Economics 241B – 2014 Final

Question 1

An ongoing debate concerns the economic returns to schooling. The goal is to ensure that the observed correlation between schooling and wage rates is not due to correlation between schooling and a worker's ability or other characteristics. To achieve the goal, one approach has been to gather data on identical twins who are raised together, so that the twins are genetically identical and have similar family backgrounds. Let the measure of performance be the logarithms of wages for the first, Y_{1i}, and second, Y_{2i}, twins in the ith pair. A general model specifies that wage rates consist of

$$\begin{aligned} Y_{1i} &= \alpha^{\mathrm{T}} X_i + \beta^{\mathrm{T}} Z_{1i} + \mu_i + U_{1i} \\ Y_{2i} &= \alpha^{\mathrm{T}} X_i + \beta^{\mathrm{T}} Z_{2i} + \mu_i + U_{2i}, \end{aligned} \tag{1}$$

where X_i represents variables that vary by family but not across twins (including age and race), Z_{1i} and Z_{2i} represent variables that may vary across the twins (including education level, union status, job tenure, marital status), μ_i represents the unobservable variables that vary by family, and U_{1i} and U_{2i} represent the unobservable variables that vary across twins. We assume that the coefficients are identical for the two twins.

Expanding the model to capture the correlation between the unobserved family component and the observable variables yields

$$\mu_i = \gamma^{\mathrm{T}} Z_{1i} + \gamma^{\mathrm{T}} Z_{2i} + \delta^{\mathrm{T}} X_i + V_i, \tag{2}$$

where V_i represents the remaining unobservable component of μ_i.

As education levels are contained in Z_{1i} and Z_{2i}, the coefficient of interest is β.

a. Identification: Establish that β is not identified in (1) alone. Establish that β can be identified by adding (2) and be clear about the assumption needed to ensure identification.

b. The reduced-form equations combine (1) and (2):

$$\begin{aligned} Y_{1i} &= (\alpha + \delta)^{\mathrm{T}} X_i + (\beta + \gamma)^{\mathrm{T}} Z_{1i} + \gamma^{\mathrm{T}} Z_{2i} + U'_{1i} \quad U'_{1i} = V_i + U_{1i} \\ Y_{2i} &= (\alpha + \delta)^{\mathrm{T}} X_i + \gamma^{\mathrm{T}} Z_{1i} + (\beta + \gamma)^{\mathrm{T}} Z_{2i} + U'_{2i} \quad U'_{2i} = V_i + U_{2i}. \end{aligned}$$

The reduced form could be estimated by OLS on each equation or by a Seemingly Unrelated Regressions (SUR) estimator. Write down each estimator and determine which estimator is preferred.

c. One could also estimate β using the difference in wages between the twins

$$Y_{1i} - Y_{2i} = \beta^{\mathrm{T}} (Z_{1i} - Z_{2i}) + U_{1i} - U_{2i}. \tag{3}$$

Establish the variance of the OLS estimator in (3).

Question 2

The Pacific Fishery Management Council is responsible for managing fisheries in the US states of California, Oregon, and Washington. They are trying to understand fisherman behavior by studying how harvest responds to changes in fishing effort and stock. In order to do this, they gather data on many fisherman, i, from many different harbors h. There are n_h observations in harbor h, and a total of $n = \sum_{h=1}^{H} n_h$ observations across all H harbors. In particular, the Council is interested in how tuna harvest per fisherman, $Y_{i,h}$, is affected by inputs of fishing effort (which is like labor), $L_{i,h}$, per fisherman and the local stock of fish, S_h, which only varies between harbors.

To model harvest, the Council decides to use the following equation:

$$\ln Y_{i,h} = a + \beta l_{i,h} + \gamma s_h \tag{4}$$

where $a = \ln A$, $l_{i,h} = \ln L_{i,h}$, and $s_h = \ln S_h$. This can be estimated using OLS with the addition of an error term:

$$\ln Y_{i,h} = a + \beta l_{i,h} + \gamma s_h + \epsilon_{i,h} \tag{5}$$

Observe that equation (4) is the natural logarithm of the following structural production function, which the Council has decided is Cobb-Douglas:

$$Y_{i,h} = A \cdot L_{i,h}^{\beta} \cdot S_h^{\gamma} \tag{6}$$

where $A > 0$ captures other components of production. The parameters of interests are (β, γ). A particular concern of the PFMC is whether or not the tuna fishery display constants returns to scale in production.

a. One approach is to estimate (β, γ) by Maximum Likelihood. To find the ML estimator, you must make an assumption on the error structure. Define the error as:

$$error_{i,h} = f(\epsilon_{i,h}) = \frac{Y_{i,h}}{A \cdot L_{i,h}^{\beta} \cdot S_h^{\gamma}}$$

Rewrite equation (6) making an assumption on the error structure ($f(\cdot)$ and $\epsilon_{i,h}$) that will let you perform ML estimation. Then write out the log-likelihood that you would use to estimate (β, γ). Given your assumption on the error structure, are there any other parameters that must be estimated? (Hint: with your assumption, the modified version of equation (6) that you write should easily map to equation (5).)

There may be more than one way to answer this. An answer should note that equation (6) rewritten is:

$$Y_{i,h} = A \cdot L_{i,h}^{\beta} \cdot S_h^{\gamma} \cdot e^{\epsilon_{i,h}}$$

where $\epsilon_{i,h}$ has some distribution (that is, $f(x) = e^x$). Assume that $\epsilon_{i,h} \sim N(0, \sigma^2)$, so that the density (pdf) of $\epsilon_{i,h}$ can be represented by $\phi(\epsilon_{i,h}/\sigma)$. As a side note $\epsilon_{i,h} \sim N(0, \sigma^2)$ implies that $e^{\epsilon_{i,h}}$ is log-normally distributed.

Assuming that the $e^{\epsilon_{i,h}}$ is normal, the likelihood of each observations is $e^{\phi(\epsilon_{i,h}/\sigma)}$. Taking logs, we have

$$\phi(\epsilon_{i,h}/\sigma) = \phi((\ln Y_{i,h} - a - \beta l_{i,h} - \gamma s_h)/\sigma)$$

To find the maximum likelihood estimator, we must find the values of $\xi = (a, \beta, \gamma, \sigma)$ that maximum the joint likelihood. Assuming observations are i.i.d., we get

$$\hat{\xi}_{ML} = \max_{\xi} \prod \phi((\ln Y_{i,h} - a - \beta l_{i,h} - \gamma s_h)/\sigma)$$

The log-likelhood is:

$$\begin{aligned} \sum \ln \phi((\ln Y_{i,h} - &a - \beta l_{i,h} - \gamma s_h)/\sigma) \\ &= -\frac{n}{2} \ln 2\pi - n \ln \sigma - \frac{1}{2\pi\sigma^2} \sum (\ln Y_{i,h} - a - \beta l_{i,h} - \gamma s_h) \end{aligned}$$

We have to additionally estimate a and σ.

b. Under what assumptions is the OLS estimator identical the ML estimator? How does this inform your interpretation of A?

Under the assumption of normality made above, the OLS and ML estimators are identical. Note that this requires conditional homoscedasticity. The parameter A captures whatever misspecification error or stochastic noise from model. This should make it clear that, because we require ϵ to be mean-zero, A just ensures that this condition is met. Thus economic interpretation is difficult.

c. Suppose the results about γ are suggestive but have standard errors that are too large. There is funding to collect 100 more samples. Should these samples be taken from harbors currently represented in the data set, or should these samples be taken from different, new harbors? Support your answer.

First, we should go to new harbors if the underlying DGP in the new harbors is similar to those in the current sample. If the new harbors are in Mexico, for example, they may be uninformative about harvest and effort-stock substitution in US harbors.

If this condition is met, it is most likely better to use new harbors as this likely increase the sampling variance on β, decreasing standard errors. In the rare case that going to more harbors would decrease sampling variance, it would be better to sample from harbors already in the data set.

As with many natural resources, knowledge about tuna stocks is often imperfect or unobservable. In order to decide on the best approach for studying the questions proposed by the Council, you consider the following situations:

[A1] s_h is unobservable

[A2] s_h is observable but measured with error

d. Is the OLS estimator consistent under A1? under A2? If A1 leads to inconsistency, what assumption restores consistency? If A2 leads to inconsistency, what assumption restores consistency? If the OLS estimator is inconsistent under A2, show what it converges in probability to.

This question extended beyond what we've done in class. Therefore the full answer has an expected component that you should be able to get at, and a bonus component.

Expected Component

Under A1, the OLS estimators are inconsistent. Because s is unobserved, l is endogenous – unless $\mathbb{E}[s_h \epsilon_{i,h} | l_{i,h}] = 0$ or $Cov(s_h, l_{i,h}) = 0$. This restores consistency. Because it is highly unlikely that fishers put out the same level of effort when stocks are low (they may put out more, or less), under A1, the zero conditional covariance assumption is not reasonable.

Under A2, the OLS estimators are inconsistent. If we can assume that measurement error is mean zero, then we have another moment condition that can restore identification in a GMM framework (see homework 7).

Because A2 is inconsistent, we want to show what it converges in probability to. This is a hard question: the important intuition is that measurement error in s makes *both β and γ inconsistent.* Put another way, endogeneity in s induces endogeneity in l.

Bonus Component

Maybe this seems unfair, but you might find yourself in a prelim or talk and in over your head. How do you respond?//

Assume deviations from means form to remove the nuisance a. We're going to show.

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \to_p \Lambda \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$$

where $\Lambda \neq I_k$ is the attenuation matrix. Solving for the OLS estimator yields:

$$\Lambda = \begin{pmatrix} \sigma_l^2 & \sigma_{ls} \\ \sigma_{ls} & \sigma_s^2 + \sigma_u^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_l^2 & \sigma_{ls} \\ \sigma_{ls} & \sigma_s^2 \end{pmatrix}$$

where σ_u^2 is the error in the measurement. These are just Var and Cov terms. Note that when $\sigma_u^2 = 0$, then $\Lambda = I$. When inverting the matrix, the determinate will include $\sigma_u^2 = 0$. Thus both terms will be inconsistent. It just so happens that this reduces down to:

$$\Lambda = \begin{pmatrix} \frac{\sigma_l^2 \sigma_s^2 - (\sigma_{ls})^2}{\sigma_l^2(\sigma_s^2 + \sigma_u^2) - (\sigma_{ls})^2} & 0 \\ 0 & \frac{\sigma_l^2 \sigma_s^2 - (\sigma_{ls})^2}{\sigma_l^2(\sigma_s^2 + \sigma_u^2) - (\sigma_{ls})^2} \end{pmatrix}$$

That is, both terms are affected equally by the measurement error.

Now suppose that there exists a possible instrument – the number of reported shark attacks near harbor h, denoted by z_h. The more shark attacks there are, the larger the stock of tuna presumably is.

e. What are the requirements for z_h to be a valid instrument? Do you find this plausible or not? Why?

For z_h to be a valid instrument, it must be the case that:

- $Cov(\epsilon, z|l) = 0$
- $Cov(l, z) \neq 0$ under A1, $Cov(s, z|l) \neq 0$ under A2

This highlights the difference between the assumptions A1 and A2. Under A1, we don't observe the underlying stock, which is correlated with fishing effort. There is thus endogeneity bias on the estimation of β. We will be unable to recover γ. Under A2, we have poor measures of the underlying stock, inducing endogeneity bias in the estimates of γ and β.

The exclusion restriction (first bullet) is somewhat plausible, thought the presence of sharks could drive fisherman away or reduce the efficacy of fishing effort. If those effects are ignored, then bullet 1 is satisfied. As for the second bullet, under A1 it seems like that labor and shark attacks are correlated (shark follow fishers, and fishers avoid sharks), though the sign may be difficult to determine *ex ante*. Under A2, because sharks follow fish stocks around (and so do fishers), this seems reasonable.

f. Assume that z_h is a valid instrument. Under assumption A1, which parameter does z_h restore identification for?

Under A1, there is an endogeneity problem fishing effort, l. The instrument allows resolution of endogeneity and permits consistent estimate of β and β alone. We cannot address γ.

g. Maintain the assumption that z_h is a valid instrument. Under assumption A2, which parameter does z_h restore identification for? You decide to use the 2SLS estimator. Show whether or not this is consistent.

Under A2, there is an endogeneity problem with stock level s, which has the additional effect of causing endogeneity in fishing effort, l. With the instrument, we resolve the endogeneity concern with respect to stock level, and thus for fishing effort as well. That is, we can know estimate γ consistently, and we are also good to estimate β. We require some information in order to be able to estimate the stock elasticity.

Can proceed noting that IV and 2SLS are identical for one endogenous regressor and one instrument, or performing 2SLS. The latter method shown here: The 2SLS estimator can be used to consistently estimate (β, γ). The stages:
1st $s_{i,h} = \delta_0 + \delta_1 z_h + \delta_2 l_{i,h} + \nu_{i,h}$
2nd $Y_{i,h} = a + \beta l_{i,h} + \gamma \hat{s}_{i,h} + u_{i,h}$
As an interesting side note, where as the original specification includes s_h, in the 2SLS implementation we must use $\hat{s}_{i,h}$ – there is individual level variation in the first stage projection of stock. Introducing some notation (everything vectors or matrices):

$$Z = (1, z, l), \quad S = (1, l, s), \quad \hat{S} = Z\hat{\delta} = Z(Z'Z)^{-1}Z's = P_Z s$$

The 2SLS estimator of $\xi = (a, \beta, \gamma)$ is:

$$\hat{\xi}_{2SLS} = (S'P_Z S)^{-1} S'P_Z y = (S'Z(Z'Z)^{-1}Z'S)^{-1} S'Z(Z'Z)^{-1}Z'y$$

Of course, $y = S\xi + u$, so

$$\hat{\xi}_{2SLS} = \xi + (S'P_Z S)^{-1} S'P_Z u$$

where $\hat{S}$ is satisfies $\mathbb{E}[\hat{S}u] = 0$. We can use a LLN in addition to the traditional tools to assert that $\hat{\xi}_{2SLS} \to_p \xi$, and is therefore consistent.

h. Returning the primary question of concern, you desire to test whether or not tuna harvesting shows constant returns to scale. Can you test this under assumption A1? Under assumption A2? Derive the test statistic and the corresponding test.

The test for CRS, which is $\beta + \gamma = 1$, can only be performed under A2. Under A1, there is no estimate of γ to include in a test.

The appropriate test utilizes the Wald statistic, which converges to a χ^2 distribution, here with 1 degree of freedom. Define

$$R = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}, \quad \hat{\xi} = \begin{pmatrix} \hat{a} \\ \hat{\beta} \\ \hat{\gamma} \end{pmatrix}, \quad r = 1$$

Then the Wald statistics is:

$$W = n \cdot (R\hat{\xi} - r)'(R\widehat{Avar(\hat{\xi})}R')^{-1}(R\hat{\xi} - r)$$

Under our standard asymptotic theory, $W \to_d \chi^2(1)$.