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

$$y_{it} = x'_{it}\beta + u_{it}, \qquad i = 1, \ldots, N; \; t = 1 \ldots, T,$$

and

$$u_{it} = \mu_i + v_{it}.$$

- Baltagi (2013)
- Stata command: `xtreg`

Panel Data Model with AR(1) Disturbances

Example

$$y_{it} = x_{it}'\beta + u_{it}, \qquad i = 1, \ldots, N; \ \ t = 1 \ldots, T,$$

$$u_{it} = \mu_i + v_{it},$$

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

$$y_{it} = x_{it}'\beta + u_{it}, \qquad i = 1, \ldots, N; \; t = 1 \ldots, T,$$

$$u_{it} = \mu_i + v_{it},$$

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

$$y = X\beta + u \tag{1}$$

and

$$u = \left(I_N \otimes \iota_T\right)\mu + \nu. \tag{2}$$

Variance–covariance matrix

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 $E\left(v_i v_i'\right) = \sigma^2 V$.

Transformed error

Given a $T \times T$ matrix C, such that $CVC' = I_T$. The transformed error becomes

$$u^* = (I_N \otimes C)\, u = (I_N \otimes \iota_T^\alpha)\, \mu + (I_N \otimes C)\, \nu, \tag{4}$$

where $\iota_T^\alpha = C\iota_T = (\alpha_1, \ldots, \alpha_T)'$ is a $T \times 1$ vector.

Transformation matrix

For AR(1), C is the Prais-Winsten transformation matrix in Baltagi and Li (1991).

$$C = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & -\rho & 1 & 0 \\ 0 & 0 & 0 & 0 & -\rho & 1 \end{bmatrix}$$

Variance–covariance matrix of transformed error

The variance-covariance matrix for the transformed disturbance u^* becomes

$$\Omega^* = I_N \otimes \Lambda^*, \tag{5}$$

where

$$\Lambda^* = C\Lambda C' = \sigma_\mu^2 J_T^\alpha + \sigma^2 I_T, \tag{6}$$

and $J_T^\alpha = \iota_T^\alpha \iota_T^{\alpha\prime}$. Define $d^2 = \iota_T^{\alpha\prime}\iota_T^\alpha = \sum_{t=1}^T \alpha_t^2$, $\bar{J}_T^\alpha = J_T^\alpha / d^2$ and $E_T^\alpha = I_T - \bar{J}_T^\alpha$. We have

$$\Lambda^* = \sigma_\alpha^2 \bar{J}_T^\alpha + \sigma^2 E_T^\alpha, \tag{7}$$

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

Two-step transformation

Therefore,

$$\sigma\Omega^{*-1/2} = \frac{\sigma}{\sigma_\alpha}\left(I_N \otimes \bar{J}_T^\alpha\right) + \left(I_N \otimes E_T^\alpha\right) = \left(I_N \otimes I_T^\alpha\right) - \delta\left(I_N \otimes \bar{J}_T^\alpha\right), \qquad (8)$$

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

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

FE estimator if $\delta = 1$.

Parameters estimation

Baltagi and Li (1991) proposed best quadratic unbiased estimators of σ^2 and σ_α^2

$$\hat{\sigma}_\alpha^2 = u^{*\prime}\left(I_N \otimes \bar{J}_T^\alpha\right) u^* / N \text{ and } \hat{\sigma}^2 = u^{*\prime}\left(I_N \otimes E_T^\alpha\right) u^* / N\left(T-1\right). \tag{10}$$

Transformation

Following Baltagi and Li (1994), the (*) transformation defined in (4), is obtained recursively as follows:

$$\begin{aligned} y^*_{i1} &= y_{i1} \\ y^*_{it} &= \left(y_{it} - b_{t,t-1}y^*_{i,t-1} - \cdots - b_{t,1}y^*_{i,1}\right)/\sqrt{a_t} \quad \text{for } t = 2,\ldots,p \\ y^*_{it} &= \left(y_{it} - \rho_1 y_{i,t-1} - \cdots - \rho_p y_{i,t-p}\right)/\sqrt{a} \quad \text{for } t = p+1,\ldots,T, \end{aligned} \tag{11}$$

where $a = \sigma^2_\epsilon / \gamma_0$.

Transformation

a_t and $b_{t,s}$ are determined recursively as

$$a_t = 1 - b_{t,t-1}^2 - \cdots - b_{t,2}^2 - b_{t,1}^2 \quad \text{for } t = 2, \ldots, p \tag{12}$$

and

$$\begin{aligned} b_{t,1} &= r_{t-1} \\ b_{t,s} &= \left(r_{t-s} - b_{s,s-1} b_{t,s-1} - \cdots - b_{s,1} b_{t,1}\right) / \sqrt{a_s} \quad \text{for } s = 2, \ldots, t-1 \end{aligned} \tag{13}$$

for $t = 2 \ldots, p$.

Transformation

Similar to y_{it}^*, we can get $\iota_T^\alpha = C\iota_T = (\alpha_1, \ldots, \alpha_T)'$ as follows:

$$\begin{array}{ll} \alpha_1 = 1 & \\ \alpha_t = \left(1 - b_{t,t-1}\alpha_{t-1} - \cdots - b_{t,1}\alpha_1\right) / \sqrt{a_t} & \text{for } t = 2, \ldots, p \\ \alpha_t = \left(1 - \sum_{s=1}^{p} \rho_s\right) / \sqrt{a} & \text{for } t = p+1, \ldots, T. \end{array} \tag{14}$$

Auto-covariance function

The above transformation depends upon the auto-covariance function of v_{it}, that is, γ_s for $t = 1 \ldots, p$.

$$\hat{\gamma}_s = \sum_{i=1}^{N} \sum_{t=s+1}^{T} \frac{\tilde{v}_{it} \tilde{v}_{i,t-s}}{N(T-s)} \tag{15}$$

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 ρ'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

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

and

$$a = \sigma_\epsilon^2/\gamma_0 = 1 - \rho_1 r_1 - \rho_2 r_2 - \cdots - \rho_p r_p. \tag{17}$$

Precedures

1. Step (i): Use the within residuals to compute $\hat{\gamma}_s$ as given in (15). From $\hat{\gamma}_s$ $(s = 1...., p)$, we can get a_t, $b_{t,t-s}$ and α_t from (12), (13) and (14).

1. Step (i): Use the within residuals to compute $\hat{\gamma}_s$ as given in (15). From $\hat{\gamma}_s$ $(s = 1...., p)$, we can get a_t, $b_{t,t-s}$ and α_t from (12), (13) and (14).
2. Step (ii): Get $\rho_1, \rho_2, \cdots, \rho_p$ from the OLS regression of $\tilde{v}_{it}$ on $\tilde{v}_{i,t-1}, \tilde{v}_{i,t-2}, \cdots, \tilde{v}_{i,t-p}$ $(t > p)$. 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).

1. Step (i): Use the within residuals to compute $\hat{\gamma}_s$ as given in (15). From $\hat{\gamma}_s$ $(s = 1...., p)$, we can get a_t, $b_{t,t-s}$ and α_t from (12), (13) and (14).
2. Step (ii): Get $\rho_1, \rho_2, \cdots, \rho_p$ from the OLS regression of $\tilde{v}_{it}$ on $\tilde{v}_{i,t-1}, \tilde{v}_{i,t-2}, \cdots, \tilde{v}_{i,t-p}$ $(t > p)$. 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).
3. 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).

Syntax

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
. xtregarp 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                            Obs per group: min =        20
       between = 0.7992                                           avg =      20.0
       overall = 0.7902                                           max =        20
                                                   Wald chi2(3)       =    380.31
corr(u_i, Xb)      = 0.0000                        Prob > chi2        =    0.0000
------------------------------------------------------------------------------
      invest |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      mvalue |   .0858281   .0077689    11.05   0.000     .0706014    .1010548
      kstock |   .3170181   .0232755    13.62   0.000      .271399    .3626371
       _cons |   -31.2444   25.06929    -1.25   0.213    -80.37931     17.8905
-------------+----------------------------------------------------------------
        rho1 |  .81710709   (estimated autocorrelation coefficient)
        rho2 | -.24028523   (estimated autocorrelation coefficient)
        rho3 |  -.0337094   (estimated autocorrelation coefficient)
     sigma_u |  74.714532
     sigma_e |  41.221855
     rho_fov |   .7666359   (fraction of variance due to u_i)
       theta |  .74992556
------------------------------------------------------------------------------
```

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)

An Application of Cornwell and Rupert (1988)

- 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 rhotype
FE (within) regression with AR(1) disturbances  Number of obs          =      3570
Group variable: id                              Number of groups       =       595
R-sq:  within  = 0.5095                         Obs per group: min =         6
       between = 0.0194                                        avg =       6.0
       overall = 0.0319                                        max =         6
                                                F(9,2966)              =    342.38
corr(u_i, Xb)  = -0.9092                        Prob > F               =    0.0000
```

lwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
occ	-.0216596	.0153898	-1.41	0.159	-.0518355	.0085162
south	.0351867	.0421693	0.83	0.404	-.0474973	.1178707
smsa	-.0386588	.0231637	-1.67	0.095	-.0840774	.0067598
ind	.0110341	.017063	0.65	0.518	-.0224225	.0444907
exp	.1062692	.0036503	29.11	0.000	.0991119	.1134266
exp2	-.0003063	.0000787	-3.89	0.000	-.0004606	-.000152
wks	.0003698	.0006845	0.54	0.589	-.0009724	.0017119
ms	-.0216163	.0220885	-0.98	0.328	-.0649267	.0216941
union	.0153562	.0166579	0.92	0.357	-.017306	.0480184
_cons	4.743534	.0516744	91.80	0.000	4.642213	4.844856
rho_ar	.14650642					
sigma_u	1.0196127					
sigma_e	.14794958					
rho_fov	.97937909	(fraction of variance because of u_i)				

```
F test that all u_i=0:     F(594,2966) =     24.91             Prob > F = 0.0000
```

```
. 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
corr(u_i, Xb)       = -0.9097                      Prob > chi2           =    0.0000
------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         occ |   -.022311   .0127311    -1.75   0.080    -.0472635    .0026414
       south |  -.0071538   .0331086    -0.22   0.829    -.0720455     .057738
        smsa |  -.0440674   .0185212    -2.38   0.017    -.0803684   -.0077665
         ind |   .0205403   .0143986     1.43   0.154    -.0076805     .048761
         exp |   .1134939   .0024702    45.95   0.000     .1086525    .1183353
        exp2 |  -.0004294   .0000546    -7.87   0.000    -.0005364   -.0003224
         wks |   .0005792   .0005452     1.06   0.288    -.0004894    .0016478
          ms |  -.0332211   .0181076    -1.83   0.067    -.0687114    .0022692
       union |   .0293732    .013791     2.13   0.033     .0023434     .056403
-------------+----------------------------------------------------------------
        rho1 |  .15024986   (estimated autocorrelation coefficient)
     sigma_u |  .46063021
     sigma_e |  .50364606
     rho_fov |  .45547913   (fraction of variance due to u_i)
       theta |          1
------------------------------------------------------------------------------
```

	RE	FE	REAR1	FEAR1-CO	FEAR1-PW
occ	-0.0501	-0.0215	-0.0690	-0.0217	-0.0223
	(0.0166)	(0.0138)	(0.0167)	(0.0154)	(0.0127)
south	-0.0166	-0.00186	-0.0406	0.0352	-0.00715
	(0.0265)	(0.0343)	(0.0218)	(0.0422)	(0.0331)
smsa	-0.0138	-0.0425	0.0435	-0.0387	-0.0441
	(0.0200)	(0.0194)	(0.0183)	(0.0232)	(0.0185)
ind	0.00374	0.0192	0.0144	0.0110	0.0205
	(0.0173)	(0.0154)	(0.0163)	(0.0171)	(0.0144)
exp	0.0821	0.1130	0.0664	0.1060	0.1130
	(0.00285)	(0.00247)	(0.00289)	(0.00365)	(0.00247)
exp2	-0.0008	-0.0004	-0.0009	-0.0003	-0.0004
	(0.00006)	(0.00005)	(0.00006)	(0.00008)	(0.00005)
wks	0.0010	0.0008	0.0012	0.0004	0.0006
	(0.0008)	(0.0006)	(0.0008)	(0.0007)	(0.0005)
ms	-0.0746	-0.0297	-0.0668	-0.0216	-0.0332
	(0.0230)	(0.0190)	(0.0237)	(0.0221)	(0.0181)
union	0.0632	0.0328	0.0682	0.0154	0.0294
	(0.0171)	(0.0149)	(0.0164)	(0.0167)	(0.0138)
fem	-0.3390		-0.3980		
	(0.0513)		(0.0401)		
blk	-0.2100		-0.1890		
	(0.0580)		(0.0424)		
ed	0.0997		0.0806		
	(0.0058)		(0.0044)		
N	4165	4165	4165	3570	4165

- 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.

An Application of Gravity Data Set in Serlenga and Shin (2007)

The FEAR1 estimator

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

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

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

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

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

where D is the first difference matrix.

- Let ρ choose between the FE and FD estimators.

An Application of Gravity Data Set in Serlenga and Shin (2007)

- bilateral trade flows among 15 European countries over the period 1960–2001.
- 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).

	OLS	RE	FE	FD	FEAR1
Gdp	1.538	2.224	3.053	1.279	2.160
	(0.0130)	(0.0536)	(0.0786)	(0.116)	(0.111)
Sim	0.839	1.279	1.422	0.596	1.051
	(0.0171)	(0.0495)	(0.0551)	(0.104)	(0.0857)
Rlf	0.0205	0.0235	0.0181	-0.00247	-0.00124
	(0.00833)	(0.00731)	(0.00718)	(0.00469)	(0.00485)
Rer	0.0878	0.0562	0.0836	0.402	0.0612
	(0.00388)	(0.00938)	(0.0102)	(0.0172)	(0.0202)
Cee	0.167	0.305	0.319	0.0493	0.0681
	(0.0264)	(0.0169)	(0.0167)	(0.0145)	(0.0161)
Emu	0.210	0.274	0.218	-0.0192	0.0333
	(0.0702)	(0.0348)	(0.0342)	(0.0167)	(0.0219)
Dist	-0.698	-0.439			
	(0.0224)	(0.123)			
Bor	0.536	0.277			
	(0.0334)	(0.196)			
Lan	0.260	0.655			
	(0.0336)	(0.190)			
ρ					0.866
N	3822	3822	3822	3822	3822

Transforming the data to remove the AR(1) component

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

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

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 μ^*.

An Application of Grunfeld Data Set

- 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)

	FE	FEAR1	FEAR2	FEAR3
mvalue	0.110	0.0917	0.0836	0.0827
	(0.0119)	(0.00867)	(0.00808)	(0.00828)
kstock	0.310	0.322	0.315	0.320
	(0.0174)	(0.0250)	(0.0228)	(0.0225)
ρ_1		0.664	0.868	0.817
ρ_2			-0.296	-0.240
ρ_3				-0.034
RMSE	52.768	50.551	50.009	50.692
N	200	200	200	200

Conclusion

- We introduce a new user-written Stata command `xtregarp`.

Conclusion

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

Conclusion

- We introduce a new user-written Stata command `xtregarp`.
- It perfrom 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 perfrom 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!!!