ECON0019 Past Paper 2023

Question A1

1. No, in order to analyze the relative attractiveness of parties, one may wish to learn marginal effect of a characteristics on the probability of choosing a certain party. However, the model given estimates

$$\mathbb{E}[y|x] = \sum_{j=0}^{2} j \cdot \Pr(y = j|x) = \Pr(y = 1|x) + 2\Pr(y = 2|x)$$

Which blends the voter share of party A or B into one estimates.

2. The outcome variable is defined as the percentage of voters who vote for party *j* in segment represented by a series of characteristics *x*

$$\pi_j(x) = \frac{\#\{\text{Voters who vote for } j \text{ given } x\}}{\#\{\text{Subpopulation size given } x\}} \times 100$$

Which is the conditional expectation of an individual-level indicator. We define d_{ij} for each voter i

$$d_{ij} = \begin{cases} 100, & \text{if vote for } j \\ 0, & \text{otherwise} \end{cases}$$

By construction

$$\mathbb{E}\left[d_{ij}\middle|x\right] = \pi_j(x) = 100\Pr(d = 100|x)$$

Therefore, we may regress d_{ij} on all covariates

$$d_{ij} = \beta_{j0} + \beta_{j1}x_1 + \cdots + \beta_{jk}x_k \Rightarrow \mathbb{E}\left[d_{ij}\middle|x\right] = \pi_j(x)$$

3. Given that $u = d_j - \mathbb{E}\left[d_j\middle|X\right]$, the variance of the error term:

$$\begin{aligned}
\text{Var}(u|x) &= \mathbb{E}[u^2|x] - \mathbb{E}[u|x]^2 \\
&= \underbrace{\mathbb{E}[(d - \mathbb{E}[d|x])^2|x]}_{\text{Var}(d|x)} - \mathbb{E}[d - \mathbb{E}[d|x]|x]^2 \\
&= \mathbb{E}[d^2|x] - \mathbb{E}[d|x]^2 \\
&= \Pr(d = 100|x) \cdot 100^2 - (\Pr(d = 100|x) \cdot 100)^2 \\
&= 100\pi_j(x) - \pi_j^2(x) \\
&= \pi_j(x)\left(1 - \pi_j(x)\right)
\end{aligned}$$

Therefore we have shown that the variance of the error term is not constant across observations, we must use heteroskedasticity robust standard error. Let $\hat{r}_{ik}$ be the first-stage residual of the partialling-out method

$$\hat{r}_{ik} = x_{ik} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{i1} - \cdots - \hat{\alpha}_{k-1} x_{i,k-1}$$

$$\operatorname{Var}(\hat{\beta}_k | \chi_n) = \frac{\sum_{i=1}^{n} \hat{r}_{ik}^2 \operatorname{Var}(u_i|\chi_n)}{(\sum_{i=1}^{n} \hat{r}_{ik}^2)^2} = \frac{\frac{1}{n}\sum_{i=1}^{n} \hat{r}_{ik}^2 \operatorname{Var}(u_i|\chi_n)}{\left(\frac{1}{n}\sum_{i=1}^{n} \hat{r}_{ik}^2\right)^2} \times \frac{1}{n}$$

Apply LLN to the denominator, we see that LHS may consistently estimate the RHS in an asymptotic setting:

$$\frac{1}{n}\sum_{i=1}^{n} \hat{r}_{ik}^2 \xrightarrow{p} \mathbb{E}[r_k^2]$$

Similarly, apply LLN to the numerator

$$\frac{1}{n}\sum_{i=1}^{n} \hat{r}_{ik}^2 \operatorname{Var}(u_i|\chi_n) \xrightarrow{p} \mathbb{E}[r_k^2 \operatorname{Var}(u_i|\chi_n)]$$

Therefore, in large sample size

$$\operatorname{Var}(\hat{\beta}_k | \chi_n) \xrightarrow{p} \frac{\mathbb{E}[r_k^2 \operatorname{Var}(u_i|\chi_n)]}{\mathbb{E}[r_k^2]} \times \frac{1}{n}$$

We wish to consistently estimate $\operatorname{Var}(\hat{\beta}_k | \chi_n)$. We have seen that the denominator can be consistently estimated, we now focus on the numerator. Given that

$$\operatorname{Var}(u_i|\chi_n) = \pi_j(x)\left(100 - \pi_j(x)\right)$$

Therefore, the numerator could be consistently estimated by

$$\frac{1}{n}\sum_{i=1}^{n} \hat{r}_{ik}^2 \hat{\pi}_j(x)\left(100 - \hat{\pi}_j(x)\right)$$

And

$$\widehat{\operatorname{Var}}_{\mathrm{HR}}(\hat{\beta}_k | \chi_n) = \frac{\frac{1}{n^2}\sum_{i=1}^{n} \hat{r}_{ik}^2 \hat{\pi}_j(x)\left(100 - \hat{\pi}_j(x)\right)}{\left(\frac{1}{n}\sum_{i=1}^{n} \hat{r}_{ik}^2\right)^2}$$

$$\widehat{\mathrm{se}}_{\mathrm{HR}}(\hat{\beta}_k | \chi_n) = \sqrt{\frac{\sum_{i=1}^{n} \hat{r}_{ik}^2 \hat{\pi}_j(x)\left(100 - \hat{\pi}_j(x)\right)}{\sum_{i=1}^{n} \hat{r}_{ik}^2}}$$

4. Feasible general least square (WLS) would achieve lower variance under LPM model since the form of heteroskedasticity is known under LPM. First, we run OLS and obtain fitted value of the outcome variable $\hat{\pi}_j(x)$. Then estimate variance $\text{Var}\left(d_{ij}\middle|X\right)$
$$\hat{h}_i = \widehat{\text{Var}}\left(d_{ij}\middle|X\right) = \hat{\pi}_j(x)\left[100 - \hat{\pi}_j(x)\right]$$
Run the following OLS regression, which is effectively a WLS regression
$$\frac{d_j}{\sqrt{\hat{h}_{ij}}} = \frac{\beta_{j0}}{\sqrt{\hat{h}_{ij}}} + \beta_{j1}\frac{x_1}{\sqrt{\hat{h}_{ij}}} + \cdots + \beta_{jk}\frac{x_k}{\sqrt{\hat{h}_{ij}}} + \tilde{u}$$
5. Given that
$$\hat{\pi}_A(x) = 100\,\Pr(d_A = 100|X) = \#\text{ who vote for } A \text{ in } 100 \text{ people}$$
$$\hat{\pi}_B(x) = 100\,\Pr(d_B = 100|X) = \#\text{ who vote for } B \text{ in } 100 \text{ people}$$
The percentage of subpopulation who doesn't vote for neither party is
$$\hat{\pi}_0(x) = 100 - \hat{\pi}_A(x) - \hat{\pi}_B(x)$$
6. The percentage of voters who doesn't vote in the whole population can be estimated by
$$\pi_0 = 100 - \frac{1}{n}\sum_{i=1}^{n}\hat{\pi}_A(x) - \frac{1}{n}\sum_{i=1}^{n}\hat{\pi}_B(x) = \frac{100}{n}\sum_{i=1}^{n}\mathbf{1}\{y_i = 0\}$$
With variance
$$\text{Var}(\pi_0) = \frac{\hat{\pi}_0(1-\hat{\pi}_0)}{n}$$

Question A2

1. The point prediction is
$$\widehat{lwage} = 1.90 + 0.35 \cdot 5 + 0.11 \cdot 25 + 0.32 \cdot 10 = 9.6$$
Assuming MLR.6 holds: $u \sim \mathcal{N}(0, \sigma^2)$, it could be shown that an adjustment must be applied to consistently predict wage in level-form
$$wage = \exp\left(\frac{\hat{\sigma}^2}{2}\right)\exp\left(\widehat{lwage}\right) = 14871.5$$
2. The marginal effect of experience on log wage is
$$\frac{\partial lwage}{\partial exper} = 0.35 + 0.22exper$$

We see that holding education constant, the marginal effect is strictly increasing for positive value of years of experience. Therefore, there is no optimal level of years of experience.

3. The t-statistics of the experience squared term is 1.222, which is not statistically significant even at the 10% level. Therefore, this squared term may not be relevant and a linear model appears to be a better choice.

4. The interaction term captures the marginal effect of experience if it depend on the level of education, vice versa.

$$\frac{\partial lwage}{\partial exper} = 0.41 + 0.21educ$$

$$\frac{\partial lwage}{\partial educ} = 0.23 + 0.21exper$$

The first equation reflects that workers with higher level of education tends to have greater reward for experience. Similarly, second equation demonstrates that more experienced workers tend to earn greater reward for more schooling.

5. I would recommend the model with interaction term as the interaction term is statistically significant at 1% level of significance, while the quadratic term is not significant. Furthermore, the model with interaction has greater adjusted R squared.

6. We put the models together

$$lwage = \beta_0 + \beta_1 exper + \beta_2 exper^2 + \beta_3 edu + \beta_4(edu \times exper) + u$$

Use t-test to test whether β_2 or β_4 are statistically different from naught.

Question B1

1. We have the following simultaneous equation system

$$\begin{cases} q_i = \alpha_d(p_i + \tau_i) + d_i \ (1) \\ q_i = \alpha_s p_i + s_i \qquad (2) \end{cases}$$

Substitute (2) into (1) and solve for p_i

$$\alpha_s p_i - \alpha_d p_i = \alpha_d \tau_i + d_i - s_i$$

$$p_i = \frac{\alpha_d}{\alpha_s - \alpha_d}\tau_i + \frac{d_i - s_i}{\alpha_s - \alpha_d}$$

Similarly, rearrange (2) and substitute $p_i = \frac{q_i - s_i}{\alpha_s}$ into the demand equation

$$\alpha_s q_i - \alpha_d q_i = \alpha_s \alpha_d \tau_i - \alpha_d s_i + \alpha_s d_i$$

$$q_i = \frac{\alpha_s \alpha_d}{\alpha_s - \alpha_d} \tau_i + \frac{\alpha_s d_i - \alpha_d s_i}{\alpha_s - \alpha_d}$$

2. OLS regression of the export supply equation will not yield a consistent estimate for α_s because we have seen in the previous question that the regressor p_i correlates with the supply error term s_i, therefore the exogeneity assumption is violated and OLS will no longer be unbiased and consistent. To be precise, we observe from the reduced form equation of p_i, the coefficient of s_i in the supply equation

$$-\frac{1}{\alpha_s - \alpha_d} < 0$$

Therefore, p_i negatively correlates with s_i and therefore would bias the OLS estimator toward naught since $\alpha_s > 0$.
Similarly, since p_i is positively correlated with d_i and $\alpha_d > 0$, the endogeneity issue will bias the OLS estimator toward zero.

3. We may use τ_i as an instrumental variable for the endogenous regressor p_i in the export supply equation. τ_i satisfies the following IV assumptions
   - Exogeneity: Satisfied if we assume tariff policy is independently decided
   - Relevance: From the reduced form equation, $\mathrm{Cov}(p_i, \tau_i) \neq 0$ unless $\frac{\alpha_d}{\alpha_s - \alpha_d} = 0$

   Therefore, we may use 2SLS estimation protocol. The first-stage regression is the reduced form of p_i with an intercept

$$p_i = \pi_0 + \underbrace{\frac{\alpha_d}{\alpha_s - \alpha_d}}_{\pi_1} \tau_i + \underbrace{\frac{d_i - s_i}{\alpha_s - \alpha_d}}_{\omega_i}$$

   We then compute the fitted value of p_i

$$\hat{p}_i = \hat{\pi}_0 + \hat{\pi}_1 \tau_i$$

   For the second stage, since the supply equation is just-identified, we regress q_i on the estimate of endogenous regressor only.

$$q_i = \alpha_s \hat{p}_i + \epsilon_i$$

This 2SLS procedure provides a consistent estimate of $\hat{\alpha}_s$. However, when export supply is highly elastic ($\alpha_s \to \infty$), meaning that this would render $\hat{\pi}_1 \to 0$ and the first stage F-statistics would collapse to zero. Therefore, τ_i becomes a weak instrument when $\hat{\alpha}_s \to \infty$, undermining the consistency of 2SLS estimator.

From an economics standpoint, when the supply curve is basically horizontal, higher import tariff would not alter the producer price p_i. Therefore, the instrument becomes irrelevant to p_i and the relevance condition of IV fails, meaning that 2SLS estimator will no longer be consistent.

4. Since we observe the consumer price $p_{ic} = p_i + \tau_i$, we may rewrite the demand equation

$$q_i = \alpha_d \cdot p_{ic} + d_i$$

The reduced form equation of consumer price in terms of tariff is

$$p_{ic} = p_i + \tau_i = \frac{\alpha_s}{\alpha_s - \alpha_d}\tau_i + \frac{d_i - s_i}{\alpha_s - \alpha_d}$$

We then consider two IV assumptions

- Relevance: Satisfied because p_{ic} positively correlates with τ_i unless $\frac{\alpha_s}{\alpha_s - \alpha_d} = 0$
- Exogeneity: Satisfied by construction of the reduced form.

As both exogeneity and relevance assumptions are satisfied, we may consistently estimate $\hat{\alpha}_d$ using 2SLS. In the classic example of demand and supply, τ_i has a dedicate coefficient and we need another excluded exogenous variable to identify the demand equation. However, in our example, since p_i and τ_i share the same coefficient, τ_i becomes an exogeneous regressor and we don't need another one to identify the demand equation as it does in the classical model.

5. As we assume α_d continue to be homogenous, 2SLS procedure will still consistently estimate $\hat{\alpha}_d$ in the demand equation. However, 2SLS protocol applied to the supply equation

$$q_i = \alpha_i p_i + s_i$$

now specifically identifies the local average treatment effect due to heterogenous effect. The IV estimator for (3) is now weighted by the heterogeneous effect of π_{1i} on p_i

$$\hat{\alpha}_{i,IV} = \frac{\mathbb{E}[\alpha_{1i}\pi_{1i}]}{\mathbb{E}[\pi_{1i}]} = \mathbb{E}\left[\alpha_{1i} \cdot \frac{\pi_{1i}}{\mathbb{E}[\pi_{1i}]}\right]$$

Where π_{1i} is the treatment effect for individual i in the first stage regression

$$p_i = \pi_0 + \underbrace{\frac{\alpha_d}{\alpha_i - \alpha_d}}_{\pi_{1i}}\tau_i + \frac{d_i - s_i}{\alpha_i - \alpha_d}$$

Notice that as π_{i1} is a monotonically decreasing function of α_i, which means $\mathrm{Cov}(\alpha_i, \pi_{1i}) < 0$. Therefore,

$$\mathrm{Cov}(\alpha_i, \pi_{1i}) = \mathbb{E}[\alpha_i\pi_{1i}] - \mathbb{E}[\alpha_i]\mathbb{E}[\pi_{1i}] < 0$$

$$\mathbb{E}[\alpha_i\pi_{1i}] < \mathbb{E}[\alpha_i]\mathbb{E}[\pi_{1i}]$$

$$\hat{\alpha}_{i,IV} = \frac{\mathbb{E}[\alpha_{1i}\pi_{1i}]}{\mathbb{E}[\pi_{1i}]} < \frac{\mathbb{E}[\alpha_i]\mathbb{E}[\pi_{1i}]}{\mathbb{E}[\pi_{1i}]} = \mathbb{E}[\alpha_i]$$

Therefore, the IV estimator suffers a downward bias.

Question B.2

The same as Tutorial Exercise 5 of Term 2