Panel Econometrics

Paulo M. M. Rodrigues

February 2022

Today's outline

Specification tests
Unobserved individual-specific effects

1. Wooldridge's test of unobserved individual-specific effects
2. Breusch-Pagan LM test of unobserved individual-specific effects
3. Tests for autocorrelation
4. Up next

Outline

1. Wooldridge's test of unobserved individual-specific effects
2. Breusch-Pagan LM test of unobserved individual-specific effects
3. Tests for autocorrelation
4. Up next

Why test for unobserved individual-specific effects?

Recall the error-components model:

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + v_{it}, \qquad v_{it} = c_i + u_{it}$$

If $\sigma_c^2 = 0$, then c_i vanishes and the variance matrix of v_i becomes

$$\boldsymbol{\Omega} = \sigma_u^2 \mathbf{I}_T.$$

This leads to

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{RE} &= \left(\sum_{i=1}^{N} \mathbf{X}_i' \hat{\boldsymbol{\Omega}}^{-1} \mathbf{X}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i' \hat{\boldsymbol{\Omega}}^{-1} \mathbf{y}_i\right) = \left(\frac{1}{\hat{\sigma}_u^2} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{X}_i\right)^{-1} \frac{1}{\hat{\sigma}_u^2} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{y}_i \\ &= \left(\sum_{i=1}^{N} \mathbf{X}_i' \mathbf{X}_i\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i' \mathbf{y}_i\right) = \hat{\boldsymbol{\beta}}_{POLS}\end{aligned}$$

Hence, pooled OLS is consistent and efficient.

Hypotheses

$H_0 : \sigma_c^2 = 0$ versus $H_1 : \sigma_c^2 \neq 0$

This implies for $\mathrm{E}(\mathbf{v}_i'\mathbf{v}_i) = \mathbf{\Omega}$

$$H_0 : \mathbf{\Omega} = \begin{bmatrix} \sigma_u^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_u^2 \end{bmatrix} \quad \text{vs.} \quad H_1 : \mathbf{\Omega} = \begin{bmatrix} \sigma_c^2 + \sigma_u^2 & & \sigma_c^2 \\ & \ddots & \\ \sigma_c^2 & & \sigma_c^2 + \sigma_u^2 \end{bmatrix}$$

Comments:

- This is a two-sided alternative.
- Theoretically, $\sigma_c^2 < 0$ is not sensible.
- Empirically, $\hat{\sigma}_c^2 < 0$ may happen, indicating some more general deviation from the implicit null $\mathbf{\Omega} = \sigma_u^2 \mathbf{I}_T$.

A simple test statistic

We can use any lower triangular element of $\mathrm{E}(\mathbf{v}_i'\mathbf{v}_i) = \mathbf{\Omega}$ as the basis for a test statistic because the hypotheses imply

$$H_0 : \mathrm{E}(v_{it}v_{is}) = 0 \quad \text{vs.} \quad H_1 : \mathrm{E}(v_{it}v_{is}) = \sigma_c^2, \qquad \text{for some } s > t.$$

If we were able to observe v_{it}, a simple statistic would be, for some choice of $s > t$,

$$N^{-1/2} \sum_{i=1}^{N} v_{it}v_{is} \xrightarrow{\mathrm{d}} \text{Normal}(0, w^2)$$

where $w^2 = \mathrm{Var}(v_{it}v_{is}) = \mathrm{E}(v_{it}^2 v_{is}^2)$.

Replacing v_{it} by the POLS residuals $\hat{\hat{v}}_{it}$, changes nothing asymptotically under the null:

$$N^{-1/2} \sum_{i=1}^{N} \hat{\hat{v}}_{it}\hat{\hat{v}}_{is} \xrightarrow{\mathrm{d}} \text{Normal}(0, w^2)$$

and we can estimate w^2 by $N^{-1} \sum_{i=1}^{N} \hat{\hat{v}}_{it}^2 \hat{\hat{v}}_{is}^2$.

Wooldridge's test statistic

Instead of using one of the triangular elements (which one?), Wooldridge suggests to use their sum because the hypotheses imply

$$H_0 : \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} \mathrm{E}(v_{it} v_{is}) = 0 \quad \text{vs.} \quad H_1 : \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} \mathrm{E}(v_{it} v_{is}) = \frac{T(T-1)}{2} \sigma_c^2 .$$

Note that the random variable

$$h_i = \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} v_{it} v_{is}$$

has mean zero under the null and is independent and identically distributed. Hence, a CLT applies:

$$N^{-1/2} \sum_{i=1}^{N} h_i \xrightarrow{\mathrm{d}} \mathsf{Normal}(0, \mathrm{Var}(h_i)),$$

$$\mathrm{Var}(h_i) = \mathrm{Var}\left(\textstyle\sum_{t=1}^{T-1} \sum_{s=t+1}^{T} v_{it} v_{is} \right) = \mathrm{E}\left(\textstyle\sum_{t=1}^{T-1} \sum_{s=t+1}^{T} v_{it} v_{is} \right)^2 .$$

No residual effect

Replacing v_{it} by the POLS residuals $\hat{\hat{v}}_{it}$ has no effect asymptotically:

$$N^{-1/2}\sum_{i=1}^{N}\hat{h}_i = N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\hat{\hat{v}}_{it}\hat{\hat{v}}_{is} \xrightarrow{\mathrm{d}} \text{Normal}(0, w^2)$$

where w^2 can be estimated as

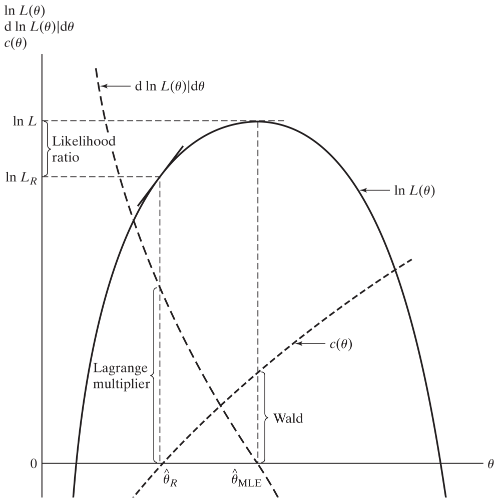
$$\hat{w}^2 = N^{-1}\sum_{i=1}^{N}\hat{h}_i^2 = N^{-1}\sum_{i=1}^{N}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\hat{\hat{v}}_{it}\hat{\hat{v}}_{is}\right)^2 .$$

Dividing $N^{-1/2}\sum_{i=1}^{N}\hat{h}_i$ by $\hat{w}$ yields asymptotic standard normality,

$$q = \frac{\sum_{i=1}^{N}\hat{h}_i}{\sqrt{\sum_{i=1}^{N}\hat{h}_i^2}} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\hat{\hat{v}}_{it}\hat{\hat{v}}_{is}}{\sqrt{\sum_{i=1}^{N}\left(\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\hat{\hat{v}}_{it}\hat{\hat{v}}_{is}\right)^2}} \xrightarrow{\mathrm{d}} \text{Normal}(0, 1).$$

Outline

LM principle



Source: Greene (2012) Econometric Analysis, p. 525.

Lagrange multiplier tests (score tests)

Recall that the Lagrange multiplier (LM) statistic offers a way to test $H_0 : \mathbf{c}(\boldsymbol{\theta}_o) = \mathbf{0}$ by only estimating the model under the null.

(In our case, estimation under the null is just pooled OLS.)

Let $\tilde{\mathbf{s}}_i = \mathbf{s}_i(\tilde{\boldsymbol{\theta}})$ be the score evaluated at the POLS estimate $\tilde{\boldsymbol{\theta}}$. This implies that you

(1) have to set up the likelihood function of the unconstrained model,

(2) compute the partial derivatives of $\ell_i(\boldsymbol{\theta})$ with respect to *each* of the parameters (also with respect to those which might be restricted under H_0), and

(3) evaluate this vector of partial derivatives at the *restricted* estimates.

Hence, you still have to write down the unrestricted model but you only have to estimate the restricted one.

Which takes us to

The **Lagrange multiplier (LM) statistic**

$$LM \equiv N \left(N^{-1} \sum_{i=1}^{N} \tilde{\mathbf{s}}_i \right)' \tilde{\mathbf{A}}^{-1} \left(N^{-1} \sum_{i=1}^{N} \tilde{\mathbf{s}}_i \right)$$

is distributed asymptotically as χ^2_Q under H_0, where Q is the number of restrictions.

Breusch-Pagan LM test

Hypotheses: $H_0 : \sigma_c^2 = 0$ versus $H_1 : \sigma_c^2 \neq 0$

Estimation under the null: restricted ML estimator = POLS estimator

POLS yields regression residuals $\tilde{v}_{it}$ and a ML estimator of σ_u^2 under the null:

$$\tilde{\sigma}_u^2 = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{v}_{it}^2.$$

Test statistic:

$$LM \equiv \frac{NT}{2(T-1)} \left(\frac{\frac{1}{N} \sum_{i=1}^{N} \tilde{\mathbf{v}}_i' \mathbf{J}_T \tilde{\mathbf{v}}_i}{\tilde{\sigma}_u^2} - 1 \right)^2 .$$

Under the null it is asymptotically χ_1^2 distributed.

An alternative view

It can be shown that the LM statistic can also be written as

$$LM = \frac{NT(T-1)}{2}\left(\frac{\ddot{\sigma}_u^2}{\tilde{\sigma}_u^2} - 1\right)^2,$$

where

$$\tilde{\sigma}_u^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{v}_{it}^2$$

is the usual POLS variance estimator and

$$\ddot{\sigma}_u^2 = \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=1}^{T}(\tilde{v}_{it} - \bar{\tilde{v}}_i)^2$$

is a within-type variance estimator (but still based on the POLS residuals).

Under the null, both estimators converge towards σ_u^2 but under the alternative, only $\ddot{\sigma}_u^2$ does. (See tutorial.)

Discussion

- Potential drawback of this test: it is based on the normal distribution.
- Fortunately, Honda (1985, Rev Econ Stud 52(4), 681-690) shows that it is robust to non-normality.
- A drawback is the two-sided alternative $H_1 : \sigma_c^2 \neq 0$ because $\sigma_c^2 < 0$ does not make sense.
- Therefore, Honda derives a uniformly most powerful test for $H_0 : \sigma_c^2 = 0$ versus $H_1 : \sigma_c^2 > 0$.
- The test statistic turns out to be the root of the Breusch-Pagan LM statistic:

$$H \equiv \sqrt{\frac{NT}{2(T-1)}} \left(\frac{\frac{1}{N} \sum_{i=1}^{N} \tilde{\mathbf{v}}_i' \mathbf{J}_T \tilde{\mathbf{v}}_i}{\tilde{\sigma}_u^2} - 1 \right) \xrightarrow{\mathrm{d}} \text{Normal}(0, 1).$$

It rejects for values of H that exceed the, say, 95% critical value of the standard normal distribution.

Implementation in Stata

Use the command `xttest0` after a random effects estimation.

First compute the RE estimator:

```
xtreg y x1 x2 x3, re
```

Then run the Breusch-Pagan LM test:

```
xttest0
```

Example: Effects of job training grants on scrap rates

Example 10.4 taken from Wooldridge's textbook

Stata code:

```
*** load data ***
use "jtrain1.dta", clear

*** set panel ***
xtset fcode year

*** run RE regression with robust s.e.'s, display theta ***
xtreg lscrap d88 d89 union grant grant_1, re vce(robust)
theta

*** run Breusch-Pagan LM test ***
xttest0
```

Stata output

```
Breusch and Pagan Lagrangian multiplier test for random effects

        lscrap[fcode,t] = Xb + u[fcode] + e[fcode,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                  lscrap |   2.209597       1.486471
                       e |   .2477493       .4977442
                       u |    1.93218       1.390029

        Test:   Var(u) = 0
                             chibar2(01) =   124.28
                          Prob > chibar2 =   0.0000
```

Outline

Why test for autocorrelation?

Recall that we can always use variance estimators that are robust. Still:

- Economically, it may be interesting whether a kind of partial adjustment is going on. Recall, e.g., that the transmission of a policy measure may take time if the model has a distributed lag structure. It can be of particular interest to infer this structure.
- Recall that $v_{it} = c_i + u_{it}$ is autocorrelated with $\mathrm{E}(v_{it}v_{is}) = \sigma_c^2$, $t \neq s$, whenever there is an unobserved individual effect c_i. Hence, an autocorrelation test is another way (even if not an efficient one) to test for $\sigma_c^2 > 0$.
- Without autocorrelation of v_{it}, OLS might be the best we can do to estimate a panel equation—the RE estimator is more efficient than OLS only if there is an error components structure (and thus autocorrelation) in v_{it}.
- Also, remember that the FE estimator is efficient if u_{it} is white noise and the FD estimator is efficient if $e_{it} \equiv \Delta u_{it}$ is white noise.

Autocorrelation in the error components model

Consider the error components structure

$$v_{it} = c_i + u_{it}$$

and assume $\mathrm{E}(c_i) = 0$, $\mathrm{E}(u_{it}) = 0$, and $\mathrm{Cov}(c_i, u_{it}) = 0$. Then v_{it} can be autocorrelated for different reasons:

(a) If u_{it} is white noise and $\sigma_c^2 > 0$, we obtain autocovariance

$$\mathrm{E}(v_{it}v_{is}) = \sigma_c^2.$$

(b) If u_{it} is autocorrelated, $\mathrm{E}(u_{it}u_{is}) = \rho_s$ and $\sigma_c^2 = 0$, we obtain autocovariance

$$\mathrm{E}(v_{it}v_{is}) = \mathrm{E}(u_{it}u_{is}) = \rho_s.$$

(c) If u_{it} is autocorrelated, $\mathrm{E}(u_{it}u_{is}) = \rho_s$ and $\sigma_c^2 > 0$, we obtain autocovariance

$$\mathrm{E}(v_{it}v_{is}) = \sigma_c^2 + \rho_s.$$

Testing for autocorrelation after POLS

Recall that POLS is consistent as long as $\mathrm{E}(\mathbf{x}_{it}' u_{it}) = \mathbf{0}$ and $\mathrm{E}(\mathbf{x}_{it}' c_i) = \mathbf{0}$.

However, compared to the RE estimator it is inefficient if $\sigma_c^2 > 0$.

While the Breusch-Pagan test is better suited, it is straightforward to test $H_0 : \mathrm{E}(v_{it} v_{is}) = 0$ against $H_1 : \mathrm{E}(v_{it} v_{is}) \neq 0$.

Just keep in mind that rejection of the null does not necessarily imply that $\sigma_c^2 > 0$. It may also be caused by $\mathrm{E}(u_{it} u_{is}) \neq 0$.

Residual autocorrelation

Here is a very simple way to test for first-order autocorrelation:

$$H_0 : \mathrm{E}(v_{it}v_{it-1}) = 0 \quad \text{against} \quad H_1 : \mathrm{E}(v_{it}v_{it-1}) \neq 0$$

- Estimate model $y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + v_{it}$ by POLS and save the residuals $\hat{v}_{it}$.
- Regress $\hat{v}_{it}$ on $\hat{v}_{it-1}$ by POLS and compute the coefficient t statistic (preferably with a robust variance estimator).
- Reject the null if the absolute t statistic exceeds the critical values of the standard normal distribution.
- Note that the use of a lagged $\hat{v}_{it-1}$ as regressor takes away one observation per individual, thus N observations altogether.
- (In a statistical software, be sure that the correct observations are deleted. In Stata, execute the `xtset` command first.)
- Important: consistency of this test hinges on the strict exogeneity assumption, see Wooldridge (p. 199) for details.

Implementation in Stata

Example: Data set has identifier for each individual denoted `id` and for each time period denoted `year`.

First tell Stata that you have panel data:

```
xtset id year
```

Perform POLS:

```
regress y x1 x2 x3
```

Compute residuals:

```
predict v if e(sample), residuals
```

Regress residuals on own lag (use the "l." operator):

```
regress v l.v, noconstant vce(cluster id)
```

Example: Effects of job training grants on scrap rates

Example 10.5 taken from Wooldridge's textbook

Stata code:

```
*** load data ***
use "jtrain1.dta", clear

*** set panel ***
xtset fcode year

*** run POLS regression ***
regress lscrap d88 d89 grant grant_1, vce(cluster fcode)

*** Compute residuals ***
predict v if e(sample), residuals

*** Regress residuals on own lag ***
regress v l.v, noconstant vce(cluster fcode)
```

Stata output

```
Linear regression                               Number of obs =        108
                                                F(  1,     53) =    180.37
                                                Prob > F       =    0.0000
                                                R-squared      =    0.8413
                                                Root MSE       =    .56436

                           (Std. Err. adjusted for 54 clusters in fcode)
```

v	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
v						
L1.	.8671465	.0645667	13.43	0.000	.737642	.996651

Testing for autocorrelation after FE estimation

Since the within transformation wipes out c_i, this is effectively a test for autocorrelation in the remainder disturbance u_{it}.

A complication arises because the within transformation yields $\ddot{u}_{it} = u_{it} - \bar{u}_i$ which is autocorrelated even if u_{it} is white noise.

In this case, you will be asked to show in the tutorial that

$$\mathrm{E}(\ddot{u}_{it}\ddot{u}_{it-1}) = -\sigma_u^2/T$$

and

$$\mathrm{Corr}(\ddot{u}_{it}, \ddot{u}_{it-1}) = (1-T)^{-1}.$$

Note that this implies that for $T = 2$, $\mathrm{Corr}(\ddot{u}_{i2}, \ddot{u}_{i1}) = -1$. In fact, this is independent of the autocorrelation of u_{it}. Hence, for $T = 2$ there is no sense to test for autocorrelation.

Residual correlation again

For $T > 2$, we can proceed as follows to test the hypotheses

$$H_0 : \mathrm{Corr}(\ddot{u}_{it}, \ddot{u}_{it-1}) = (1-T)^{-1} \quad \text{vs.} \quad H_1 : \mathrm{Corr}(\ddot{u}_{it}, \ddot{u}_{it-1}) \neq (1-T)^{-1}.$$

- Estimate the model $\ddot{y}_{it} = \ddot{\mathbf{x}}_{it}\boldsymbol{\beta} + \ddot{u}_{it}$ by FE and save the residuals $\hat{\ddot{u}}_{it}$.
- Regress $\hat{\ddot{u}}_{it}$ on $\hat{\ddot{u}}_{it-1}$ by POLS and compute the coefficient t statistic (use the robust variance estimator because there is autocorrelation under the null).
- Reject the null if the absolute t statistic exceeds the critical values of the standard normal distribution.
- Note that the use of a lagged $\hat{\ddot{u}}_{it-1}$ as regressor takes away one observation per individual, thus N observations altogether.
- (In a statistical software, be sure that the correct observations are deleted. In Stata, execute the `xtset` command first.)
- Important: consistency of this test hinges on the strict exogeneity assumption, see Wooldridge (p. 311) for details.

Implementation in Stata

Example: Data set has identifier for each individual denoted `id` and for each time period denoted `year`.

First tell Stata that you have panel data:

```
xtset id year
```

Perform FE estimation:

```
xtreg y x1 x2 x3, fe
```

Compute FE residuals:

```
predict u_fe if e(sample), e
```

Regress FE residuals on own lag (use the “l.” operator):

```
regress u_fe l.u_fe, noconstant vce(cluster id)
```

Example: Effects of job training grants on scrap rates

Example 10.5 taken from Wooldridge's textbook

Question: How do job training grants affect scrap rates (recall: $T = 3$)?

Stata code:

```
*** load data and set panel ***
use "jtrain1.dta", clear
xtset fcode year

*** run POLS regression and compute residuals ***
xtreg lscrap d88 d89 grant grant_1, fe vce(robust)
predict u_fe if e(sample), e

*** Regress residuals on own lag with robust s.e.'s ***
regress u_fe l.u_fe, noconstant vce(cluster fcode)

*** test deviation of coefficient from -1/(T-1) = -1/2 ***
```

```
. regress u_fe l.u_fe, noconstant

      Source |       SS           df       MS      Number of obs   =       108
-------------+----------------------------------   F(1, 107)       =      4.72
       Model |  .607552012         1  .607552012   Prob > F        =    0.0320
    Residual |  13.7716104       107  .128706639   R-squared       =    0.0423
-------------+----------------------------------   Adj R-squared   =    0.0333
       Total |  14.3791624       108  .133140393   Root MSE        =    .35876

------------------------------------------------------------------------------
        u_fe |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        u_fe |
         L1. |  -.2058466   .0947442    -2.17   0.032    -.3936659   -.0180274
------------------------------------------------------------------------------

. test l.u_fe=-1/2

 ( 1)  L.u_fe = -.5

       F(  1,   107) =    9.64
            Prob > F =    0.0024
```

Other tests for autocorrelation

Baltagi's textbook (Chapter 5.2.7) describes several test for autocorrelation, most of them based on the LM principle.

A good choice is the locally best invariant test by Baltagi and Wu (1999).

You may also use the Durbin-Watson test for panel data.

Both test are implemented in Stata (use the `xtregar` command with the option `lbi`) but critical values need to be looked up.

Therefore, it should be sufficient to either use autocorrelation-consistent standard errors anyway, or to perform one of the tests described in detail above and use non-robust standard errors in case the null is not rejected.

If one is interested in the "amount" of autocorrelation, one may specify and estimate an appropriate model like AR(1) or MA(1).

Outline

Coming up

IV Estimation