov}(X, \tilde{X}) = \sigma_X^2$ makes the scatter too flat, and causes the regression estimate to be too close to zero (negative bias when $\beta > 0$)

Example:

$$E\left[X \mid \tilde{X} \right] = \gamma_0 + \frac{\operatorname{cov}(X, \tilde{X})}{\operatorname{var}(\tilde{X})} \tilde{X} \quad \text{(simple linear regression)}$$

$$E[Y \mid X] = E\left[Y \mid X, \tilde{X} \right] = \beta_0 + \beta_1 X$$

$$\Downarrow$$

$$E\left[Y \mid \tilde{X} \right] = E\left[E\left[Y \mid X, \tilde{X} \right] \mid \tilde{X} \right]$$

$$= E\left[\beta_0 + \beta_1 X \mid \tilde{X} \right] = \beta_0 + \beta_1 E\left[X \mid \tilde{X} \right]$$

$$= \beta_0 + \gamma_0 \beta_1 + \beta_1 \frac{\sigma_X^2}{\sigma_X^2 + \sigma_w^2}$$

Correct for the bias when you know how much of the variance in the measured variable is due to mesurement error!

Biased sampling

- Units are often “self-selected”
- non-response is widespread, and increasing, often about 40% of the sampled
    - A.N. Kiær (1895) developed "The representative method" of survey sampling
    - Kiær argued in vain for his representative method at four sessions of the ISI from 1895 to 10903 – survey sampling was thought contrary to “la statistique serieuse” of counting all units

- The actuary Hiorth showed that Kiær's estimate of the cost of a universal health insurance was negatively biased - due to biased sampling.
  - Kiær did not understands Hjorth's mathematical arguments, and thought
  - Survey sampling was not used by Statistics Norway until the 1930s (Gallup)

Simultaneous causality

In an endogenous system $Y \leftrightarrow X$,

$$Y = \beta_0 + \beta_1 X + u$$
$$X = \gamma_0 + \gamma_1 Y + v$$

A large residual u makes Y large, which makes X large ($\gamma_1 > 0$) and thus $\text{cov}(X,U) > 0$. Bias!

Example: A key for distributing income from the state to Norwegian communes was based on linear regressions of communal expenditure on demographic and other "objective" variables characterizing the individual commune (the unit). The regression coefficients were used to construct the key. Why could we have a simultaneous causality bias here? What could be done? (Schweder, T. 1996. Rattsø-utvalgets regresjoner. Sosialøkonomen, nr 4: 15-18.)

- Controlled experiment – manipulate X
- Natural experiment
- Instrument variables (SW chapter 12)

Fishing

Enough typewrighters given to enough monkeys…

External validity

Test scores and class sizes