

Best Linear Unbiased Estimator: Lecture IX

Charles B. Moss ¹

¹University of Florida

January 30, 2012

- 1 Readings
- 2 Best Linear Unbiased Estimators
 - BLUE: A Statistical Solution
- 3 Matrix Form of Gauss-Markov
 - Variance of the Estimator
 - Gauss-Markov Theorem

Readings

- * Amemiya, T. 1994. *Introduction to Statistics and Econometrics*. Harvard University Press. (Chapter 12).
- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition. (Chapter 4.1-4.3 pp. 51 - 63)
- * Hansen, B.E. 2011. *Econometrics* Draft Textbook Website: <http://www.ssc.wisc.edu/bhansen>. (Chapter 5).
- Hayashi, F. 2000. *Econometrics* Oxford University Press (Section 1.5 pp. 47-54)

Readings - Continued

- * Moss, C.B. 2012. Mathematical Statistics for Applied Econometrics Draft Textbook Website:
<http://ricardo.ifas.ufl.edu/aeb6571.econometrics/MathStat2012.pdf>. (Chapter 10, pp. 155 - 164).
- * Popper, K. 2010. The Logic of Scientific Discovery Routledge Classics [Chapter 10: Corroboration, or How a Theory Stands Up to Tests, 248-282].

Best Linear Unbiased Estimators: A Statistical Solution

- The linear relationship between the x s and y s

$$E[y_i] = \alpha + \beta x_i \quad (1)$$

and we assume that

$$V(y_i) = \sigma^2 \quad (2)$$

- The implications of this variance assumption are significant.
 - Note that we assume that each observation has the same variance irregardless of the value of the independent variable.
 - In traditional regression terms, this implies that the errors are homoscedastic.

BLUE: A Statistical Solution - Continued

- One way to state these assumptions is

$$y_i = \alpha + \beta x_i + \epsilon_i \quad (3)$$

$$E[\epsilon_i] = 0, \quad V(\epsilon_i) = \sigma^2$$

- This specification is consistent with our assumptions, since the model is homoscedastic and linear in the parameters.
- Based on this formulation, we can define the linear estimators of α and β as

$$\sum_{i=1}^n d_i y_i \quad (4)$$

BLUE: A Statistical Solution - Continued

- An unbiased estimator of β can further be defined as those linear estimators whose expected value is the true value of the parameter

$$E \left[\sum_{i=1}^n d_i y_i \right] = \beta \quad (5)$$

BLUE: A Statistical Solution - Continued

- This implies that

$$\begin{aligned}
 \beta &= E \left[\sum_{i=1}^n d_i y_i \right] \\
 &= \sum_{i=1}^n d_i E[y_i] \\
 &= \sum_{i=1}^n d_i (\alpha + \beta x_i) \tag{6} \\
 &= \alpha \left(\sum_{i=1}^n d_i \right) + \beta \left(\sum_{i=1}^n d_i x_i \right) \Rightarrow \begin{cases} \sum_{i=1}^n d_i = 0 \\ \sum_{i=1}^n d_i x_i = 1 \end{cases}
 \end{aligned}$$

BLUE: A Statistical Solution - Continued

- The linear estimator that satisfies these unbiasedness conditions and yields the smallest variance of the estimate is referred to as the best linear unbiased estimator (or BLUE). In this example, we need to show that

$$d_i = \frac{(x_i - \bar{x})}{S_{xx}} \Rightarrow \hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{S_{xx}} \quad (7)$$

- Minimizes the variance for all such linear models. Given that the y_i s are uncorrelated, the variance of linear model can be written as

$$V \left(\sum_{i=1}^n d_i y_i \right) = \sum_{i=1}^n d_i^2 V(y_i) = \sigma^2 \sum_{i=1}^n d_i^2 \quad (8)$$

BLUE: A Statistical Solution - Continued

- The problem of minimizing the variance then becomes choosing the d_i s to minimize this sum subject to the unbiasedness constraints

$$\min_{d_i} \sigma^2 \sum_{i=1}^n d_i^2$$

$$\text{s.t. } \sum_{i=1}^n d_i x_i = 1 \quad (9)$$

$$\sum_{i=1}^n d_i = 0$$

BLUE: A Statistical Solution - Continued

- Transforming Equation 9 into a Lagrangian form

$$L = \sigma^2 \sum_{i=1}^n d_i^2 + \lambda \left(1 - \sum_{i=1}^n d_i x_i \right) - \mu \left(\sum_{i=1}^n d_i \right)$$

$$\frac{\partial L}{\partial d_i} = 2\sigma^2 d_i - \lambda x_i - \mu = 0$$

$$\Rightarrow d_i = \frac{\lambda}{2\sigma^2} x_i + \frac{\mu}{2\sigma^2} \tag{10}$$

$$\frac{\partial L}{\partial \lambda} = 1 - \sum_{i=1}^n d_i x_i = 0$$

$$\frac{\partial L}{\partial \mu} = - \sum_{i=1}^n d_i = 0$$

BLUE: A Statistical Solution - Continued

- Using the results from the first n first-order conditions and the second constraint first, we have

$$\begin{aligned}
 -\sum_{i=1}^n \left(\frac{\lambda}{2\sigma^2} x_i + \frac{\mu}{2\sigma^2} \right) &= 0 \\
 \Rightarrow \frac{\lambda}{2\sigma^2} \sum_{i=1}^n x_i &= \frac{n\mu}{2\sigma^2} \\
 \lambda \sum_{i=1}^n x_i &= n\mu \\
 \Rightarrow \mu &= \frac{\lambda \sum_{i=1}^n x_i}{n} = \lambda \bar{x}
 \end{aligned} \tag{11}$$

- Substituting this result into the first n first-order conditions yields

$$d_i = \frac{\lambda}{2\sigma^2} x_i - \frac{\lambda}{2\sigma^2} \bar{x} = \frac{\lambda}{2\sigma^2} (x_i - \bar{x}) \tag{12}$$

BLUE: A Statistical Solution - Continued

- Substituting these conditions into the first constraint, we get

$$\begin{aligned}
 1 - \sum_{i=1}^n \frac{\lambda}{2\sigma^2} (x_i - \bar{x}) x_i &= 0 \\
 \Rightarrow \lambda &= \frac{2\sigma^2}{\sum_{i=1}^n (x_i - \bar{x}) x_i} \\
 \Rightarrow d_i &= \frac{(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x}) x_i} = \frac{(x_i - \bar{x})}{S_{xx}}
 \end{aligned} \tag{13}$$

BLUE: A Statistical Solution - Continued

- This proves that simple least squares is BLUE on a fairly global scale.
 - Note that we did not assume normality in this proof.
 - The only assumptions were that the expected error term is equal to zero and that the variances were independently and identically distributed.

Variance of the Estimator

- Variance of the estimated parameters
 - The variance of the parameter matrix can be written as

$$V(\hat{\beta}) = E \left[(\hat{\beta} - \beta) (\hat{\beta} - \beta)' \right] \quad (14)$$

- Working backward

$$\begin{aligned} y &= X\beta + \epsilon \Rightarrow \hat{\beta} = (X'X)^{-1} X'y \\ &= (X'X)^{-1} X' (X\beta + \epsilon) \\ &= (X'X)^{-1} X'X\beta + (X'X)^{-1} X'\epsilon \\ &= \beta + (X'X)^{-1} X'\epsilon \end{aligned} \quad (15)$$

Variance of the Estimator - Continued

- Continued
 - Continued
 - Substituting this back into the variance relationship in Equation 14 yields

$$V(\hat{\beta}) = E \left[(X'X)^{-1} X' \epsilon \epsilon' X (X'X)^{-1} \right] \quad (16)$$

- Note that $\epsilon \epsilon' = \sigma^2 I$ therefore

$$\begin{aligned} V(\hat{\beta}) &= E \left[(X'X)^{-1} X' \epsilon \epsilon' X (X'X)^{-1} \right] \\ &= (X'X)^{-1} X' \sigma^2 I X (X'X)^{-1} \\ &= \sigma^2 (X'X)^{-1} X' X (X'X)^{-1} \\ &= \sigma^2 (X'X)^{-1} \end{aligned} \quad (17)$$

- Note that this formulation is based on the true variance. If you substitute the sample estimate of the variance, you have the traditional students t test.

Gauss-Markov Theorem

- **Theorem 12.2.1** (Gauss-Markov) Let $\beta^* = C'y$ where C is a $T \times K$ constant matrix such that $C'X = I$. Then, $\hat{\beta}$ is better than β^* if $\beta^* \neq \hat{\beta}$.
- Proof: Starting with

$$\beta^* = \beta + C'u \quad (18)$$

- This choice of C guarantees that the estimator β^* is an unbiased estimator of β . The variance of β^* can then be written as

$$\begin{aligned} V(\beta^*) &= E[C'uu'C] \\ &= C'E[uu']C \\ &= \sigma^2 C'C \end{aligned} \quad (19)$$

Gauss-Markov Theorem - Continued

- Continued

- To complete the proof, we want to add a special form of zero. Specifically, we want to add $\sigma^2 (X'X)^{-1} - \sigma^2 (X'X)^{-1} = 0$

$$V(\beta^*) = \sigma^2 (X'X)^{-1} + \sigma^2 C' C - \sigma^2 (X'X)^{-1} \quad (20)$$

- Focusing on the last terms, we note that by the orthogonality conditions for the C matrix

$$\begin{aligned} Z'Z &= \left(C - X(X'X)^{-1} \right)' \left(C - X(X'X)^{-1} \right) \\ &= C'C - C'X(X'X)^{-1} - (X'X)^{-1}X'C + (X'X)^{-1}X'X(X'X)^{-1} \end{aligned} \quad (21)$$

- Continued
 - Substituting backwards

$$\begin{aligned}
 \sigma^2 C' C - \sigma^2 (X' X)^{-1} &= \sigma^2 \left[C' C - (X' X)^{-1} \right] \\
 &= \sigma^2 \left[C' C - C' X (X' X)^{-1} - (X' X)^{-1} X' C + (X' X)^{-1} X' X (X' X)^{-1} \right] \\
 &= \sigma^2 \left[\left(C - X (X' X)^{-1} X' \right)' \left(C - X (X' X)^{-1} X' \right) \right] \tag{22}
 \end{aligned}$$

Thus,

$$V(\beta^*) = \sigma^2 (X' X)^{-1} + \sigma^2 \left[\left(C' - (X' X)^{-1} X' \right) \left(C - X (X' X)^{-1} X' \right) \right] \tag{23}$$

- The minimum variance estimator is then $C = X (X' X)^{-1}$ which is the ordinary least squares estimator.