Econometrics Qualifying Exam
May 18, 2017

**Answer all questions**. Show work fully and write neatly. Good luck.

1. Answer the following questions.

   a. What does it mean for an estimator to be BLUE?

   b. Under what data generating process is OLS BLUE?

   c. Generalized Least Squares (GLS) extends the OLS estimator to be BLUE in what setting?

   d. True, Partly True, or False: GLS is to OLS as the Generalized Method of Moments is to the Method of Moments. Explain.

2. You have a simple linear demand model for cigarette at the product level:

   $$lnQ_{ijt} = \alpha_i + \beta_i lnp_{ijt} + \varepsilon_{ijt},$$

   where $Q_{ijt}$ and $p_{ijt}$ are the quantity and per pack price of cigarette product $i$ (e.g., full strength Marlboro in pack, light Marlboro in carton, etc.) in state $j$ in year $t$, respectively. You have sales data on 121 products from 48 contiguous states for 10 years. You also have data on the attributes of each cigarette product (e.g., tar level, whether it is mentholated, premium vs. generic, etc.). Suppose you have applied the proper econometric technique to obtain consistent estimates of the coefficients. Let $\hat{\beta}$ be the $121 \times 1$ column vector of slope coefficient estimates whose $i$th element is $\hat{\beta}_i$; and $\Omega$ be the $121 \times 121$ matrix of variance-covariance matrix for $\hat{\beta}$.

   With the above empirical results in hand, you are asked to investigate whether the product-level price elasticities are statistically associated with product attributes. Describe how you would proceed with your analysis? Would OLS work best in this case? If not, what is a better alternative estimator and why?

3. Assume a linear regression model that accurately represents the true data generating process:

$$y_t = \beta_o + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + \varepsilon_t \qquad \text{where } \varepsilon_t \sim N(0, \sigma^2).$$

You have some information, from economic theory, that guides you toward likely values of $\beta_2$ and $\beta_3$.

a) Show how to estimate the model if the values of $\beta_2$ and $\beta_3$ are restricted to $b_2$ and $b_3$.

b) Compare the restricted estimator to the OLS estimator and describe fully when one is preferable to the other and how you propose to measure "preferable."

4. Consider two non-nested models:

(M1) $y1 = X\beta + \varepsilon$ and (M2) $y2 = Z\theta + \omega$

a) How would you go about choosing between these two non-nested models if they had the same dependent variable (y1 = y2)?

b) What would you do if y1 = ln(y2)?

5. Consider the model

$$y_i = y_i^* \times \mathbb{I}\{y_i^* \geq 0\}$$
$$y_i^* = \alpha + \beta x_i + \varepsilon_i$$
$$\varepsilon_i \sim N(0, \sigma^2)$$

where you observe i.i.d. realizations of $(y_i, x_i)$ and $\mathbb{I}$ is an indicator function that equals 1 when the term in the {}s is true. Assume that $x_i$ has full support.

a. Which of $(\alpha, \beta, \sigma)$ are identified? Explain.

b. How would the answer to part a change if we only observed $(\mathbb{I}\{y_i^* \geq 0\}, x_i)$?

c. A researcher attempts to estimate $(\alpha, \beta)$ by running an OLS regression of $y_i$ on $x_i$. Explain why the estimator is biased. What is the sign of the bias?