ECON0019 Past Paper 2021

Question A1

1. We may use 2SLS estimation protocol to estimate the $\hat{\beta}$ and $\hat{\gamma}$. First, we regress each endogenous regressors (T_i, W_i) on all instruments and a constant.

$$T_i = \pi_1 + \pi_2 P_i + \pi_3 S_i + u$$
$$W_i = \gamma_1 + \gamma_2 P_i + \gamma_3 S_i + v$$

Obtain the fitted value $\widehat{T}_i$ and $\widehat{W}_i$. Then we regress the outcome variable on the fitted value of endogenous regressors

$$\log Y_i = \alpha + \beta \widehat{T}_i + \gamma \widehat{W}_i + \epsilon_i$$

2. For these 2SLS estimates to be consistent, we would need satisfaction relevance condition and exogeneity condition. The relevance condition states that each instrument must correlates with one endogenous regressors, so that the first stage coefficients of instruments are non-zero. This could be tested by conducting F-test on the first stage coefficients. The exogeneity condition requires each instrument to be uncorrelated with ϵ_i. However, this condition cannot be tested using a J-test since our model is just-identified with 2 endogenous regressors instrumented by 2 exogenous variables.

3. First, the order condition is satisfied because we have two endogenous regressors instrumented by 2 excluded exogenous variables, meaning that our model is just-identified.

Second, we need both instruments to correlate with each endogenous regressors, stronger the correlation, lower the risk of inconsistency introduced due to weak instrument. In our example, P_i tends to be positively correlated with international trade because we would expect more trade if two countries are closer to each other. Furthermore, S_i tends to be positively correlated with the volume of within-country trade while negatively correlated with international trade.

Finally, the exogeneity condition requires both instruments to be independently determined of any variables that could correlate with T_i and W_i. This condition tends to hold true because the boarder and size of countries are determined certainly independently of any trade factors.

4. We substitute the function of the proxy into the original specification

$$\log Y_i = \alpha + \beta T_i + \gamma\eta + \gamma\lambda S_i + \gamma v_i + \epsilon_i$$
$$= (\alpha + \gamma\eta) + \beta T_i + \gamma\lambda S_i + (\gamma v_i + \epsilon_i)$$

Now we have S_i as an included exogenous variable while T_i remains endogenous and has to be instrumented by P_i. We use 2SLS estimation protocol. The first stage is

$$T_i = \pi_0 + \pi_1 P_i + \pi_2 S_i + u_i$$

we compute the fitted value $\hat{T}_i$ and regress the second stage

$$\log Y_i = (\alpha + \gamma\eta) + \beta \hat{T}_i + \gamma\lambda S_i + (\gamma v_i + \epsilon_i)$$

Therefore the coefficient of $\hat{T}_i$ will be a consistent estimate of β when the following conditions are satisfied

$$\text{Cov}(P_i, \epsilon_i) = 0, \text{Cov}(P_i, v_i) = 0$$
$$\text{Cov}(P_i, T_i) \neq 0$$

5. Due to heterogenous effect of international trade (T_i) on income (Y_i), β now depends on i

$$\log Y_i = \alpha + \beta_i T_i + \epsilon_i$$

We still estimate this model using our usual 2SLS estimation protocol, the first stage regression is

$$T_i = \pi_0 + \pi_{1i} P_i + u_i$$

Instead of the average treatment effect, the IV estimator now consistently estimate the Local Average Treatment Effect (LATE), which is an average of β_i weighted by the first stage effect π_{1i}.

$$\text{plim}\,\hat{\beta} = \frac{\text{Cov}(P_i, Y_i)}{\text{Cov}(P_i, T_i)} = \frac{\mathbb{E}[\beta_i \pi_{1i}]}{\mathbb{E}[\pi_{1i}]} = \mathbb{E}\left[\beta_i \frac{\pi_{1i}}{\mathbb{E}[\pi_{1i}]}\right]$$

The ATE is asymptotically recovered if and only if

1) The first-stage effect is homogenous for all countries: $\beta_i = \beta$
2) Instrument affects each country equally: $\pi_{1i} = \pi_1$

The first requirement has been falsified, leaving us to consider the second one. We may argue that the effect of proximity to other countries on international trade may be heterogenous across countries. For example, proximity matters more (higher weight) for rich countries. Therefore, condition (2) is not likely to be satisfied. Furthermore, trading with rich countries tends to be more rewarding because rich countries may offer a higher price for the same goods compared with a poor country. Therefore, trading with rich countries may have high trade-responsiveness to proximity while enjoying larger reward of trade, meaning that $\text{Cov}(\beta_i, \pi_{1i}) > 0$. Furthermore, it's reasonable to infer that higher proximity means greater trade, then $\mathbb{E}[\pi_{1i}] > 0$. Hence

$$\text{plim}\ \hat{\beta}_{IV} - \mathbb{E}[\beta_{1i}] = \frac{\text{Cov}(\beta_{1i}, \pi_{1i})}{\mathbb{E}[\pi_{1i}]} > 0$$

and the IV estimator overestimate the average causal effect.

Question A2

1. We may perform a Durbin (1970) Wald test for autocorrelation. We first regress the original model

$$S_t = \beta_0 + \beta_1 S_{t-1} + E_t$$

and save the OLS residual $\hat{E}_t = S_t - \hat{\beta}_0 - \hat{\beta}_1 S_{t-1}$. Then we run the auxiliary regression of $\hat{E}_t$ on all original regressors and $\hat{u}_{t-1}$

$$\hat{E}_t = \gamma_0 + \gamma_1 S_{t-1} + \rho_1 \hat{E}_{t-1} + v_t$$

Finally we test the null hypothesis $\mathcal{H}: \rho = 0$ with a *t*-statistics on the lagged-residual coefficient $\hat{\rho}$.

$$t = \frac{\hat{\rho}}{se(\hat{\rho})}$$

We reject $\mathcal{H}_0$ if the test statistics is statistically significant, meaning that there exists autocorrelation in our model.

2. Contemporaneous exogeneity requires the following

$$\mathbb{E}[E_t | S_{t-1}] = Cov(E_t, S_{t-1}) = 0$$

Expand the covariance term

$$\begin{aligned}\text{Cov}(E_t, S_{t-1}) &= \text{Cov}(\alpha_0 + \alpha_1 E_{t-1} + V_t, S_{t-1}) \\ &= \alpha_1 \text{Cov}(E_{t-1}, S_{t-1}) + \text{Cov}(V_t, S_{t-1})\end{aligned}$$

Since V_t is assumed to be independently distributed across generations, the second term is zero. However, from the original equation we observe that S_{t-1} depends on E_{t-1}. Therefore, the contemporaneous exogeneity assumption will not be satisfied when $\alpha_1 \neq 0$. In order to get the $\alpha_1\beta_1$ term, we substitute $\mathbb{E}_{t-1}$ into the family equation to get an $AR(2)$ model of S

$$\begin{aligned} S_t &= \beta_0 + \beta_1 S_{t-1} + \alpha_0 + \alpha_1 E_{t-1} + V_t \\ &= \beta_0 + \beta_1 S_{t-1} + \alpha_0 + \alpha_1 (S_{t-1} - \beta_0 - \beta_1 S_{t-2}) + V_t \\ &= \underbrace{(\beta_0 + \alpha_0 - \alpha_1 \beta_0)}_{\gamma_0} + \underbrace{(\alpha_1 + \beta_1)}_{\gamma_1} S_{t-1} \underbrace{-\alpha_1 \beta_1}_{\gamma_2} S_{t-2} + V_t \\ &= \gamma_0 + \gamma_1 S_{t-1} + \gamma_2 S_{t-2} + V_t \end{aligned}$$

Given that V_t is *i.i.d.* across generations, it must be orthogonal to S_{t-1} and S_{t-2}. Therefore, the contemporaneous assumption $\mathbb{E}[V_t | S_{t-1}, S_{t-2}] = 0$ holds, meaning that γ_2 can be consistently estimated by its OLS estimator $\hat{\gamma}_2$. Hence, we can use a *t*-test of null hypothesis $\mathcal{H}_0: \gamma_2 = 0$ to test whether $\alpha_1 \beta_1 = 0$.

3. The outcome variable of interest is

$$S^{C^\star} = \gamma_0 + \gamma_1 \ln S^P + U, \ \ U \sim \mathcal{N}(0,1)$$

However, we only observe due to censoring from the right at $c = 12$

$$S^C = \min\left(12, S^{C^\star}\right)$$

Define the indicator variable $HE^C = \mathbf{1}\{S^{C^\star} \geq 12\}$ of whether the children has 12 or more years of education.

For each observation *i*:

$$\begin{aligned} \Pr\left(HE_i^C = 1 | S^P\right) &= \Pr(\gamma_0 + \gamma_1 \ln S_i^P + U \geq 12 \big| S^P) \\ &= \Pr(U \geq 12 - \gamma_0 - \gamma_1 \ln S_i^P \, | S^P) \\ &= 1 - \Phi(12 - \gamma_0 - \gamma_1 \ln S_i^P) \\ &= \Phi(\gamma_0 + \gamma_1 \ln S_i^P - 12) \\ \Pr(HE_i^C = 0 | S^P) &= \Pr(\gamma_0 + \gamma_1 \ln S_i^P + U < 12 \big| S^P) \\ &= \Pr(U < 12 - \gamma_0 + \gamma_1 \ln S_i^P \, | S^P) \\ &= \Phi(12 - \gamma_0 + \gamma_1 \ln S_i^P) \end{aligned}$$

The log likelihood function is

$$\ell = \sum_{i=1}^{N} \left[HE_i^C \ln[1 - \Phi(12 - \gamma_0 - \gamma_1 \ln S_i^P)] + \left(1 - HE_i^C\right) \ln[\Phi(12 - \gamma_0 + \gamma_1 \ln S_i^P)] \right]$$

The marginal effect of S^P on the conditional probability of children having 12+ years of education is

$$\frac{\partial \Pr(HE^C = 1|S^P)}{\partial S^P} = \frac{\gamma_1}{S^P}\phi(\gamma_0 + \gamma_1 \ln S_i^P - 12)$$

We may estimate the average partial effect $\mathbb{E}\left[\frac{\partial \Pr(HE^C = 1|S^P)}{\partial S^P}\right]$ by its sample mean

$$\frac{\gamma_1}{N}\sum_{i=1}^{N}\frac{1}{S_i^P}\phi(\gamma_0 + \gamma_1 \ln S_i^P - 12)$$

4. OLS will not be consistent since we are selecting data based on the value of the dependent variable, hence the selection rule will correlate with the error term, breaking the exogeneity assumption and making the OLS estimates biased and inconsistent.

5. We will use Heckman correction to restore consistency. First, we use MLE to estimate the Probit model of the selection rule $s_i = \mathbf{1}\{\gamma_0 + \gamma_1 S_i^P + v_i < 12\}$

$$\hat{\gamma}_0, \hat{\gamma}_1 = \arg\max_{\gamma_0, \gamma_1} \ell$$

From this fit compute the inverse Mills' ratio

$$\hat{\lambda}_i = \frac{\phi(\hat{\gamma}_0 + \hat{\gamma}_1 S^P)}{\Phi(\hat{\gamma}_0 + \hat{\gamma}_1 S^P)}$$

Then, we regress the model with sample selection issue plus the inverse Mills' ratio we estimated earlier as an additional regressor.

$$S_i^C = \beta_0 + \beta_1 S_i^P + \rho\hat{\lambda}_i + E$$

Running OLS of this second stage regression would give us consistent estimates of β_0 and β_1.