Alastair Hall | **ECON61001: Econometric Methods** | Semester 1

Study Group Questions # 4

In these Study Group questions we return to the log-level specification in Study Group Questions #2. Our focus here is on prediction of the level of the dependent variable based on the results from the log-level model.

Consider the multiple linear regression model,

$$log(w_t) \; = \; x_t'\beta_0 \, + \, u_t, \tag{1}$$

where the dependent variable is $y_t = log(w_t)$, x_t is a $k \times 1$ vector of explanatory variable, and u_t is the error. Let $\hat{\beta}_T$ and $\hat{\sigma}_T^2$ be the OLS estimators of β_0 and $\sigma_0^2 = Var[u_t]$ respectively based on the model in (1) and observations $t \; = \; 1, 2 \ldots, T$. Suppose now that x_{T+1} is known and it is desired to predict w_{T+1}. Assume that the model in (1) holds for $t = 1, 2, \ldots, T, T+1$ and that $\{x_t\}_{t=1}^{T+1}$ are fixed in repeated samples and $(u_1, u_2, \ldots, u_{T+1})' \sim N\left(0, \, \sigma_0^2 I_{T+1}\right)$. Let X denote the $T \times k$ matrix with t^{th} row x_t'.

To facilitate the presentation of the questions below, we need to introduce some additional terminology: the random variable a_{T+1} is said to be an *unbiased predictor* of w_{T+1} if $E[a_{T+1}] = E[w_{T+1}]$, otherwise a_{T+1} is said to be a biased predictor with prediction bias equal to $E[a_{T+1}] - E[w_{T+1}]$.

In addition, you may quote the following two results as needed in your answer:

- *Result 1:* If v is a $m \times 1$ random vector with the multivariate normal distribution $N(\mu, \Sigma)$ then for any constant $c \in \mathbb{R}^m$, we have: $E[e^{c'v}] = e^{c'\mu + 0.5c'\Sigma c}$.
- *Result 2:* If $h \sim \chi_n^2$ then for any constant $c < 0.5$ we have $E[e^{ch}] = \{1/(1-2c)\}^{n/2}$.

Please enter the details of the group:

Group name	Matrix
Student ID 1	11335127
Student ID 2	10710007
Student ID 3	10704589
Student ID 4	11465531
Student ID 5	

1. Consider the following predictor of w_{T+1}: $\tilde{w}_{T+1} = exp\{x'_{T+1}\hat{\beta}_T\}$. Evaluate $E[\tilde{w}_{T+1}]$; your answer should be expressed as a function of all or a subset of $\{\beta_0, \sigma_0^2, x_{T+1}, X, T, k, u\}$. Is $\tilde{w}_{T+1}$ an unbiased predictor of w_{T+1}? If not then what is the prediction bias?

From (1) $\Rightarrow$ $w_t = exp(x_t' \beta_0 + u_t)$

Since (1) holds for $T+1$ $\Rightarrow$ $w_{T+1} = exp(x_{T+1}' \beta_0 + u_{T+1})$

Starting by evaluating $E[w_{T+1}]$:

$$E[w_{T+1}] = E[exp(x_{T+1}' \beta_0 + u_{T+1})]$$

$$= exp(x_{T+1}'\beta_0) E[e^{u_{T+1}}] \quad (\text{as } x'_{T+1}\beta_0 \text{ is fixed})$$

We can rewrite $E[e^{u_{T+1}}]$ as:

$$E[e^{u_{T+1}}] = \int_{-\infty}^{+\infty} e^{u_{T+1}} \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{u_{T+1}^2}{2\sigma_0^2}} du$$

$$(\text{as } u_{T+1} \sim N(0, \sigma_0^2))$$

$$= \frac{1}{\sigma_0\sqrt{2\pi}} \int_{-\infty}^{+\infty} exp\left(u_{T+1} - \frac{u_{T+1}^2}{2\sigma_0^2}\right) du \quad (a)$$

Since we can rewrite:

$$u_{T+1} - \frac{u^2_{T+1}}{2\sigma_0^2} = \frac{1}{2}\sigma_0^2 - \frac{(u_{T+1} - \sigma^2)^2}{2\sigma_0^2}$$

(a) $\Rightarrow$ $$E[e^{u_{T+1}}] = \frac{1}{\sigma_0\sqrt{2\pi}} \int_{-\infty}^{+\infty} exp\left(\frac{1}{2}\sigma_0^2\right) exp\left(-\frac{(u_{T+1}-\sigma^2)^2}{2\sigma_0^2}\right) du$$

$$= exp\left(\frac{1}{2}\sigma_0^2\right) \frac{1}{\sigma_0\sqrt{2\pi}} \int_{-\infty}^{+\infty} exp\left(-\frac{(u_{T+1}-\sigma^2)^2}{2\sigma_0^2}\right) du$$

$$= exp\left(\frac{1}{2}\sigma_0^2\right) \int_{-\infty}^{+\infty} \phi(u_{T+1}; \sigma^2; \sigma^2)\, du$$

$$= exp\left(\frac{1}{2}\sigma_0^2\right)$$

$\Rightarrow$ $$E[w_{T+1}] = exp\left(x'_{T+1}\beta_0 + \frac{1}{2}\sigma_0^2\right)$$

Next, we evaluate $E[\tilde{w}_{T+1}] = E[\exp(x'_{T+1}\hat{\beta}_T)]$

Look at $E[\hat{\beta}_T]$ and $var[\hat{\beta}_T]$:

$$E[\hat{\beta}_T] = E[(X'X)^{-1}X'y]$$
$$= E[\beta_0 + (X'X)^{-1}X'u]$$
$$= \beta_0 + (X'X)^{-1}X'E[u]$$
$$= \beta_0 \quad (\text{as } X \text{ is fixed and } E[u] = 0).$$

$$Var[\hat{\beta}_T] = E[\hat{\beta}_T - E[\hat{\beta}_T]][\hat{\beta}_T - E[\hat{\beta}_T]]'$$
$$= E[(X'X)^{-1}Xu + \beta_0 - \beta_0][(X'X)^{-1}Xu + \beta_0 - \beta_0]'$$
$$= E[(X'X)^{-1}X'uu'X(X'X)^{-1}]$$
$$= (X'X)^{-1}X'E[uu']X(X'X)^{-1}$$
$$= \sigma_0^2(X'X)^{-1}$$

$\Rightarrow \hat{\beta}_T \sim N(\beta_0, \sigma_0^2(X'X)^{-1})$ (Theorem 2.3, Lecture Notes, p26)

Therefore, using Result 1, we have:

$$E[\tilde{w}_{T+1}] = E[\exp(x'_{T+1}\hat{\beta}_T)] = \exp\left(x'_{T+1}\beta_0 + \frac{\sigma_0^2 x'_{T+1}(X'X)^{-1}x_{T+1}}{2}\right)$$

$\Rightarrow E[\tilde{w}_{T+1}] \neq E[w_{T+1}]$, $\tilde{w}_{T+1}$ is not an unbiased predictor of w_{T+1} and the prediction bias is:

$$E[\tilde{w}_{T+1}] - E[w_{T+1}] =$$
$$\exp(x'_{T+1}\beta_0)\left[\exp(0{,}5\,\sigma_0^2\, x'_{T+1}(X'X)^{-1}x_{T+1}) - \exp(0{,}5\,\sigma_0^2)\right]$$

2. Now consider the following predictor of w_{T+1}: $\hat{w}_{T+1} = exp\{x'_{T+1}\hat{\beta}_T + 0.5\hat{\sigma}_T^2\}$. Assuming $0.5\sigma_0^2/(T-k) < 0.5$, evaluate $E[\hat{w}_{T+1}]$; your answer should be expressed as a function of all or a subset of $\{\beta_0,\ \sigma_0^2,\ x_{T+1},\ X, T, k, u\}$. Is $\hat{w}_{T+1}$ an unbiased predictor of w_{T+1}? If not then what is the bias? *Hint: under the conditions above $\hat{\beta}_T$ and $\hat{\sigma}_T^2$ are independent.* In answering this question, you may quote, as needed, from Lecture Notes (with appropriate reference): (i) the sampling distribution of $\hat{\beta}_T$; (ii) the formula for $\hat{\sigma}_T^2$ as a function of u and X; (iii) distribution of quadratic forms in standard normal random vectors.

Evaluate $E[\hat{w}_{T+1}] = E[\exp(x'_{T+1}\hat{\beta}_T + 0{,}5\,\hat{\sigma}_T^2)]$

$= E[\exp(x'_{T+1}\hat{\beta}_T)\exp(0{,}5\,\hat{\sigma}_T^2)]$

$= E[\exp(x'_{T+1}\hat{\beta}_T)]E[\exp(0{,}5\,\hat{\sigma}_T^2)]$

(Because $\hat{\beta}_T$ and $\hat{\sigma}_T$ are independent)

We can rewrite the 2^{nd} expectation as:

$$E[\exp(0{,}5\,\hat{\sigma}_T^2)] = E\left[\exp\left\{\frac{(T-k)\hat{\sigma}_T^2}{\sigma_0^2}\cdot\frac{0{,}5\,\sigma_0^2}{(T-k)}\right\}\right]$$

By construction, we have $\frac{0{,}5\sigma_0^2}{T-k} < 0{,}5$

Also, from Theorem 2.9 (Lecture Notes, p67), we have:

$$\frac{(T-k)\hat{\sigma}_T^2}{\sigma_0^2} \sim \chi^2_{T-k}$$

Using Result 2, we have:

$$E\left[\exp\left\{\frac{(T-k)\hat{\sigma}_T^2}{\sigma_0^2}\cdot\frac{0{,}5\,\sigma_0^2}{(T-k)}\right\}\right] = \left(\frac{1}{1-\frac{\sigma_0^2}{T-k}}\right)^{\frac{T-k}{2}} = \left(\frac{T-k}{T-k-\sigma_0^2}\right)^{\frac{T-k}{2}}$$

$$\Rightarrow E[\hat{w}_{T+1}] = \exp\left(x'_{T+1}\beta_0 + \frac{\sigma_0^2 x'_{T+1}(X'X)^{-1}x_{T+1}}{2}\right)\left(\frac{T-k}{T-k-\sigma_0^2}\right)^{\frac{T-k}{2}}$$

Thus, $\hat{w}_{T+1}$ is not an unbiased estimator of w_{T+1}

The prediction bias is:

$E[\hat{w}_{T+1}] - E[w_{T+1}] =$

$$\exp(x'_{T+1}\beta_0)\left[\exp\left(\frac{\sigma_0^2 x'_{T+1}(X'X)^{-1}x_{T+1}}{2}\right)\left(\frac{T-k}{T-k-\sigma_0^2}\right)^{\frac{T-k}{2}} - \exp\left(\frac{\sigma_0^2}{0{,}5}\right)\right]$$

But from the programming assignment, we see that when $\sigma_0^2 = 1$ and when T is large, the second component of the prediction bias can get asymptotically close to 0, which means $\hat{w}_{T+1}$ is closer to w_{T+1}.

So therefore, even $\hat{w}_{T+1}$ is not an unbiased estimator of w_{T+1} for every σ_0^2, it's a better estimator compared to $\tilde{w}_{T+1}$.

3. Consider the following prediction interval for w_{T+1},

$$\left(exp\{y^{LB}_{T+1}\}, exp\{y^{UB}_{T+1}\} \right),$$

where

$$\begin{aligned} y^{LB}_{T+1} &= x'_{T+1}\hat{\beta}_T - \tau_{T-k,1-\alpha/2}\hat{\sigma}_T\sqrt{1 + x'_{T+1}(X'X)^{-1}x_{T+1}}, \\ y^{UB}_{T+1} &= x'_{T+1}\hat{\beta}_T + \tau_{T-k,1-\alpha/2}\hat{\sigma}_T\sqrt{1 + x'_{T+1}(X'X)^{-1}x_{T+1}}, \end{aligned}$$

and $\tau_{T-k,1-\alpha/2}$ is the $100(1-\alpha/2)^{th}$ percentile of the Student's t distribution with $T-k$ degrees of freedom. Let probability that this prediction interval contains w_{T+1} equal γ that is,

$$P\left(w_{T+1} \in \left(exp\{y^{LB}_{T+1}\}, exp\{y^{UB}_{T+1}\} \right) \right) = \gamma.$$

Assess which of the following statements is true and provide a justification for your conclusion:

(i) $\gamma > (1-\alpha)$;

(ii) $\gamma = (1-\alpha)$;

(iii) $\gamma < (1-\alpha)$;

(iv) any of (i)-(iii) can occur.

The answer box is on the next page.

Consider the prediction model :

$$\log(w_{T+1}) = y_{T+1} = x'_{T+1}\beta_0 + u_{T+1}.$$

We have the prediction error of y_{T+1} :

$$\begin{aligned} e^P_{T+1} &= y_{T+1} - y^P_{T+1} = x'_{T+1}\beta_0 + u_{T+1} - x'_{T+1}\hat{\beta}_T \\ &= u_{T+1} - x'_{T+1}(\hat{\beta}_T - \beta_0) \\ &= u_{T+1} - x'_{T+1}(X'X)^{-1}X'u \end{aligned}$$

As prediction error is a linear combination of normal random variables, we have:

$$e^P_{T+1} \sim N(0, \sigma_0^2(1 + x'_{T+1}(X'X)^{-1}x_{T+1}))$$

Therefore, we have the $100(1-\alpha)\%$ prediction interval of y_{T+1}

$$\left(y^P_{T+1} \pm T_{T-k}(1-\alpha/2)\,\hat{\sigma}_T\sqrt{(1 + x'_{T+1}(X'X)^{-1}x_{T+1})} \right)$$

Which means: $P(y^P_{T+1} \in (y^{LB}_{T+1};\ y^{UB}_{T+1})) = 1-\alpha$

or $P(y^{LB}_{T+1} < y^P_{T+1} < y^{UB}_{T+1}) = 1-\alpha.$

As $\log(w_{T+1}) = y_{T+1} \Rightarrow w^P_{T+1} = \exp(y^P_{T+1})$

And since $f(x) = e^x$ is a monotonically increasing function,

we have $P(\exp(y^{LB}_{T+1}) < \exp(y^P_{T+1}) < \exp(y^{UB}_{T+1})) = 1-\alpha$

$\Leftrightarrow P(w^P_{T+1} \in (\exp(y^{LB}_{T+1});\ \exp(y^{UB}_{T+1})) = 1-\alpha.$

Therefore, $\gamma = 1-\alpha$.

The correct statement is (ii).