Applied Panel Data Analysis – Lecture 6

Christopher F. Parmeter

AGRODEP
March 10-14th, 2015
Addis Ababa, Ethiopia

UmiamI

- Discuss estimation and inference when both individual and time specific heterogeneity is present
- Learn about additional impacts of having both forms of effects present in the unobserved effects model

- Recall the unobserved effects model introduced in the previous lectures

$$y_{it} = x_{it}'\beta + c_i + \varepsilon_{it} \qquad (1)$$

- This model contains unobserved, individual specific heterogeneity
- But it is just as conceivable that we could also have time specific heterogeneity
- That is, there could be factors which vary across time, but influence all individuals equally
- Consider a wage regression and there are changes in the tax structure that apply equally to all workers
- Or consider an output regression for farmers in a given region that periodically suffers from drought

- Our goal in this lecture is to generalize the unobserved effects model to control for both individual specific and time specific heterogeneity
- The unobserved effects model in this case is

$$y_{it} = x'_{it}\beta + c_i + d_t + \varepsilon_{it} \tag{2}$$

- As before, we will need to clarify how we think of the relationship between the unobserved heterogeneity and the observable variables in our model
- The random effects framework will assume $E[c_i|x_{it}] = E[c_i]$ for all t and $E[d_t|x_{it}] = E[d_t]$ for all i

- Before discussing estimation within the random effects framework, we introduce estimation under the fixed effects framework
- Much of the notation here is similar to the one way unobserved effects model we discussed in our previous lectures
- Here the challenge is to account for both time invariant and individual invariant effects in a tractable manner

- To estimate this model we will need to determine the matrix form of the collection of d_t
- Lets call this collection D
- D is an $NT \times T$ matrix that can be written $D = \imath_N \otimes I_T$

APDE
    A Two-Way Unobserved Effects Model
        The Fixed Effects Framework

- Recall from the one way unobserved effects model that the time demeaning transformation eliminated the presence of the fixed effects
- It might be tempting to perform a second transformation that demeans across time periods (instead of across individuals)
- However, this will not suffice
- Consider the following three equations

$$y_{it} = x'_{it}\beta + c_i + d_t + \varepsilon_{it} \tag{3}$$

$$\bar{y}_{i.} = \bar{x}'_{i.}\beta + c_i + T^{-1} + \varepsilon_{i.} \tag{4}$$

$$\bar{y}_{.t} = \bar{x}'_{.t}\beta + N^{-1} + d_t + \bar{\varepsilon}_{.t} \tag{5}$$

- Subtracting (4) from (3) eliminates c_i but introduces T^{-1}
- Subtracting (5) from (3) eliminates d_t but introduces N^{-1}
- So we cannot use just time demeaning within individuals and individual meaning within time periods, we must also account for overall variation across individuals and time

- Wallace and Hussain (1969) develop the appropriate demeaning to eliminate both the time constant and individual constant unobserved heterogeneity
- Their transformation, for a variable z is

$\tilde{z}_{it} = z_{it} - \bar{z}_{i\cdot} - \bar{z}_{\cdot t} + \bar{z}_{\cdot\cdot}$ where $\bar{z}_{\cdot\cdot} = (NT)^{-1} \sum\limits_{i=1}^{N} \sum\limits_{t=1}^{T} z_{it}$

- This amounts to a fourth equation

$$\bar{y}_{\cdot\cdot} = \bar{x}_{\cdot\cdot}'\beta + N^{-1} + T^{-1} + \bar{\varepsilon}_{\cdot\cdot} \tag{6}$$

- In matrix notation our transformation matrix is

$$Q = E_N \otimes E_T, \tag{7}$$

  where $E_N = I_N - \bar{J}_N$ and $E_T = I_T - \bar{J}_T$

- The within estimator has the same structure as in the individual specific unobserved effects model

$$\tilde{\beta} = \left(X'QX\right)^{-1} X'Qy = \left(\tilde{X}'\tilde{X}\right)^{-1} \tilde{X}'\tilde{y} \tag{8}$$

  we just use a different transformation matrix

- In this setup we refer to this as the two way within estimator

- This matrix formulation is convenient because it is easy to describe any estimator
- Letting Q define our transformation matrix, we have the following characterizations

$$Q = I_N \otimes I_T \quad \text{Pooled OLS} \tag{9}$$

$$Q = I_N \otimes E_T \quad \text{within individual} \tag{10}$$

$$Q = E_N \otimes I_T \quad \text{within time} \tag{11}$$

$$Q = E_N \otimes E_T \quad \text{two way within} \tag{12}$$

- We can recover the global intercept as well as the individual and time specific unobserved components

$$\tilde{\alpha} = \bar{y}_{..} - \bar{x}'_{..}\beta \tag{13}$$

$$\tilde{c}_i = (\bar{y}_{i\cdot} - \bar{y}_{..}) - (\bar{x}_{i\cdot} - \bar{x}_{..})'\beta \tag{14}$$

$$\tilde{d}_t = (\bar{y}_{\cdot t} - \bar{y}_{..}) - (\bar{x}_{\cdot t} - \bar{x}_{..})'\beta \tag{15}$$

- As in the one way component case, the two way within estimator prevents estimation of the impact of any variable which is either time invariant (such as gender or region) or individual invariant (such as a tax rate change or a union strike)
- If both c_i and d_t belong in the model then both pooled OLS and the one way within estimator will produce biased estimates
- While the fixed effects framework produces an estimator that is easy to interpret and work with, if the random effects framework is appropriate, then the two way within estimator is no longer efficient

- Under the random effects framework we assume $E[c_i|x_{it}] = E[c_i]$ for all t and $E[d_t|x_{it}] = E[d_t]$ for all i
- Further, we will assume that $d_t \sim IID(0, \sigma_t^2)$ to couple with our initial one way error component assumptions of $c_i \sim IID(0, \sigma_c^2)$ and $\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$
- Our two way error component under the random effects framework is now

$$u_{it} = c_i + d_t + \varepsilon_{it} \qquad (16)$$

- As before, we must determine the variance-covariance structure of this error term and then perform GLS to recover β

- With the given assumptions we have

$$E[u_{it}u_{js}|X] = \begin{cases} \sigma_c^2 + \sigma_t^2 + \sigma_\varepsilon^2, & i = j, t = s \\ \sigma_c^2, & i = j, t \neq s \\ \sigma_t^2, & i \neq j, t = s \\ 0, & i \neq j, t \neq s \end{cases}$$

- In matrix form this is represented as

$$\Omega = \sigma_c^2(I_N \otimes J_T) + \sigma_t^2(J_N \otimes I_T) + \sigma_\varepsilon^2(I_N \otimes I_T) \qquad (17)$$

APDE
    A Two-Way Unobserved Effects Model
        The Random Effects Framework

- For GLS we need Ω^{-1} so we once again need to write Ω in its spectral decomposition form so that inversion is tractable
- Following Fuller and Battese (1974) we replace J_T with $T\bar{J}_T$, J_N with $N\bar{J}_N$, I_N with $E_N + \bar{J}_N$ and I_T with $E_T + \bar{J}_T$
- Collecting terms we have

$$
\begin{aligned}
\Omega =& \sigma_c^2 \left((E_N + \bar{J}_N) \otimes T\bar{J}_T \right) + \sigma_t^2 \left(N\bar{J}_N \otimes (E_T + \bar{J}_T) \right) \\
& + \sigma_\varepsilon^2 \left((E_N + \bar{J}_N) \otimes (E_T + \bar{J}_T) \right) \\
=& T\sigma_c^2 \left((E_N + \bar{J}_N) \otimes \bar{J}_T \right) + N\sigma_t^2 \left(\bar{J}_N \otimes (E_T + \bar{J}_T) \right) \\
& + \sigma_\varepsilon^2 \left((E_N + \bar{J}_N) \otimes (E_T + \bar{J}_T) \right) \\
=& \sigma_\varepsilon^2 \left(E_N \otimes E_T \right) + \left(T\sigma_c^2 + \sigma_\varepsilon^2 \right) \left(E_N \otimes \bar{J}_T \right) \\
& + \left(N\sigma_t^2 + \sigma_\varepsilon^2 \right) \left(\bar{J}_N \otimes E_T \right) + \left(\sigma_\varepsilon^2 + T\sigma_c^2 + N\sigma_t^2 \right) \left(\bar{J}_N \otimes \bar{J}_T \right)
\end{aligned}
$$

- We can write this succinctly as

$$\Omega = \sum_{j=1}^{4} \sigma_j^2 Q_j, \qquad (18)$$

  where $\sigma_1^2 = \sigma_\varepsilon^2$, $\sigma_2^2 = T\sigma_c^2 + \sigma_\varepsilon^2$, $\sigma_3^2 = N\sigma_t^2 + \sigma_\varepsilon^2$ and
  $\sigma_4^2 = \sigma_\varepsilon^2 + T\sigma_c^2 + N\sigma_t^2$
- Also, $Q_1 = E_N \otimes E_T$, $Q_2 = E_N \otimes \bar{J}_T$, $Q_3 = \bar{J}_N \otimes E_T$ and
  $Q_4 = \bar{J}_N \otimes \bar{J}_T$
- Each Q_j is symmetric, idempotent and is orthogonal to the
  other Q_j
- This spectral decomposition implies

$$\Omega^r = \sum_{j=1}^{4} (\sigma_j^2)^r Q_j \qquad (19)$$

- Using $\sigma_\varepsilon \Omega^{-1/2} = \sum\limits_{j=1}^{4} (\sigma_\varepsilon / \sigma_j^{1/2}) Q_j$ for our GLS transformation
  we have

$$\hat{\beta}_{GLS} = (\check{X}'\check{X})^{-1}\check{X}'\check{y} \qquad (20)$$

- Here, $\check{z}$ has typical element

$$\check{z}_{it} = z_{it} - \theta_1 \bar{z}_{i\cdot} - \theta_2 \bar{z}_{\cdot t} - \theta_3 \bar{z}_{\cdot\cdot} \qquad (21)$$

  with $\theta_1 = 1 - \sigma_\varepsilon / \sigma_2$, $\theta_2 = 1 - \sigma_\varepsilon / \sigma_3$ and
  $\theta_3 = 1 - \theta_1 - \theta_2 - \sigma_\varepsilon / \sigma_4$

- As before, note that the fixed effects estimator essentially treats $\theta_1 = \theta_2 = 1$ and $\theta_3 = -1$
- When both N and T are large, σ_2^2, σ_3^2 and σ_4^2 are large, implying that θ_1 and θ_2 are 1 and $\theta_3 = -1$
- Thus, each θ provides information on the importance of demeaning in a particular direction within the panel, θ_1 is for time demeaning, θ_2 is for individual demeaning and θ_3 is for full demeaning

- As in the one way case, the random effects estimator for the two way unobserved effects model can be easily obtained as OLS regression of $\check{y}$ on $\check{X}$
- The only remaining implementation point is how to estimate σ_j
- Following the discussion in Lecture 4 for the estimation of the one way error component model, we use

$$\hat{\sigma}_j^2 = u'Q_j u / tr(Q_j) \tag{22}$$

for $j = 1, 2, 3$
- We can recover $\hat{\sigma}_4^2$ from the other three estimators

- Estimation of the unknown variance parameters follows again from Wallace and Hussain (1969), Amemiya (1971), Nerlove (1971) and Swamy and Aurora (1972)
- The Wallace and Hussain (1969) and Amemiya (1971) approaches are identical to the one way error component setup while the Nerlove (1971) and Swamy and Aurora (1972) approaches require a bit more discussion

- Wallace and Hussain (1969) proposed replacing u with the residuals obtained from OLS estimation of the unobserved effects panel data model
- Under the random effects framework the OLS estimator of β is a consistent estimator so the residuals are reasonable estimates for the unknown u
- The Wallace and Hussain (1969) procedure is
  - Step 1: Estimate the unobserved effects model using pooled OLS, obtain residuals
  - Step 2: Use residuals in place of u in (22)
  - Step 3: Use estimates of σ_j^2 and σ_ε^2 to construct Ω
  - Step 4: Obtain the GLS estimator

- Amemiya (1971) shows that the Wallace and Hussain (1969) approach suffers some theoretical drawbacks
- Amemiya (1971) proposed replacing u with the residuals obtained from within estimation of the unobserved effects panel data model
- Under the random effects framework the within estimator of β is a consistent estimator so the residuals are reasonable estimates for the unknown u
- The Amemiya (1971) procedure is
    - Step 1: Estimate the unobserved effects model using the within estimator, obtain residuals
    - Step 2: Use residuals in place of u in (22)
    - Step 3: Use estimates of σ_j^2 and σ_ε^2 to construct Ω
    - Step 4: Obtain the GLS estimator

- Nerlove (1971) constructs estimates of σ_1^2, σ_2^2 and σ_3^2 by using director estimators of σ_c^2 and σ_t^2
- The Nerlove (1971) procedure is
  - Step 1: Estimate the unobserved effects model using the within estimator, obtain residuals, estimated fixed effects and estimated time effects
  - Step 2a: Construct $\hat{\sigma}_c^2 = \sum_{i=1}^{N} \left(\hat{c}_i - \bar{\hat{c}} \right)^2 / (N - 1)$
  - Step 2b: Construct $\hat{\sigma}_t^2 = \sum_{t=1}^{T} \left(\hat{d}_t - \bar{\hat{d}} \right)^2 / (T - 1)$
  - Step 2c: Construct $\hat{\sigma}_\varepsilon^2 = \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\varepsilon}_{it}^2 / NT$
  - Step 3: Use estimates of σ_c^2, σ_t^2 and σ_ε^2 to construct σ_1^2, σ_2^2 and σ_3^2 and Ω
  - Step 4: Obtain the GLS estimator

- To discuss the Swamy and Arora (1972) GLS implementation we must first discuss between estimation in the two way error component model; the Between estimator that we discussed in Lecture 4 can now be further extended
- When we first discussed the between estimator it was for assessing variation across individual specific means
- With a two way setup we can also have a between estimator for assessing variation purely across time specific means
- Following Maddala (1971) we can decompose the two way random effects GLS estimator into a weighted average of the within estimator and both between estimators

- Consider the following system of $3NT$ observations

$$
\begin{pmatrix} Q_1 y \\ Q_2 y \\ Q_3 y \end{pmatrix} = \begin{pmatrix} Q_1 X \\ Q_2 X \\ Q_3 X \end{pmatrix} \beta + \begin{pmatrix} Q_1 u \\ Q_2 u \\ Q_3 u \end{pmatrix} \tag{23}
$$

- The variance-covariance matrix of this error structure is

$$
\Sigma = \begin{bmatrix} \sigma_1^2 Q_1 & 0 & 0 \\ 0 & \sigma_2^2 Q_2 & 0 \\ 0 & 0 & \sigma_3^2 Q_3 \end{bmatrix}
$$

with inverse

$$
\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} Q_1 & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} Q_2 & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} Q_3 \end{bmatrix}
$$

- GLS estimation using this inverse matrix produces

$$\hat{\beta}_{GLS} = \left(\frac{1}{\sigma_1^2} X'Q_1 X + \frac{1}{\sigma_2^2} X'Q_2 X + \frac{1}{\sigma_3^2} X'Q_3 X \right)^{-1}$$
$$\left(\frac{1}{\sigma_1^2} X'Q_1 y + \frac{1}{\sigma_2^2} X'Q_2 y + \frac{1}{\sigma_3^2} X'Q_3 y \right)$$
$$= \left(X'Q_1 X + \frac{\sigma_1^2}{\sigma_2^2} X'Q_2 X + \frac{\sigma_1^2}{\sigma_3^2} X'Q_3 X \right)^{-1}$$
$$\left(X'Q_1 y + \frac{\sigma_1^2}{\sigma_2^2} X'Q_2 y + \frac{\sigma_1^2}{\sigma_3^2} X'Q_3 y \right) \quad (24)$$

- This derivation can be further decomposed
- Define $\phi_2^2 = \frac{\sigma_1^2}{\sigma_2^2}$ and $\phi_3^2 = \frac{\sigma_1^2}{\sigma_3^2}$
- Let $W = \left(X'Q_1X + \phi_2^2 X'Q_2X + \phi_3^2 X'Q_3X\right)$
- We have

$$
\begin{aligned}
\hat{\beta}_{GLS} =& W^{-1} \Big(X'Q_1X \left(X'Q_1X\right)^{-1} X'Q_1y \\
&\qquad + \phi^2 X'Q_2X \left(\phi^2 X'Q_2X\right)^{-1} \phi^2 X'Q_2y \\
&\qquad + \phi^3 X'Q_3X \left(\phi^2 X'Q_3X\right)^{-1} \phi^2 X'Q_3y \Big) \\
=& W^{-1} \left(X'Q_1X\tilde{\beta} + \phi^2 X'Q_2X\hat{\beta}_{B,i} + \phi^3 X'Q_3X\hat{\beta}_{B,t} \right) \\
=& W_1\tilde{\beta} + W_2\hat{\beta}_{B,i} + W_3\hat{\beta}_{B,t}
\end{aligned}
\tag{25}
$$

- An interesting question is what do these Between estimators capture/measure?
- Notice that the Between individual regression is

$$\bar{y}_{i\cdot} = \alpha + \bar{X}_{i\cdot}'\beta + \bar{u}_{i\cdot}$$

while the Between period regression is

$$\bar{y}_{\cdot t} = \alpha + \bar{X}_{\cdot t}'\beta + \bar{u}_{\cdot t}$$

- Thus, β is identified off of either time variation in the mean of each variable (Between individual) or individual variation in the time means of each variable (Between period)

- Swamy and Arora (1972) proposed estimating σ_1^2, σ_2^2 and σ_3^2 using three different estimators
- The Swamy and Arora (1972) approach does not replace the errors in (22) but constructs entirely different estimators altogether
- The Swamy and Arora (1972) procedure is
    - Step 1: Construct $\hat{\sigma}_1^2$ from the residuals from within estimation of the unobserved effects model
    - Step 2: Construct $\hat{\sigma}_2^2$ from the residuals from between individual estimation of the unobserved effect model
    - Step 3: Construct $\hat{\sigma}_3^2$ from the residuals from between period estimation of the unobserved effect model
    - Step 4: Use these estimates of σ_1^2, σ_2^2 and σ_3^2 to construct Ω
    - Step 5: Obtain the GLS estimator

- Baltagi (1981) presents a Monte Carlo analysis that investigates performance of the four approaches detailed here
- A negative variance estimate for either σ_c^2 or σ_t^2 is more concerning here than in the one-way error component setting
- The reason being that misspecification here cannot be distinguished from the case where the model is properly specified but with a variance component that is small
- While replacing a negative variance estimate with 0 will not impact estimation of β it does interfere with interpretation of the model
- Take suggestion of Maddala and Mount (1973) seriously, when working with the two-way error component model the analyst should always use two different estimation strategies to determine how widely the GLS estimates differ

- In the two way unobserved effects model testing poolability offers many avenues to exploit
- One could test for poolability in the individual dimension (as in the one way case)
- One could test for poolability in the time dimension
- Or one could test in both individual and time dimensions
- Need to be careful not to use up too many degrees of freedom

- In most applied micro settings T is small so we would in general not be interested in testing poolability across each time period as

$$y_t = Z_t \delta_t + \varepsilon_t \qquad (26)$$

  where $y'_t = (y_{1t}, y_{2t}, \ldots, y_{Nt})$, $Z_t = [\imath_N, X_t]$ and $\varepsilon'_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Nt})$

- With a large number of covariates δ_t would not be identified

- If we restrict our focus to strictly additive time and individual effects then three separate hypothesis of poolability that we have are

$$H_{0,1}: c_1 = c_2 = \cdots = c_{N-1} = 0; d_1 = d_2 = \cdots = d_{T-1} = 0$$
$$H_{0,2}: c_1 = c_2 = \cdots = c_{N-1} = 0; d_t \neq 0, \forall t$$
$$H_{0,3}: c_i \neq 0, \forall i; d_1 = d_2 = \cdots = d_{T-1} = 0$$

- $H_{0,1}$ test for the existence of both time and individual effects, $H_{0,2}$ tests for the existence of individual effects while $H_{0,3}$ test for the existence of time effects

- For all three hypotheses let the residual sum of squares from the two way within estimator be denoted $\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* = RSS_{UR}$

- Our statistics for testing each hypothesis are

$$F_1 = \frac{\hat{\varepsilon}_1'\hat{\varepsilon}_1 - \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} \frac{((N-1)(T-1) - K)}{N + T - 2}$$

$$F_2 = \frac{\hat{\varepsilon}_2'\hat{\varepsilon}_2 - \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} \frac{((N-1)(T-1) - K)}{N - 1}$$

$$F_3 = \frac{\hat{\varepsilon}_3'\hat{\varepsilon}_3 - \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} \frac{((N-1)(T-1) - K)}{T - 1}$$

where $\hat{\varepsilon}_j$ are the residuals obtained from the appropriately restricted regression

- The Hausman test for the two way model works identically to the one-way error component model
- A squared statistic based off of the difference between the within estimator and the GLS estimator (both for the two way model) is used
- The regression model from Lecture 5 can also be used:

$$\check{y} = \check{X}\beta + \tilde{X}\delta + \omega, \tag{27}$$

where $\check{z}$ is the GLS two way transformation and $\tilde{z}$ is the two way within transformation

- Kang (1985) also develops a hierarchy of Hausman tests that can performed
  - Test $Cov(c_i, x_{it}) = 0$ assuming $Cov(d_t, x_{it}) \neq 0$ using $\tilde{\beta} - \hat{\beta}_{B,i}$
  - Test $Cov(c_i, x_{it}) = 0$ assuming $Cov(d_t, x_{it}) = 0$ using $\hat{\beta}_{B,i} - \hat{\beta}_{GLS}$
  - Test $Cov(d_t, x_{it}) = 0$ assuming $Cov(c_i, x_{it}) \neq 0$ using $\tilde{\beta} - \hat{\beta}_{B,t}$
  - Test $Cov(d_t, x_{it}) = 0$ assuming $Cov(c_i, x_{it}) = 0$ using $\hat{\beta}_{B,t} - \hat{\beta}_{GLS}$
  - Test $Cov(d_t, x_{it}) = 0$ and $Cov(c_i, x_{it}) = 0$ using $\hat{\beta}_{GLS} - \tilde{\beta}$

- Following Guggenberger (2010), it is not clear in practice if all of these tests should be performed
- Further, if one were to test sequentially, there is no theoretical guidelines for the order in which the tests should be conducted
- Notice that several of the tests are non-nested
- It is not clear which dimension of the model leads to a rejection/failure to reject for a given hypothesis

- The two way unobserved effects model incorporates both time invariant and individual invariant heterogeneity
- Aside from some additional notational complications, the models are estimated in similar fashion to the one way unobserved effects model
- Two different between estimators lead to a trivariate decomposition of the random effects estimator
- Both tests of poolability and the Hausman test offer an array of interesting hypothesis to think about the structure of heterogeneity in one's model