Econometrics - 1

Endsem Exam

Solutions

A.

Let's first provide an interpretation of model parameters in (1) and (2):

$$G_{q(y)} = \alpha + \beta T_{q(y)} + \eta E_{q(y)} + \lambda D_{q(y)} + u_{q(y)}$$

Here,

- α represents the mean value of GDP, i.e., $\overline{GDP} = \frac{\sum_y \sum_{q(y)} G_{q(y)}}{200}$ when all of the predictor variables in ~~the~~ model (1) are zero.

- $\beta = \frac{\partial G_{q(y)}}{\partial T_{q(y)}}$ represents the on-average change in GDP level, when tractor sales in a given quarter are marginally higher.

- $\eta = \frac{\partial G_{q(y)}}{\partial E_{q(y)}}$ represents the on-average change in GDP upon marginal increase in export volume in a given quarter.

- λ represents the difference in mean GDP level of q3 relative to the overall mean GDP when all other predictor variables are zero.

In other words, mathematically we have.

$$\bar{G}_{q(y) \in \{1(y), 2(y), 3(y), 4(y)\}}\Big|_{T=0, E=0} = \frac{\sum_y \sum_{q(y)} G_{q(y)}}{200} = \alpha$$

$$\bar{G}_{3(y)}\Big|_{T=0, E=0} = \frac{\sum_y G_{3(y)}}{50} = \alpha + \lambda$$

$$\Rightarrow \boxed{\lambda = \bar{G}_{3(y)} - \bar{G}_{q(y)}}$$

For model (2),

- interpretations of α and λ will remain qualitatively similar with only change being that now 'mean of log(GDP)' will replace 'mean of GDP level' in model (1).
- However, in the interpretation of β and η will be different in model (2).

$$\beta = \frac{(\partial G / G)}{\partial T} = \frac{\%\text{ change in } G}{\Delta\text{ change in } T}$$ = on-avg % change in GDP levels upon marginal change in tractor sales

$$\lambda = \frac{\partial G / G}{\partial E} = \frac{\%\text{ change in } G}{\Delta\text{ change in } E}$$ = same as above with vol. of exports

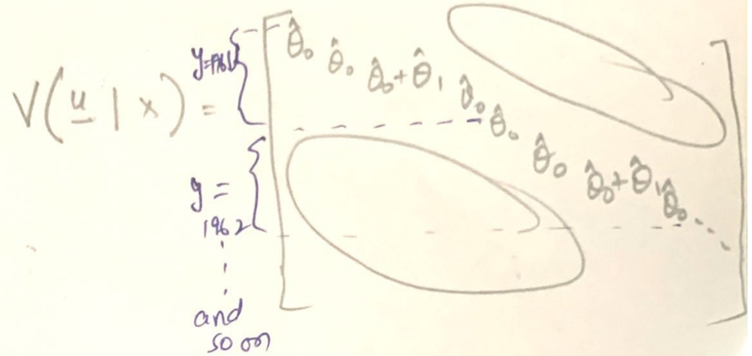
→ The second model, i.e., model (2), is more appropriate for two reasons:

① the log-level model provides the impact of T and E on G in % terms, which controls for the "levels of GDP" through 1960 – 2020. This is important because GDP levels can drastically differ across a 50-yr timeline.

② Model (1) is restrictive in that it only considers a linear change in G that is the same at all levels of T and E, whereas model (2) accounts for diminishing returns assuming T and E will increase over time.

B1. True.

Explanation :- The variance-covariance matrix can be written as under



Hence, $\hat{\sigma}^2_u$ can be negative in following cases

- $\hat{\theta}_0 < 0$ and $\hat{\theta}_1 < 0$ then $\hat{\sigma}^2_u < 0$ for all quarters in all years (OR $\hat{\theta}_0 < 0$, $\hat{\theta}_1 > 0$ and $\hat{\theta}_0 + \hat{\theta}_1 < 0$)
- $\hat{\theta}_0 > 0$, $\hat{\theta}_1 < 0$, $\hat{\theta}_0 + \hat{\theta}_1 < 0$ then $\hat{\sigma}^2_u < 0$ for 3rd quarter of all years
- $\hat{\theta}_0 < 0$, $\hat{\theta}_1 > 0$, $\hat{\theta}_0 + \hat{\theta}_1 > 0$ then $\hat{\sigma}^2_u < 0$ for 1st, 2nd and 4th quarter in all years.

B2: $\sigma_u^2 = \exp(\theta_0 + \theta_1 D_{q(y)})$

step 1: Regress G on T, £ and D w/ intercept,

recover OLS estimates of all model parameters,

recover $\hat{u}_{q(y)}$'s

step 2: Regress $\ln(\hat{u}^2_{q(y)})$ on $D_{q(y)}$ w/ intercept using OLS algo.;

recover $\hat{\theta}_0$ and $\hat{\theta}_1$

step 3 The variance-covariance matrix $\hat{\Omega}$ as follows:

$$\hat{\Omega} = \begin{bmatrix} \exp(\hat{\theta}_0) & & & & & \\ & \exp(\hat{\theta}_0) & & & & \\ & & \exp(\hat{\theta}_0+\hat{\theta}_1) & & & \\ & & & \exp(\hat{\theta}_0) & & \\ & & & & \ddots & \end{bmatrix}$$

y = 1961

and so on

and so on for later years

Step 4 and onwards follow the steps as described in class notes.

$\boxed{C1}$

$$t\text{-stat} = \frac{\hat{\lambda}_{OLS} - 0}{s.e\,(\hat{\lambda}_{OLS})} \sim t_{\underbrace{46}_{\substack{=\text{degrees}\\ \text{of freedom}}}}$$

[50 - 4 = n - k]

where $s.e.(\hat{\lambda}_{OLS}) = \dfrac{\hat{\sigma}^2_u}{\sum_y \sum_{q(y)} (D_{q(y)} - \bar{D})^2}$

Inference :

If $|t\text{-stat}| > t^*_{46,\, 0.025}$ then Reject H_0

If $|t\text{-stat}| \leq t^*_{46,\, 0.025}$ then fail to reject H_0.

where $t^*_{46,0.025}$ is the critical value at 95% confidence level for a two-sided t-test. (to be recovered from a standard t-table)

$\boxed{C2}$ True

Step1 : Run full model as given in (1); recover SSE_{FULL}

Step2 : Run restricted model by putting $\lambda = 0$ in (1); recover $SSE_{Restd.}$

Step3 : $F = \dfrac{(SSE_{RESTD} - SSE_{FULL})/1}{SSE_{FULL}/46} \sim F_{1,46}$

Step4 : steps in inference (as in C1) ← Full credit only if all steps fully described.

C3 Breusch-Pagan-Godfrey Test

Step 1: obtain $\hat{u}_{q(y), OLS}$

Step 2: Regress $\hat{u}^2_{q(y), OLS}$ on $D_{q(y)}$ w/ intercept

Step 3: Compute R^2 for regression in step 2

Step 4: $nR^2 \sim \chi^2_1$

Step 5: steps in inference

if $nR^2 > \chi^2_{1,0.05}$ then Reject H_0

if $nR^2 \leq \chi^2_{1,0.05}$ then fail to reject H_0

D1. Consider model (1) with suggested heteroskedastic structure:

$$G_{q(y)} = \alpha + \beta T_{q(y)} + \eta E_{q(y)} + \lambda D_{q(y)} + u_{q(y)}$$

s.t. $$u_{q(y)} \sim N(0, \theta_0 + \theta_1 D_{q(y)})$$

and $$\underbrace{Cov(u_{q(y)}, u_{q(y')}) = Cov(u_{q(y)}, u_{q'(y)}) = 0,}_{\text{Basically independence of errors across all distinct } q(y)\text{'s.}}$$

$$f(u_{q(y)}) = \frac{1}{\sqrt{2\pi \underbrace{\sigma_u^2}_{=[\theta_0+\theta_1 D_{q(y)}]}}} \exp\left(-\frac{1}{2}\left(\frac{u_i^2}{[\theta_0+\theta_1 D_{q(y)}]}\right)\right)$$

Hence, we can write

$$f(G_{q(y)}) = \frac{1}{\sqrt{2\pi[\theta_0+\theta_1 D_{q(y)}]}} \exp\left(\frac{-1}{2}\left(\frac{[G_{q(y)} - \alpha - \beta T_{q(y)} - \eta E_{q(y)} - \lambda D_{q(y)}]^2}{[\theta_0+\theta_1 D_{q(y)}]}\right)\right) \longrightarrow (1)$$

Now, the likelihood of observing the sample

= Joint probability of observing $G_{q(y)}$'s across all quarters in all years.

Therefore,

$$L(\underbrace{\alpha,\beta,\eta,\lambda,\theta_0,\theta_1}_{\text{parameters}};\underbrace{G,T,E,D}_{\text{Data}}) = \prod_{y}\prod_{q(y)}\left[\underbrace{f(G_{q(y)})}_{\text{from eq}^n\ (1).}\right]$$

and the log-likelihood f^{xn}:-

$$l(\underbrace{\alpha,\beta,\eta,\lambda,\theta_0,\theta_1}_{\text{parameters}};\underbrace{G,T,E,D}_{\text{Data}}) = \ln L(\underbrace{\alpha,\beta,\eta,\lambda,\theta_0,\theta_1}_{\text{parameters}};\underbrace{G,T,E,D}_{\text{Data}})$$

$$= \prod_{y}\prod_{q(y)}\left\{2\pi\left[\theta_0+\theta_1 D_{q(y)}\right]\right\}^{-\frac{1}{2}} - \sum_y\sum_{q(y)}\left\{\underbrace{\frac{(G_{q(y)}-\alpha-\beta T_{q(y)}-\eta E_{q(y)}-\lambda D_{q(y)})^2}{2\left[\theta_0+\theta_1 D_{q(y)}\right]}}_{=A_{q(y)}}\right\}$$

$$= (2\pi)^{-200/2}(\theta_0)^{\frac{-150}{2}}(\theta_0+\theta_1)^{\frac{-50}{2}} - \sum_y\sum_{q(y)}\left\{A_{q(y)}\right\}$$

As a next step we will maximize the log-likelihood of observing the sample given error distribution and choose our model parameters.

That is, $\max\limits_{\alpha,\beta,\eta,\lambda,\theta_0,\theta_1} l(\alpha,\beta,\eta;\lambda,\theta_0,\theta_1;G,T,E,D)$

→ write first order conditions w.r.t each parameters:-

$\partial l/\partial\alpha=0$; $\partial l/\partial\beta=0$; $\partial l/\partial\eta=0$; $\partial l/\partial\lambda=0$; $\partial l/\partial\theta_0=0$; $\partial l/\partial\theta_1=0$ } 6 eq^{ns}, 6 unknowns, SOLVE!

Bonus Question :

Electric vehicles' demand in New Delhi will depend on factors that can be categorised into following categories:

1. Technology : Battery life; mileage ; safety parameters
2. Market segmentation : Luxury ; commercial; etc.
3. Economic information : Prices of electricity and alternative fuels; Price of electric vehicles and their counterparts ; household income; endowments ; etc
4. Public infrastructure : charging stations ; width of roads ; parking spaces, etc.
5. Climatic factors and suitability of electric vehicles

For the second aspect of differential demand structure, one can specify regression models with dummy variables for two-wheelers; three-wheelers and so on.

But a better alternative is to run separate models for each category simultaneously. (why?)