Estimation in panel data with individual effects and AR\((p)\) remainder disturbances

Long Liu, University of Texas at San Antonio

2015 Stata Conference
Outline

- Literature review
- The Model and Estimation
- Applications
- Conclusion
Panel Data Model

Example

\[ y_{it} = x_{it}' \beta + u_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \]

and

\[ u_{it} = \mu_i + \nu_{it}. \]

- Baltagi (2013)
- Stata command: xtreg
Panel Data Model with AR(1) Disturbances

Example

\[ y_{it} = x_{it}' \beta + u_{it}, \quad i = 1, \ldots, N; \ t = 1, \ldots, T, \]
\[ u_{it} = \mu_i + v_{it}, \]
and
\[ v_{it} = \rho v_{i,t-1} + \epsilon_{it} \]

- Baltagi and Li (1991)
- Stata command: *xtregar*
Panel Data Model with AR(p) Disturbances

Example

\[ y_{it} = x_{it}' \beta + u_{it}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \]

\[ u_{it} = \mu_i + v_{it}, \]

and

\[ v_{it} = \rho_1 v_{i,t-1} + \rho_2 v_{i,t-2} + \cdots + \rho_p v_{i,t-p} + \epsilon_{it}. \]

- Baltagi and Liu (2013)
- New user-written Stata command: `xtregarp`
Model in matrix forms

\[ y = X\beta + u \]  \hspace{1cm} (1)

and

\[ u = (l_N \otimes \iota_T) \mu + \nu. \]  \hspace{1cm} (2)

Long Liu, University of Texas at San Antonio

Panel data with AR(p) remainder disturbances
The variance–covariance matrix of \( u \) is

\[
\Omega = I_N \otimes \Lambda, \tag{3}
\]

where

\[
\Lambda = \sigma_\mu^2 J_T + \sigma^2 V,
\]

\( J_T \) is a matrix of ones of dimension \( T \) and \( E(\nu_i \nu_i') = \sigma^2 V \).
Given a $T \times T$ matrix $C$, such that $CVC' = I_T$. The transformed error becomes

$$u^* = (I_N \otimes C) u = (I_N \otimes \iota_T^\alpha) \mu + (I_N \otimes C) \nu,$$

(4)

where $\iota_T^\alpha = C\iota_T = (\alpha_1, \ldots, \alpha_T)'$ is a $T \times 1$ vector.
For AR(1), $C$ is the Prais-Winsten transformation matrix in Baltagi and Li (1991).

$$C = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & -\rho & 1 & 0 \\ 0 & 0 & 0 & 0 & -\rho & 1 \end{bmatrix}$$
The variance-covariance matrix for the transformed disturbance $u^*$ becomes

$$\Omega^* = I_N \otimes \Lambda^*, \quad (5)$$

where

$$\Lambda^* = C\Lambda C' = \sigma^2_\mu J^\alpha_T + \sigma^2 I_T, \quad (6)$$

and $J^\alpha_T = \iota_T^\alpha \iota_T^{\alpha'}$. Define $d^2 = \iota_T^{\alpha'} \iota_T^\alpha = \sum_{t=1}^T \alpha_t^2$, $\bar{J}^\alpha_T = J^\alpha_T / d^2$ and $E_T^\alpha = I_T - \bar{J}^\alpha_T$. We have

$$\Lambda^* = \sigma^2_\alpha \bar{J}^\alpha_T + \sigma^2 E_T^\alpha, \quad (7)$$

where $\sigma^2_\alpha = \sigma^2_\mu d^2 + \sigma^2$. 

Long Liu, University of Texas at San Antonio

Panel data with AR(p) remainder disturbances
Therefore,

$$\sigma \Omega^{*-1/2} = \frac{\sigma}{\sigma_{\alpha}} (I_N \otimes \bar{J}_{T}^{\alpha}) + (I_N \otimes E_{T}^{\alpha}) = (I_N \otimes I_{T}^{\alpha}) - \delta (I_N \otimes \bar{J}_{T}^{\alpha}), \quad (8)$$

where $\delta = 1 - \frac{\sigma}{\sigma_{\alpha}}$. Make the error spherical. $y^{**} = \sigma \Omega^{*-1/2} y^{*}$, and $X^{**}$ and $u^{**}$ are similarly defined. The typical elements

$$y_{it}^{**} = y_{it}^{*} - \delta \alpha_t \frac{\sum_{s=1}^{T} \alpha_s y_{is}^{*}}{\sum_{s=1}^{T} \alpha_s^2}. \quad (9)$$

FE estimator if $\delta = 1$. 

---

**Long Liu, University of Texas at San Antonio**

**Panel data with AR(p) remainder disturbances**
Baltagi and Li (1991) proposed best quadratic unbiased estimators of $\sigma^2$ and $\sigma^2_\alpha$

$$\hat{\sigma}^2_\alpha = u^* (I_N \otimes J_T^{\alpha}) u^* / N \quad \text{and} \quad \hat{\sigma}^2 = u^* (I_N \otimes E_T^{\alpha}) u^* / N (T - 1). \quad (10)$$
Following Baltagi and Li (1994), the (*) transformation defined in (4), is obtained recursively as follows:

\[
\begin{align*}
    y_{i1}^* &= y_{i1} \\
    y_{it}^* &= \left( y_{it} - b_{t,t-1}y_{i,t-1}^* - \cdots - b_{t,1}y_{i,1}^* \right) / \sqrt{a_t} \quad \text{for} \quad t = 2, \ldots, p \\
    y_{it}^* &= \left( y_{it} - \rho_1 y_{i,t-1}^* - \cdots - \rho_p y_{i,t-p}^* \right) / \sqrt{a} \quad \text{for} \quad t = p + 1, \ldots, T,
\end{align*}
\]

where \( a = \sigma^2_\varepsilon / \gamma_0 \).
$a_t$ and $b_{t,s}$ are determined recursively as

$$a_t = 1 - b_{t,t-1}^2 - \cdots - b_{t,2}^2 - b_{t,1}^2 \quad \text{for } t = 2, \ldots, p$$

(12)

and

$$b_{t,1} = r_{t-1}$$

$$b_{t,s} = (r_{t-s} - b_{s,s-1} b_{t,s-1} - \cdots - b_{s,1} b_{t,1}) / \sqrt{a_s} \quad \text{for } s = 2, \ldots, t-1$$

(13)

for $t = 2 \ldots, p$. 

Long Liu, University of Texas at San Antonio  
Panel data with AR($p$) remainder disturbances
Similar to $y_{it}^*$, we can get $\alpha^T = C \alpha_T = (\alpha_1, \ldots, \alpha_T)'$ as follows:

\[
\alpha_1 = 1
\]
\[
\alpha_t = (1 - b_{t,t-1} \alpha_{t-1} - \cdots - b_{t,1} \alpha_1) / \sqrt{a_t} \quad \text{for } t = 2, \ldots, p \tag{14}
\]
\[
\alpha_t = (1 - \sum_{s=1}^{p} \rho_s) / \sqrt{a} \quad \text{for } t = p+1, \ldots, T.
\]
The above transformation depends upon the auto-covariance function of $v_{it}$, that is, $\gamma_s$ for $t = 1 \ldots, p$.

$$\hat{\gamma}_s = \sum_{i=1}^{N} \sum_{t=s+1}^{T} \frac{\tilde{v}_{it}\tilde{v}_{i,t-s}}{N(T-s)}$$

(15)

for $s = 0, \ldots, p$, where $\tilde{v}_{it}$ denotes the within residuals. After getting $\hat{\gamma}_s$, one can compute $\hat{r}_s = \hat{\gamma}_s / \hat{\gamma}_0$ for $s = 1 \ldots, p$. 
Next, we can estimate the $\rho$’s by running the regression of $\tilde{v}_{it}$ on $\tilde{v}_{i,t-1}, \tilde{v}_{i,t-2}, \ldots, \tilde{v}_{i,t-p}$ ($t > p$). Finally

$$\gamma_0 = E \left( v_{it}^2 \right) = \rho_1 \gamma_1 + \rho_2 \gamma_2 + \cdots + \rho_p \gamma_p + \sigma^2_\epsilon. \quad (16)$$

and

$$a = \sigma^2_\epsilon / \gamma_0 = 1 - \rho_1 r_1 - \rho_2 r_2 - \cdots - \rho_p r_p. \quad (17)$$
Step (i): Use the within residuals to compute $\hat{\gamma}_s$ as given in (15). From $\hat{\gamma}_s \ (s = 1, \ldots, p)$, we can get $a_t, b_{t,s}$ and $\alpha_t$ from (12), (13) and (14).
Step (i): Use the within residuals to compute $\hat{\gamma}_s$ as given in (15). From $\hat{\gamma}_s (s = 1, \ldots, p)$, we can get $a_t$, $b_{t,t-s}$ and $\alpha_t$ from (12), (13) and (14).

Step (ii): Get $\rho_1, \rho_2, \ldots, \rho_p$ from the OLS regression of $\tilde{v}_{it}$ on $\tilde{v}_{i,t-1}, \tilde{v}_{i,t-2}, \ldots, \tilde{v}_{i,t-p}$ ($t > p$). Obtain an estimate of $a$ from (17). We now have all the ingredients to compute $y_{it}^*$ and $x_{it}^*$ for $t = 1, \ldots, T$ from (11).
Step (i): Use the within residuals to compute $\hat{\gamma}_s$ as given in (15). From $\hat{\gamma}_s \ (s = 1, \ldots, p)$, we can get $a_t$, $b_{t, t-s}$ and $\alpha_t$ from (12), (13) and (14).

Step (ii): Get $\rho_1, \rho_2, \ldots, \rho_p$ from the OLS regression of $\tilde{v}_{it}$ on $\tilde{v}_{i,t-1}, \tilde{v}_{i,t-2}, \ldots, \tilde{v}_{i,t-p} \ (t > p)$. Obtain an estimate of $a$ from (17). We now have all the ingredients to compute $y^*_it$ and $x^*_it$ for $t = 1, \ldots, T$ from (11).

Step (iii): Compute $\hat{\sigma}^2_\alpha$ and $\hat{\sigma}^2$ in (10) using OLS residuals of $y^*_it$ on $x^*_it$. Then compute $y^{**}_{it}$ and $x^{**}_{it}$ for $t = 1, \ldots, T$ from (9). Run the OLS regression of $y^{**}_{it}$ on $x^{**}_{it}$. This is equivalent to running the GLS regression on (1).
Random Effects (RE) model
```
xtregarp depvar [indepvars] [if] [in], re
```

or Fixed Effects (FE) model
```
xtregarp depvar [indepvars] [if] [in] [weight], fe
```
. use http://www.stata-press.com/data/r13/grunfeld
. xtset
    panel variable:  company (strongly balanced)
    time variable:  year, 1935 to 1954
    delta:  1 year
. xtregar invest mvalue kstock, re p(3)
RE GLS regression with AR(3) disturbances
Number of obs      =       200
Group variable (i): company
Number of groups   =        10
R-sq: within       = 0.7626
         between   = 0.7992
         overall    = 0.7902
Wald chi2(3)       = 380.31
corr(u_i, Xb)      = 0.0000

invest               Coef.     Std. Err.     z    P>|z|     [95% Conf. Interval]
--------------------------------- ----------- ------- ------ ------------------
mvalue               0.0858281   0.0077689   11.05  0.000    0.0706014    0.1010548
kstock               0.3170181   0.0232755   13.62  0.000    0.271399    0.3626371
_cons                -31.2444    25.06929   -1.25  0.213   -80.37931   17.8905

rho1                  0.8171071  (estimated autocorrelation coefficient)
rho2                 -0.2402852  (estimated autocorrelation coefficient)
rho3                 -0.0337094  (estimated autocorrelation coefficient)
sigma_u       74.714532
sigma_e          41.221855
rho_fov             0.7666359  (fraction of variance due to u_i)
theta              0.74992556

Long Liu, University of Texas at San Antonio  Panel data with AR(p) remainder disturbances
An Application of Cornwell and Rupert (1988)

- PSID data of 595 individuals over the period 1976-82
- log wage is regressed on
- years of education (ED),
- weeks worked (WKS),
- years of full-time work experience (EXP),
- occupation (OCC=1, if in a blue-collar occupation),
- residence (SOUTH = 1, if in the South),
- metropolitan area (SMSA = 1, if metropolitan area),
- industry (IND = 1, if in a manufacturing industry),
- marital status (MS = 1, if married),
- sex (FEM = 1, if female),
- race (BLK = 1, if black),
- union coverage (UNION = 1, if in a union contract)
RE estimator: `xtreg, re`
FE estimator: `xtreg, fe`
REAR1 estimator: `xtregar, re`
FEAR1-CO estimator using Cochrane-Orcutt transformation: `xtregar, fe`
FEAR1-PW estimator using Prais-Winsten transformation: `xtregarp, fe p(1)`
. xtregar lwage occ south smsa ind exp exp2 wks ms union fem blk ed, fe rho=rhotype(onestep)
FE (within) regression with AR(1) disturbances
Number of obs = 3570
Group variable: id
Number of groups = 595
R-sq: within = 0.5095
between = 0.0194
overall = 0.0319
F(9,2966) = 342.38
corr(u_i, Xb) = -0.9092
Prob > F = 0.0000

|     | Coef.  | Std. Err. |     t  |    P>|t| |  [95% Conf. Interval] |
|-----|--------|-----------|-------|-------|----------------------|
| lwage |        |           |       |       |                      |
| occ   | -0.0216596 | 0.0153898 | -1.41 |  0.159 | -0.0518355 - 0.0085162 |
| south | 0.0351867  | 0.0421693 |  0.83 |  0.404 | -0.0474973 - 0.1178707 |
| smsa  | -0.0386588 | 0.0231637 | -1.67 |  0.095 | -0.0840774 - 0.0067598 |
| ind   | 0.0110341  | 0.017063  |  0.65 |  0.518 | -0.0224225 - 0.0444907 |
| exp   | 0.1062692  | 0.0036503 | 29.11 |  0.000 | 0.0991119 - 0.1134266 |
| exp2  | -0.0003063 | 0.0000787 | -3.89 |  0.000 | -0.0004606 - 0.000152 |
| wks   | 0.0003698  | 0.0006845 |  0.54 |  0.589 | -0.0009724 - 0.0017119 |
| ms    | -0.0216163 | 0.0220885 | -0.98 |  0.328 | -0.0649267 - 0.0216941 |
| union | 0.0153562  | 0.0166579 |  0.92 |  0.357 | -0.017306 - 0.0480184 |
| _cons | 4.743534   | 0.0516744 | 91.80 |  0.000 | 4.642213 - 4.844856  |

df = 595
rho_ar = 0.14650642
sigma_u = 1.0196127
sigma_e = 0.14794958
rho_fov = 0.97937909 (fraction of variance because of u_i)
F test that all u_i=0: F(594,2966) = 24.91 Prob > F = 0.0000

Long Liu,, University of Texas at San Antonio
Panel data with AR(p) remainder disturbances
. xtregarp lwage occ south smsa ind exp exp2 wks ms union fem blk ed, fe p(1)  
FE GLS regression with AR(1) disturbances
Number of obs = 4165
Group variable (i): id  
Number of groups = 595
R-sq: within = 0.6581  
Obs per group: min = 7
between = 0.0261  
avg = 7.0
overall = 0.0462  
max = 7
Wald chi2(9) = 6836.85  
Prob > chi2 = 0.0000

corr(u_i, Xb) = -0.9097

| lwage | Coef. | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-------|-------|-----------|-------|------|---------------------|
| occ   | -0.022311 | 0.0127311 | -1.75 | 0.080 | -0.0472635 to 0.0026414 |
| south | -0.0071538 | 0.0331086 | -0.22 | 0.829 | -0.0720455 to 0.057738 |
| smsa  | -0.0440674 | 0.0185212 | -2.38 | 0.017 | -0.0803684 to -0.0077665 |
| ind   | 0.0205403  | 0.0143986 | 1.43  | 0.154 | -0.0076805 to 0.048761 |
| exp   | 0.1134939  | 0.0024702 | 45.95 | 0.000 | 0.1086525 to 0.1183353 |
| exp2  | -0.0004294 | 0.0000546 | -7.87 | 0.000 | -0.0005364 to -0.0003224 |
| wks   | 0.0005792  | 0.0005452 | 1.06  | 0.288 | -0.0004894 to 0.0016478 |
| ms    | -0.0332211 | 0.0181076 | -1.83 | 0.067 | -0.0687114 to 0.0022692 |
| union | 0.0293732  | 0.013791  | 2.13  | 0.033 | 0.0023434 to 0.056403 |

rho1   | 0.15024986 | (estimated autocorrelation coefficient) |
sigma_u| 0.46063021  |
sigma_e| 0.50364606  |
rho_fov| 0.45547913  | (fraction of variance due to u_i) |
theta  | 1          |

Long Liu, University of Texas at San Antonio  
Panel data with AR(p) remainder disturbances
<table>
<thead>
<tr>
<th></th>
<th>RE</th>
<th>FE</th>
<th>REARI</th>
<th>FEARI-CO</th>
<th>FEARI-PW</th>
</tr>
</thead>
<tbody>
<tr>
<td>occ</td>
<td>-0.0501</td>
<td>-0.0215</td>
<td>-0.0690</td>
<td>-0.0217</td>
<td>-0.0223</td>
</tr>
<tr>
<td></td>
<td>(0.0166)</td>
<td>(0.0138)</td>
<td>(0.0167)</td>
<td>(0.0154)</td>
<td>(0.0127)</td>
</tr>
<tr>
<td>south</td>
<td>-0.0166</td>
<td>-0.00186</td>
<td>-0.0406</td>
<td>0.0352</td>
<td>-0.00715</td>
</tr>
<tr>
<td></td>
<td>(0.0265)</td>
<td>(0.0343)</td>
<td>(0.0218)</td>
<td>(0.0422)</td>
<td>(0.0331)</td>
</tr>
<tr>
<td>smsa</td>
<td>-0.0138</td>
<td>-0.0425</td>
<td>0.0435</td>
<td>-0.0387</td>
<td>-0.0441</td>
</tr>
<tr>
<td></td>
<td>(0.0200)</td>
<td>(0.0194)</td>
<td>(0.0183)</td>
<td>(0.0232)</td>
<td>(0.0185)</td>
</tr>
<tr>
<td>ind</td>
<td>0.00374</td>
<td>0.0192</td>
<td>0.0144</td>
<td>0.0110</td>
<td>0.0205</td>
</tr>
<tr>
<td></td>
<td>(0.0173)</td>
<td>(0.0154)</td>
<td>(0.0163)</td>
<td>(0.0171)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>exp</td>
<td>0.0821</td>
<td>0.1130</td>
<td>0.0664</td>
<td>0.1060</td>
<td>0.1130</td>
</tr>
<tr>
<td></td>
<td>(0.00285)</td>
<td>(0.00247)</td>
<td>(0.00289)</td>
<td>(0.00365)</td>
<td>(0.00247)</td>
</tr>
<tr>
<td>exp2</td>
<td>-0.0008</td>
<td>-0.0004</td>
<td>-0.0009</td>
<td>-0.0003</td>
<td>-0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.00006)</td>
<td>(0.00005)</td>
<td>(0.00006)</td>
<td>(0.00008)</td>
<td>(0.00005)</td>
</tr>
<tr>
<td>wks</td>
<td>0.0010</td>
<td>0.0008</td>
<td>0.0012</td>
<td>0.0004</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0006)</td>
<td>(0.0008)</td>
<td>(0.0007)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>ms</td>
<td>-0.0746</td>
<td>-0.0297</td>
<td>-0.0668</td>
<td>-0.0216</td>
<td>-0.0332</td>
</tr>
<tr>
<td></td>
<td>(0.0230)</td>
<td>(0.0190)</td>
<td>(0.0237)</td>
<td>(0.0221)</td>
<td>(0.0181)</td>
</tr>
<tr>
<td>union</td>
<td>0.0632</td>
<td>0.0328</td>
<td>0.0682</td>
<td>0.0154</td>
<td>0.0294</td>
</tr>
<tr>
<td></td>
<td>(0.0171)</td>
<td>(0.0149)</td>
<td>(0.0164)</td>
<td>(0.0167)</td>
<td>(0.0138)</td>
</tr>
<tr>
<td>fem</td>
<td>-0.3390</td>
<td>-0.149</td>
<td>-0.3980</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0513)</td>
<td>(0.0401)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>blk</td>
<td>-0.2100</td>
<td>-0.1890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0580)</td>
<td>(0.0424)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ed</td>
<td>0.0997</td>
<td>0.0806</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0058)</td>
<td>(0.0044)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>4165</td>
<td>4165</td>
<td>4165</td>
<td>3570</td>
<td>4165</td>
</tr>
</tbody>
</table>
The standard error of the FEAR1-CO estimator is even larger than the one of FE estimator. This is because the loss of the first time period. The standard error of the FEAR1-PW estimator is smaller than the one of FE estimator.
The FEAR1 estimator

$$\hat{\beta}_{FEAR1} = \left[ X^* (I_N \otimes E_T^\alpha) X^* \right]^{-1} X^* (I_N \otimes E_T^\alpha) y^*$$

If $\rho = 0$, reduces to the FE estimator

$$\hat{\beta}_{FE} = \left[ X' (I_N \otimes E_T) X \right]^{-1} X' (I_N \otimes E_T) y,$$

where $E_T$ is the within matrix, and if $\rho = 1$, reduces to the FD estimator

$$\hat{\beta}_{FD} = \left[ X' (I_N \otimes D'D) X \right]^{-1} X' (I_N \otimes D'D) y,$$

where $D$ is the first difference matrix.

- Let $\rho$ choose between the FE and FD estimators.

The general model regresses bilateral trade (Trade) is regressed on GDP (GDP), similarity in relative size (SIM), differences in relative factor endowments between trading partners (RLF), real exchange rate (RER), both countries belong to the European community (CEE), adopt a common currency (EMU), distance between capital cities (DIST), common border (BOR), common language (LAN).
<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>RE</th>
<th>FE</th>
<th>FD</th>
<th>FEAR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gdp</td>
<td>1.538</td>
<td>2.224</td>
<td>3.053</td>
<td>1.279</td>
<td>2.160</td>
</tr>
<tr>
<td></td>
<td>(0.0130)</td>
<td>(0.0536)</td>
<td>(0.0786)</td>
<td>(0.116)</td>
<td>(0.111)</td>
</tr>
<tr>
<td>Sim</td>
<td>0.839</td>
<td>1.279</td>
<td>1.422</td>
<td>0.596</td>
<td>1.051</td>
</tr>
<tr>
<td></td>
<td>(0.0171)</td>
<td>(0.0495)</td>
<td>(0.0551)</td>
<td>(0.104)</td>
<td>(0.0857)</td>
</tr>
<tr>
<td>Rlf</td>
<td>0.0205</td>
<td>0.0235</td>
<td>0.0181</td>
<td>-0.00247</td>
<td>-0.00124</td>
</tr>
<tr>
<td></td>
<td>(0.00833)</td>
<td>(0.00731)</td>
<td>(0.00718)</td>
<td>(0.00469)</td>
<td>(0.00485)</td>
</tr>
<tr>
<td>Rer</td>
<td>0.0878</td>
<td>0.0562</td>
<td>0.0836</td>
<td>0.402</td>
<td>0.0612</td>
</tr>
<tr>
<td></td>
<td>(0.00388)</td>
<td>(0.00938)</td>
<td>(0.0102)</td>
<td>(0.0172)</td>
<td>(0.0202)</td>
</tr>
<tr>
<td>Cee</td>
<td>0.167</td>
<td>0.305</td>
<td>0.319</td>
<td>0.0493</td>
<td>0.0681</td>
</tr>
<tr>
<td></td>
<td>(0.0264)</td>
<td>(0.0169)</td>
<td>(0.0167)</td>
<td>(0.0145)</td>
<td>(0.0161)</td>
</tr>
<tr>
<td>Emu</td>
<td>0.210</td>
<td>0.274</td>
<td>0.218</td>
<td>-0.0192</td>
<td>0.0333</td>
</tr>
<tr>
<td></td>
<td>(0.0702)</td>
<td>(0.0348)</td>
<td>(0.0342)</td>
<td>(0.0167)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>Dist</td>
<td>-0.698</td>
<td>-0.439</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0224)</td>
<td>(0.123)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bor</td>
<td>0.536</td>
<td>0.277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0334)</td>
<td>(0.196)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lan</td>
<td>0.260</td>
<td>0.655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0336)</td>
<td>(0.190)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ρ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.866</td>
</tr>
<tr>
<td>N</td>
<td>3822</td>
<td>3822</td>
<td>3822</td>
<td>3822</td>
<td>3822</td>
</tr>
</tbody>
</table>
Transforming the data to remove the AR(1) component

After estimating \( \rho \), Baltagi and Wu (1999) derive a transformation of the data that removes the AR(1) component. Their \( C_i(\rho) \) can be written as

\[
y_{itij}^* = \begin{cases} 
(1 - \rho^2)^{1/2} y_{itij} & \text{if } t_{ij} = 1 \\
(1 - \rho^2)^{1/2} \left[ y_{itij} \left( \frac{1}{1 - \rho^2(t_{ij} - t_{i,j-1})} \right)^{1/2} - y_{i,t_{i,j-1}} \left( \frac{\rho^2(t_{ij} - t_{i,j-1})}{1 - \rho^2(t_{i,j} - t_{i,j-1})} \right)^{1/2} \right] & \text{if } t_{ij} > 1
\end{cases}
\]

Using the analogous transform on the independent variables generates transformed data without the AR(1) component. Performing simple OLS on the transformed data leaves behind the residuals \( \mu^* \).
Panel data on 11 large US manufacturing firms over 20 years, for the years 1935–1954.

- Gross investment (invest) is regressed on
- Market value of the firm (mvalue),
- Stock of plant and equipment (kstock)
<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>FEAR1</th>
<th>FEAR2</th>
<th>FEAR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>mvalue</td>
<td>0.110 (0.0119)</td>
<td>0.0917 (0.00867)</td>
<td>0.0836 (0.00808)</td>
<td>0.0827 (0.00828)</td>
</tr>
<tr>
<td>kstock</td>
<td>0.310 (0.0174)</td>
<td>0.322 (0.0250)</td>
<td>0.315 (0.0228)</td>
<td>0.320 (0.0225)</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.664</td>
<td>0.868</td>
<td>0.817</td>
<td></td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>-0.296</td>
<td>-0.240</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_3$</td>
<td>-0.034</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE</td>
<td>52.768</td>
<td>50.551</td>
<td>50.009</td>
<td>50.692</td>
</tr>
<tr>
<td>N</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>
We introduce a new user-written Stata command `xtregarp`.

- Pros: allows autocorrelation besides AR(1); use PW transformation for FE estimator
- Cons: do not allow unbalanced panel data.
We introduce a new user-written Stata command `xtregarp`.

It performs the RE or FE estimator with AR($p$) disturbances in Baltagi and Liu (2013).
We introduce a new user-written Stata command `xtregarp`.

It performs the RE or FE estimator with AR($p$) disturbances in Baltagi and Liu (2013).

Pros: allows autocorrelation besides AR(1); use PW transformation for FE estimator.
Conclusion

- We introduce a new user-written Stata command `xtregarp`.
- It performs the RE or FE estimator with AR($p$) disturbances in Baltagi and Liu (2013)
- Pros: allows autocorrelation besides AR(1); use PW transformation for FE estimator
- Cons: do not allow unbalanced panel data.
Thank you!!!