Applied Regression

Feb 1- Feb 6, 2006

Lecture Outline

- The Econometrician's Problem revisited
  - The disturbance term
  - Samples of Y given X
  - Choose an estimator of beta
- Properties of a "Good" Estimator
  - Low cost
  - Unbiased
  - Efficient/small variance
  - Consistent

Lecture Outline, continued

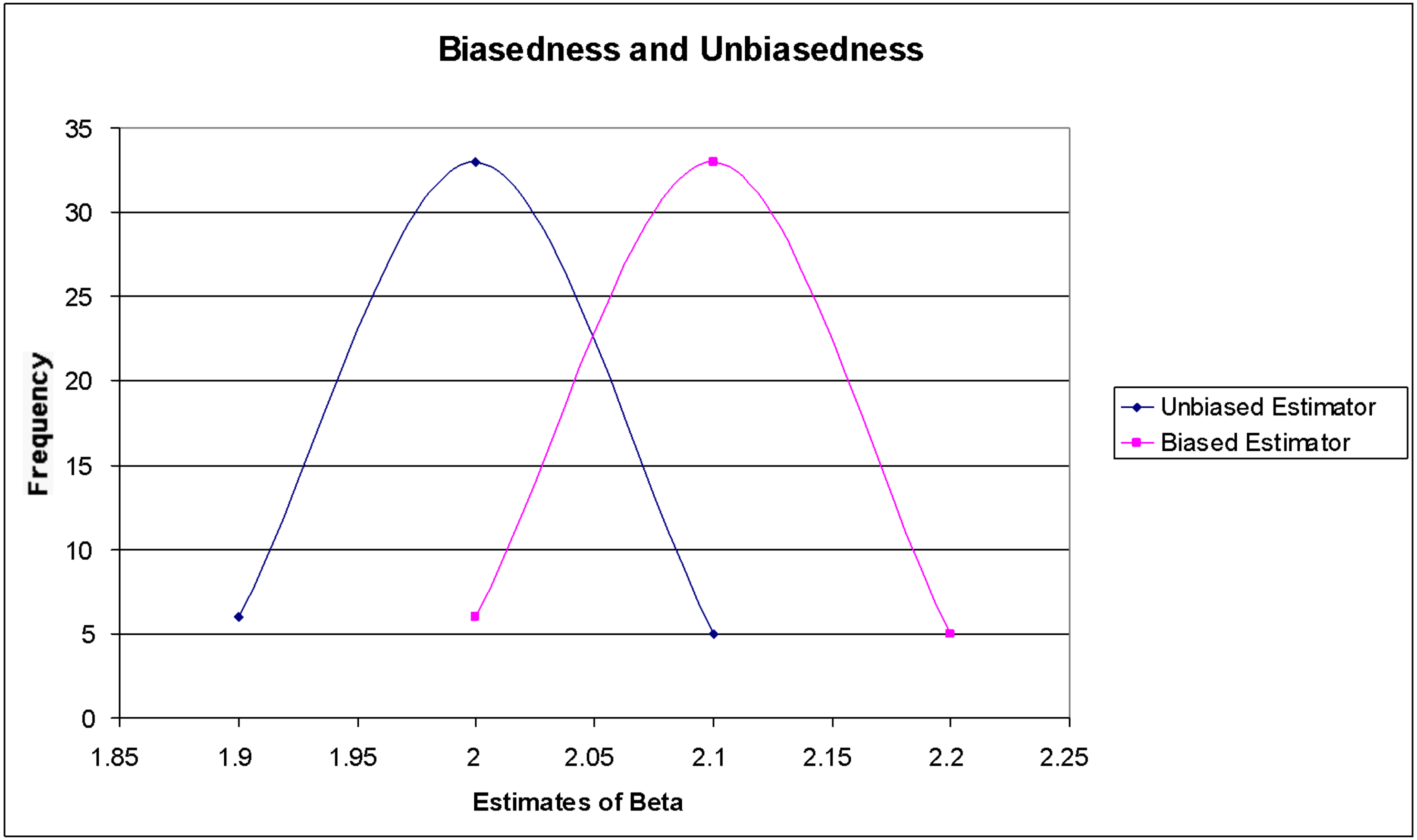
- The Ordinary Least Squares Estimator
  - Theoretical Equation
  - OLS chooses a,b,g to minimize sum of the squared residuals.
  - Derivation of OLS estimators
- Classical Assumptions
- Gauss Markov Theorem

The Econometrician’s Problem Revisited

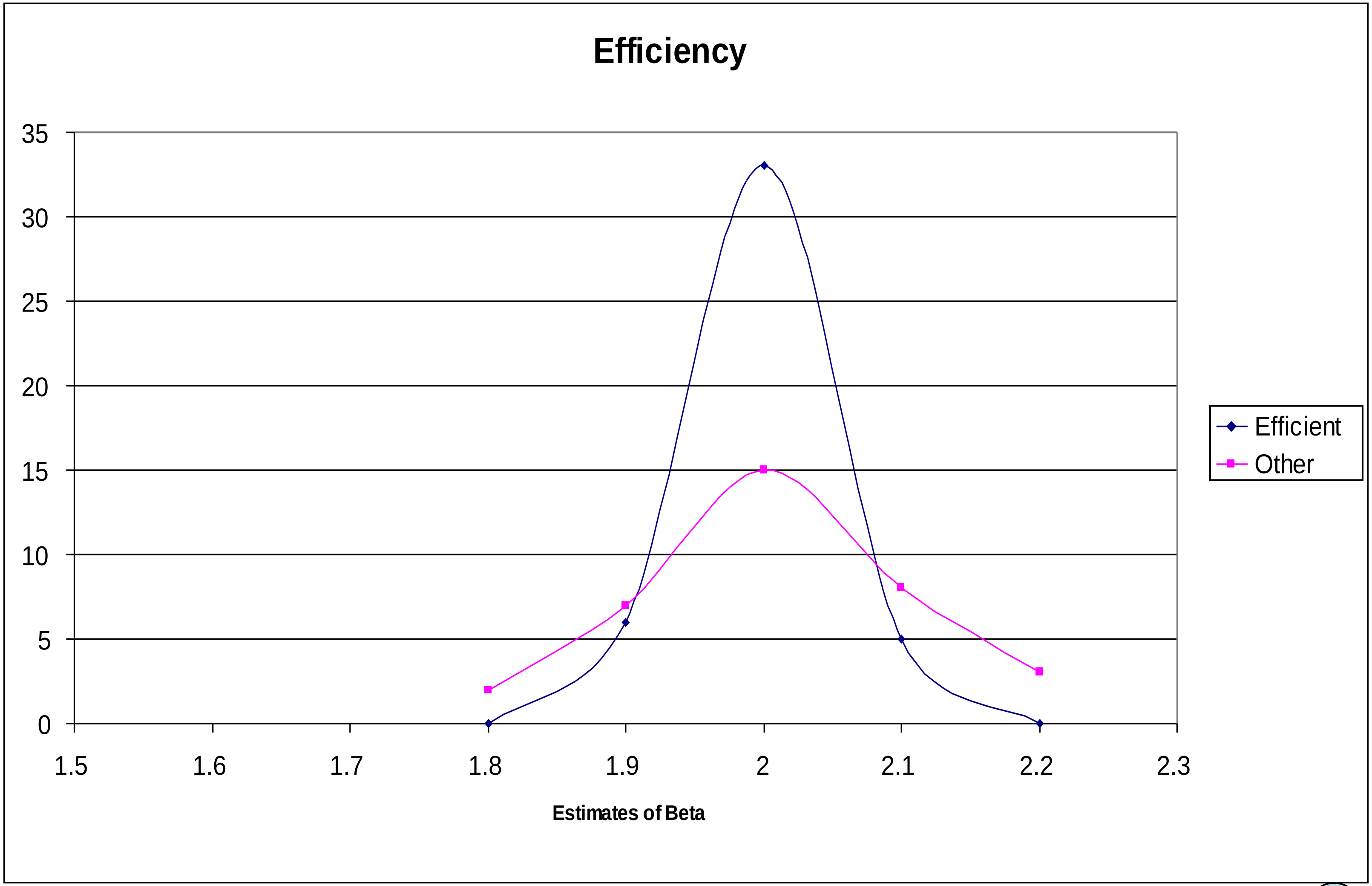
- Theoretical relationship: $Y=\alpha+\beta X$
- Econometric equation: $Y=\alpha+\beta X+\varepsilon$
- Synthetic.xls Used random number generator to obtain ε. Because of the disturbance term, Y could vary across the same observations of X
- Mock Data; Estimate β
- We want to study the properties of the sampling distributions of estimator for β.

Properties of Good Estimator

- Low cost
- Unbiased
- Efficient/small variance
- Consistent



Biasedness and Unbiasedness
Frequency
35
30
25
20
15
10
5
0
1.85
1.9
1.95
2
2.05
2.1
2.15
2.2
2.25
Estimates of Beta
Unbiased Estimator
Biased Estimator


Efficiency
35
30
25
20
15
10
5
0
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
Estimates of Beta
Efficient
Other

How Does OLS Work?

OLS chooses α,β_x,β_y to minimize the sum of the squared residuals

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$$

$$= \sum \left(Y_i - \hat{\alpha} - \hat{\beta}_x X_i - \hat{\beta}_z Z_i\right)^2$$

The OLS Estimated Coefficients

$$\hat{\beta}_x = \frac{\left(\sum_{i=1}^{n} y_i x_i\right)\left(\sum_{i=1}^{n} z_i^2\right) - \left(\sum_{i=1}^{n} y_i z_i\right)\left(\sum_{i=1}^{n} x_i z_i\right)}{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} z_i^2\right) - \left(\sum_{i=1}^{n} x_i z_i\right)^2}$$

$$\hat{\beta}_z = \frac{\left(\sum_{i=1}^{n} y_i z_i\right)\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} y_i x_i\right)\left(\sum_{i=1}^{n} z_i x_i\right)}{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} z_i^2\right) - \left(\sum_{i=1}^{n} x_i z_i\right)^2}$$

$$\alpha = \overline{Y} - \hat{\beta}_x X - \hat{\beta}_z Z$$

$$\text{where } y_i = Y_i - \overline{Y}; x_i = X_i - \overline{X}; z_i = Z_i - \overline{Z}$$

Classical Assumptions

- The regression model is linear in the coefficients and the error term.
- The error term has zero population mean
- All explanatory variables are uncorrelated with the error term
- Observations of the error term are uncorrelated with each other (no serial correlation)
- The error term has constant variance (no heteroskedasticity)
- No explanatory variable is a perfect linear function of the other explanatory variables (no perfect multicollinearity)
- The error term is normally distributed. (optional).

OLS Estimator of β

$$\beta_x = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}$$

Unbiasedness

Estimator is unbiased if $\mathrm{E}(\hat{\beta}) = \beta$.

$(x_{\mathrm{t}} = X_t - \overline{X}; y_t = Y_t - \overline{Y})$

$$E\left(\frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2}\right) = \frac{1}{\sum_{t=1}^{T} x_t^2} E\left(\sum_{t=1}^{T} x_t y_t\right)$$

Substituting $\mathrm{Y}_{\mathrm{t}} = \alpha + \beta \mathrm{X}_{\mathrm{t}} + \varepsilon_t$ and $\overline{\mathrm{Y}} = \frac{1}{\mathrm{T}}\sum_{t=1}^{T}[\alpha + \beta \mathrm{X}_{\mathrm{t}} + \varepsilon_t]$.

$$= \frac{1}{\sum_{t=1}^{T} x_t^2} E\left(\sum_{t=1}^{T} x_t(\alpha + \beta X_t + \varepsilon_t) - \sum_{t=1}^{T} x_t(\alpha + \tfrac{1}{T}\beta \sum_{t=1}^{T} X_t + \tfrac{1}{T}\sum_{t=1}^{T} \varepsilon_t)\right)$$

$$= \frac{1}{\sum_{t=1}^{T} x_t^2} \sum_{t=1}^{T}\left[x_t(\alpha + \beta X_t - \alpha - \beta \overline{X})\right] - \frac{1}{\sum_{t=1}^{T} x_t^2} \sum_{t=1}^{T}[x_t E(\varepsilon_t - \bar{\varepsilon})]$$

Unbiasedness, cont.

$$= \frac{1}{\sum_{t=1}^{T} x_t^2}\left[\beta \sum_{t=1}^{T} x_t (X_t - \bar{X})\right] + \frac{1}{\sum_{t=1}^{T} x_t^2}\left(\sum_{t=1}^{T} x_t E(\varepsilon_t)\right)$$

Because $\mathrm{E}(\varepsilon_t) = 0$ & $\mathrm{E}(\mathrm{X}_t \varepsilon_t) = 0$

$$= \beta \frac{\sum_{t=1}^{T} x_t x_t}{\sum_{t=1}^{T} x_t^2} = \beta$$

Alternative Estimator

$$\tilde{\beta} = \frac{Y_1 - Y_T}{X_1 - X_T}$$

$$E(\tilde{\beta}) = \frac{1}{X_1 - X_T} E\left(\alpha + \beta X_1 + \varepsilon_1 - \alpha - \beta X_T - \varepsilon_T\right)$$

$$= \frac{1}{X_1 - X_T} E\left(\beta(X_1 - X_T) + (\varepsilon_1 - \varepsilon_T)\right)$$

$$= \beta + E(\varepsilon_1 - \varepsilon_T) = \beta, \text{ if } \mathrm{E}(\varepsilon) = 0$$

Consistency

- What happens as the size of the sample approaches the population?
  - If X and ε are not correlated (independent) and Var (X)>0, OLS estimator gets closer to its true value.
  - Slope estimator doesn't depend on T. So it can not be consistent

OLS is Consistent

$$\hat{\beta} = \beta + \frac{\sum_{t=1}^{T} x_t(\varepsilon_t - \bar{\varepsilon})}{\sum_{t=1}^{T} x_t^2}$$

Large of law numbers says as $\mathrm{T} \to \infty, \ \sum_{t=1}^{T} x_t(\varepsilon_t - \bar{\varepsilon})$

converges to its expectation, $\mathrm{Cov}(\mathrm{X}, \varepsilon) = 0$

Gauss-Markov Theorem

OLS is BLUE

Gauss Markov Theorem

Under the Classical Assumptions, the OLS estimator of β_k is the minimum variance estimator from among the set of all linear unbiased estimators for β_k, for $k=1,\ldots,K$

Flow Chart of Proof of Gauss-Markov Theorem

Define Class of Linear Estimators:

$$\tilde{\beta} = \sum_{i=1}^{N} \gamma_i y_i$$

Recognize that OLS estimator is member of this class:

$$\hat{\beta} = \sum_{i=1}^{N} \omega_i y_i \quad \text{where } \omega_i = \frac{x_i}{\sum_{i=1}^{N} x_i}$$

Observe That

$$\tilde{\beta} = \sum_{i=1}^{N} (d_i + \omega_i) y_i$$

$$\text{where } d_i = \gamma - \omega_i$$

Identify Restrictions to Insure Unbiasedness:

$$\sum_{i=1}^{N} d_i = 0 \text{ and } \sum_{i=1}^{N} d_i X_i = 0$$

Choose d_i to minimize variance:

$$d_i = 0$$

So the best linearly unbiased estimator is

$$\tilde{\beta}^* = \sum_{i=1}^{N} (0 + \omega_i) y_i = \sum_{i=1}^{N} \omega_i y_i = \hat{\beta}$$

OLS is linear

$$\hat{\beta} = \sum_{t=1}^{T} \frac{x_t}{\sum_{t=1}^{T} x_t^2} y_t = \sum_{t=1}^{T} w_t y_t$$

General Class of Linear Estimators

$$\tilde{\beta} = \sum_{t=1}^{T} \gamma_t y_t$$

Let $\mathrm{d}_{\mathrm{t}} = \gamma_t - w_t$ so

$$\tilde{\beta} = \sum_{t=1}^{T} (w_t + d_t) y_t = \hat{\beta} + \sum_{t=1}^{T} d_t y_t$$

Unbiasedness requires

$$E(\tilde{\beta}) = E(\hat{\beta}) + \sum_{t=1}^{T} d_t E(Y_t)$$

$$= \beta + \alpha \sum_{t=1}^{T} d_t + \beta \sum_{t=1}^{T} d_t X_t$$

$$\mathrm{E}(\tilde{\beta}) = \beta \text{ if and only if } \sum_{t=1}^{T} d_t = 0 \text{ and } \sum_{\mathrm{t}=1}^{\mathrm{T}} d_t X_t = 0$$

Minimize Variance

$$Var(\tilde{\beta}) = Var(\sum_{t=1}^{T}(w_t + d_t)Y_t)$$

The variance of the sum of independent random variables is the sum of the variances. Because of assumption of no serial correlation, we have independence. Because of assumption of no heteroskedasticity, variance of error term and of Y is constant.

$$Var(\sum_{t=1}^{T}(w_t + d_t)Y_t) = \sum_{t=1}^{T} E[(w_t + d_t)Y_t - (w_t + d_t)\overline{Y})^2]$$

$$= \sum_{t=1}^{T} E[(w_t + d_t)^2 (Y_t - \overline{Y})^2]$$

$$= \sum_{t=1}^{T} (w_t + d_t)^2 E[(Y_t - \overline{Y})^2]$$

$$= \mathrm{var}(\varepsilon)\sum_{t=1}^{T}(w_t + d_t)^2$$

$$= \mathrm{var}(\varepsilon)\sum_{t=1}^{T}(w_t{}^2 + 2d_t w_t + d_t^2)$$

Second term is zero

The first term is not a choice variable. To minimize variance, must have $d_t = 0$

Summary

- OLS is BLUE
- Consistency and Unbiasedness require $E(\varepsilon)=0$ and $E(X\varepsilon)=0$
- Efficiency requires no serial correlation and homoscedastic errors