# ECONOMETRICS FINAL EXAM: B
# EXTRAORDINARY CALL 2022-23

Universidad Carlos III de Madrid
Answer the three questions in 90 minutes

1. (30%) Christensen and Greene (1976) estimated a generalized Cobb-Douglas cost function for electricity generation of the form

$$\ln(C) = \alpha + \beta \ln Q + \gamma \left[\frac{1}{2}(\ln Q)^2\right] + \delta_k \ln P_k + \delta_l \ln P_l + \delta_f \ln P_f + u,$$

where $P_k$, $P_l$ and $P_f$ indicate the unit prices of capital, labour and fuel, respectively, $q$ is output and $C$ are the total costs. The purpose of this generalization was to produce a $U$-shaped average total cost curve. To conform to the underlying theory of production, the cost function should be homogeneous of degree one in the prices, $\delta_k + \delta_l + \delta_f = 1$. (In the questions about tests below, the null and alternative hypotheses, the test statistic, its distribution under the null, the decision rule and the assumptions under which the test is valid must be provided.)

   (a) (7%) Explain in detail the relationship between the error $u$ and the explanatory variables so that $\delta_k$, $\delta_l$ and $\delta_f$ are the expected elasticities of $C$ with respect to $P_k$, $P_l$ and $P_f$. (5%) If this relationship holds, what is the expected elasticity of costs with respect to output for a level $Q = \bar{Q}$? (2%).

   ANSWER*: We notice that*

$$C = e^{\alpha + \beta \ln Q + \gamma\left[\frac{1}{2}(\ln Q)^2\right] + \delta_k \ln P_k + \delta_l \ln P_l + \delta_f \ln P_f} \cdot e^u.$$

*Therefore,*

$$E(C|Q, P_k, P_l, P_f) = e^{\alpha + \beta \ln Q + \gamma\left[\frac{1}{2}(\ln Q)^2\right] + \delta_k \ln P_k + \delta_l \ln P_l + \delta_f \ln P_f} \cdot E(e^u|Q, P_k, P_l, P_f)$$

*and*

$$\ln E(C|Q, P_k, P_l, P_f) = \alpha + \beta \ln Q + \gamma \left[\frac{1}{2}(\ln Q)^2\right] + \delta_k \ln P_k + \delta_l \ln P_l + \delta_f \ln P_f + \ln E(e^u|Q, P_k, P_l, P_f)$$

*The expected elasticity of* $C$ *with respect to* $P_j$, $j = k, l, f$, *is:*

$$\begin{aligned} \xi_{C,P_j} &= \frac{d \ln E(C|Q, P_k, P_l, P_f)}{d \ln P_j} \\ &= \delta_j + \frac{d \ln E(e^u|Q, P_k, P_l, P_f)}{d \ln P_j}, \ j = k, l, f \end{aligned}$$

*Then,*

$$\xi_{C,P_j} = \delta_j \text{ if and only if } E\left(e^u \middle| Q, P_k, P_l, P_f\right) = A, \text{ constant.}$$

*Under the previous condition, the expected elasticity of $C$* with respect to output at the level $Q = \bar{Q}$ is

$$\begin{aligned} \xi_{C,Q}\left(\bar{Q}\right) &= \left.\frac{d\ln E\left(C|\, Q, P_k, P_l, P_f\right)}{d\ln Q}\right]_{Q=\bar{Q}} \\ &= \beta + \gamma \ln \bar{Q}. \end{aligned}$$

(b) (8%) Provide a test for the hypothesis $\delta_k + \delta_l + \delta_f = 1$ against the alternative $\delta_k + \delta_l + \delta_f \neq 1$, at the $100\alpha\%$ significance level by an $F$ test (4%). How would you test that $\delta_k + \delta_l + \delta_f = 1$ against the alternative $\delta_k + \delta_l + \delta_f < 1$ using a significance test on an individual coefficient? (4%)

ANSWER*: There are two ways of constructing the $F$ test for the first part of the question on*

$$\begin{aligned} H_0 &: \delta_k + \delta_l + \delta_f = 1 \\ H_1 &: \delta_k + \delta_l + \delta_f \neq 1, \end{aligned}$$

*one using an $F$ test under homoskedasticity or other using an $F$ test constructed using the general formula, which for one restriction as here is equivalent to squaring the corresponding t-statistic for the significance of $\delta_k + \delta_l + \delta_f - 1$, possibly using robust-to-heteroskedasticity standard errors.*

*First option: Assuming homoskedasticity (which should be mentioned) run the two regressions, the non- or unrestricted and the restricted (replacing $\delta_f = 1 - \delta_k - \delta_l$) ones,*

$$\begin{aligned} \ln(C) &= \alpha + \beta \ln Q + \gamma \left[\frac{1}{2}(\ln Q)^2\right] + \delta_k \ln P_k + \delta_l \ln P_l + \delta_f \ln P_f + u \\ \ln\left(\frac{C}{P_f}\right) &= \alpha + \beta \ln Q + \gamma \left[\frac{1}{2}(\ln Q)^2\right] + \delta_k \ln\left(\frac{P_k}{P_f}\right) + \delta_l \ln\left(\frac{P_l}{P_f}\right) + u, \end{aligned}$$

*and comparing the unrestricted and restricted SCRs,*

$$F = n\frac{SSR_{restricted} - SSR_{unrestricted}}{SSR_{unrestricted}}$$

*as there is only one restriction and the $R^2$ cannot be compared because the restricted model has a different dependent variable ( $n$ can be replaced by $n-5-1$). This $F$ statistic is compared to the critical value from a $\chi^2_1/1 = \chi^2_1$ and the test rejected if $F > \chi^2_{1,\alpha}$ at the $\alpha$ significance level.*

*It is also possible to use the reparametrization of the model in the next question so that the dependent variable is the same in both regressions, i.e.*

$$\begin{aligned}
\ln\left(\frac{C}{P_f}\right) &= \alpha+\beta\ln Q+\gamma\left[\frac{1}{2}(\ln Q)^2\right]+\delta_k\ln\left(\frac{P_k}{P_f}\right)+\delta_l\ln\left(\frac{P_l}{P_f}\right)+\theta\ln P_f+u \\
\ln\left(\frac{C}{P_f}\right) &= \alpha+\beta\ln Q+\gamma\left[\frac{1}{2}(\ln Q)^2\right]+\delta_k\ln\left(\frac{P_k}{P_f}\right)+\delta_l\ln\left(\frac{P_l}{P_f}\right)+u
\end{aligned}$$

*and here the $F$ statistic based on the $R^{2\prime}s$*

$$F=n\frac{R^2_{unrestricted}-R^2_{restricted}}{R^2_{unrestricted}},$$

*could also be used, but otherwise proposing the $R^2$ version in the original model is wrong.*

*Second option: Once we have estimated the non restricted model, the $t$* test is

$$\begin{aligned}
t &= \frac{\hat{\delta}_k+\hat{\delta}_l+\hat{\delta}_f-1}{SE\left(\hat{\delta}_k+\hat{\delta}_l+\hat{\delta}_f-1\right)} \\
&= \frac{\hat{\delta}_k+\hat{\delta}_l+\hat{\delta}_f-1}{\sqrt{\widehat{Var}\left(\hat{\delta}_k\right)+\widehat{Var}\left(\hat{\delta}_l\right)+\widehat{Var}\left(\hat{\delta}_l\right)+2\cdot\widehat{Cov}\left(\hat{\delta}_k,\hat{\delta}_l\right)+2\cdot\widehat{Cov}\left(\hat{\delta}_k,\hat{\delta}_f\right)+2\cdot\widehat{Cov}\left(\hat{\delta}_l,\hat{\delta}_f\right)}}
\end{aligned}$$

*(discount 2% if the formula of $SE\left(\hat{\delta}_k+\hat{\delta}_l+\hat{\delta}_f-1\right)$* is not provided. Therefore, the $F$ statistic is

$$F=t^2=\frac{\left(\hat{\delta}_k+\hat{\delta}_l+\hat{\delta}_f-1\right)^2}{\widehat{Var}\left(\hat{\delta}_k\right)+\widehat{Var}\left(\hat{\delta}_l\right)+\widehat{Var}\left(\hat{\delta}_l\right)+2\cdot\widehat{Cov}\left(\hat{\delta}_k,\hat{\delta}_l\right)+2\cdot\widehat{Cov}\left(\hat{\delta}_k,\hat{\delta}_f\right)+2\cdot\widehat{Cov}\left(\hat{\delta}_l,\hat{\delta}_f\right)}$$

*Under the null hypothesis,*

$$F\underset{aprox.}{\sim}\chi^2_1.$$

*We reject this hypothesis if*

$$F>\chi^2_{1,\alpha}\text{ where }\Pr\left(\chi^2_1>\chi^2_{1,\alpha}\right)=\alpha.$$

*An additional 1% will be given to those how say that $\chi^2_{1,\alpha}=Z^2_{\alpha/2}$, where* $\Pr(Z>Z_a)=\alpha$*, an example ($\alpha=0.05$) is enough.*

*On the other hand, we can express the null and alternative hypotheses as*

$$H_0:\theta=0\text{ vs }H_1:\theta<0,$$

*where $\theta=\delta_k+\delta_l+\delta_f-1$. Then, $\delta_f=\theta+1-\delta_k-\delta_l$ and replacing $\delta_f$ in the model,*

$$\ln(C)=\alpha+\beta\ln Q+\gamma\left[\frac{1}{2}(\ln Q)^2\right]+\delta_k\ln P_k+\delta_l\ln P_l+(\theta+1-\delta_k-\delta_l)\ln P_f+u.$$

*Collecting terms,*

$$\ln\left(\frac{C}{P_f}\right) = \alpha + \beta \ln Q + \gamma \left[\frac{1}{2}(\ln Q)^2\right] + \delta_k \ln\left(\frac{P_k}{P_f}\right) + \delta_l \ln\left(\frac{P_l}{P_f}\right) + \theta \ln P_f + u$$

*We estimate this model by OLS and obtain an estimate of* $\theta$, $\tilde{\theta}$, *and its corresponding standard error,* $SE\left(\tilde{\theta}\right)$. *WE use the t statistic*

$$t = \frac{\tilde{\theta}}{SE\left(\tilde{\theta}\right)},$$

*which is approximately distributed as a standard normal with zero mean and unit variance when* $H_0$ *is true. We reject the null of homogeneity of degree one in prices against the alternative of smaller than one at the* $100 \cdot \alpha\%$ *significance level when* $t < -Z_{\alpha/2}$.

(c) (15%) An estimator of this generalized Cobb-Douglas cost function that satisfies the degree 1 homogeneity condition based on data from 158 firms is

$$\widehat{\ln\left(\frac{C}{P_f}\right)} = \underset{(0.252)}{-6.818} + \underset{(0.031)}{0.403} \ln Q + \underset{(0.004)}{0.061}\left[\frac{1}{2}(\ln Q)^2\right] + \underset{(0.040)}{0.162} \ln\left(\frac{P_k}{P_f}\right) + \underset{(0.047)}{0.152} \ln\left(\frac{P_l}{P_f}\right),$$

The standard errors are in parenthesis. Compute the restricted estimate of $\delta_f$ based on these results (5%). Based on the restricted estimator, provide a 95% confidence interval for $\delta_f$ taking into account that the estimated covariance between the OLS estimators of $\delta_k$ and $\delta_l$ is 0.0018. (10%).

ANSWER: *Under the hypothesis of homogeneity of degree one in prices,*

$$\delta_f = 1 - \delta_k - \delta_l.$$

*In the restricted model the OLS estimates of* $\delta_k$ *and* $\delta_l$ *are* $\hat{\delta}_k = 0.162$ *and* $\hat{\delta}_l = 0.152$. *Therefore, the corresponding estimate of* $\delta_f$ *is*

$$\hat{\delta}_f = 1 - \hat{\delta}_k - \hat{\delta}_l = 1 - 0.162 - 0.152 = 0.686$$

*Then,*

$$\begin{aligned} \widehat{Var}\left(\hat{\delta}_f\right) &= \widehat{Var}\left(\hat{\delta}_k + \hat{\delta}_l\right) \\ &= \widehat{Var}\left(\hat{\delta}_k\right) + \widehat{Var}\left(\hat{\delta}_l\right) + 2 \cdot \widehat{Cov}\left(\hat{\delta}_k, \hat{\delta}_l\right) \\ &= 0.040^2 + 0.047^2 + 2 \cdot 0.0018 \\ &= 0.0075, \end{aligned}$$

*and*

$$SE\left(\hat{\delta}_f\right) = \sqrt{\widehat{Var}\left(\hat{\delta}_f\right)} = \sqrt{0.0075} = 0.086.$$

*Then, the 95% confidence interval for $\delta_f$ is*

$$(0.786 - 1.96 \cdot 0.086, 0.786 + 1.96 \cdot 0.086) = (0.617\,44, 0.954\,56).$$

2. (20%) Consider the model

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \ i = 1, ..., n,$$

where the explanatory variable can be correlated with the error term. Suppose we have an exogenous variable $Z_i$ that is correlated with $X_i$.

(a) (6.5%) Derive algebraically an expression for $\beta_1$ in terms of the covariances of the explained and explanatory variables with respect to the exogenous variable (4.5%). From this expression obtain the instrumental variables estimator of $\beta_1$ (2%).

ANSWER*:* Using the class notes, Topic on IV, pp. 9 and 10,

$$\beta_1 = \frac{Cov\,(Y, Z)}{Cov\,(X, Z)}$$

(no credit will be given if the formula is not derived). The IV estimator is

$$\hat{\beta}_1^{VI} = \frac{\widehat{Cov}\,(Y, Z)}{\widehat{Cov}\,(X, Z)}. \tag{1}$$

(b) (6.5%) Derive an expression for $\beta_1$ in terms of the coefficients of the slopes of the reduced forms of the two endogenous variables (4.5%). From this expression, obtain the estimator of instrumental variables in terms of the OLS estimators of the slopes of the two reduced forms (2%).

ANSWER*: Using the class notes, Topic 6 on IV, (Ch 12 SW) pp. 12 and 13,*

$$\beta_1 = \frac{\gamma_1}{\pi_1},$$

*where $\pi_1$ is the slope coefficient of the reduced form of $X$* and *$\gamma_1$ is the corresponding coefficient in the reduced form of $Y$. The estimate based on the OLS estimates of $\pi_1$ and $\gamma_1$, $\hat{\pi}_1$ and $\hat{\gamma}_1$,* is

$$\hat{\beta}_1^{VI} = \frac{\hat{\gamma}_1}{\hat{\pi}_1}. \tag{2}$$

(c) (7%) Show that the two estimates in (a) and (b) are identical.

ANSWER*: To show that (1) and (2) are identical we notice that*

$$\hat{\gamma}_1 = \frac{\widehat{Cov}\,(Y, Z)}{\widehat{Var}\,(Z)} \text{ and } \hat{\pi}_1 = \frac{\widehat{Cov}\,(X, Z)}{\widehat{Var}\,(Z)},$$

*so that*

$$\hat{\beta}_1^{VI} = \frac{\hat{\gamma}_1}{\hat{\pi}_1} = \frac{\frac{\widehat{Cov}(Y,Z)}{\widehat{Var}(Z)}}{\frac{\widehat{Cov}(X,Z)}{\widehat{Var}(Z)}} = \frac{\widehat{Cov}\,(Y, Z)}{\widehat{Cov}\,(X, Z)}.$$

3. (20%) Consider the model

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i, \;\; i = 1, ..., n,$$

where $W_i$ is a control variable.

(a) (10%) Show that $\beta_1$ has a causal interpretation when $\mathbb{E}\left(u_i | X_i, W_i\right) = \mathbb{E}\left(u_i | W_i\right)$.

(b) (10%) Show that the OLS estimate of $\beta_1$ is consistent (4.5%), but that the estimate of $\beta_2$ might not (5.5%), if $\mathbb{E}\left(u_i | X_i, W_i\right) = \delta_0 + \delta_1 W_i$.

ANSWER: *Class notes for Topic 3 (Ch 7 SW), pp. 56, 58 and 59.*