Panel data models

Demidova O.

2016

Panel data models

$$Y_{it} = \alpha_i + X_{it}^{'}\beta + u_{it},$$

$$i = 1,\ldots,N - individual\ number$$

$$t = 1,\ldots,T - time$$

Pooled regression

$$
Y = \begin{pmatrix} \underset{(T\times 1)}{Y_1} \\ \underset{(T\times 1)}{Y_2} \\ \vdots \\ \underset{(T\times 1)}{Y_N} \end{pmatrix}, \qquad X = \begin{pmatrix} \underset{(T\times k)}{X_1} \\ \underset{(T\times k)}{X_2} \\ \vdots \\ \underset{(T\times k)}{X_N} \end{pmatrix}, \qquad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}
$$

$$
Y = X\beta + \varepsilon
$$

OLS, all observations

Fixed effects models

$$Y_{it} = \alpha_i + X_{it}^{'}\beta + u_{it},$$

$$E(\alpha_i) = const\text{ , }\operatorname{var}(\alpha_i) = 0$$

Dummy for each individual

NT – N – k degrees of freedom

$$\hat{\beta}_{LSDV}$$

Fixed effects model vs Pooled regression

$$H_0 : \alpha_1 = ... = \alpha_N,$$

$$H_1 : \exists\, \alpha_i \neq \alpha_j$$

$$F = \frac{(RSS_{pooled} - RSS_{FE})/(N-1)}{RSS_{FE}/(NT - N - k)}$$

Between regression

$$Y_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} Y_{it}$$

$$Y_{i\cdot} = \alpha_1 + X_{i\cdot}'\beta + \varepsilon_{i\cdot},$$

$$\hat{\beta}_B, \quad RSS_B$$

OLS

Within regression

$$Y_{it} - Y_{i\cdot} = (X'_{it} - X'_{i\cdot})\beta + \varepsilon_{it} - \varepsilon_{i\cdot},$$

$$\hat{\beta}_W = \hat{\beta}_{FE}, \quad RSS_W$$

Random effects models

$$Y_{it} = X_{it}^{'}\beta + const + \alpha_i + \varepsilon_{it},$$

$$E(\alpha_i) = 0, \ \mathrm{var}(\alpha_i) = \sigma_\alpha^2$$

$$\Sigma = \begin{pmatrix} \sigma_\alpha^2 + \sigma_\varepsilon^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \vdots & \ddots & & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 + \sigma_\varepsilon^2 \end{pmatrix}$$

Random effects models

$$Y_{it} = X_{it}^{'}\beta + const + \alpha_i + \varepsilon_{it},$$

$$E(\alpha_i) = 0, \operatorname{var}(\alpha_i) = \sigma_u^2$$

$$\Omega = \begin{pmatrix} \Sigma & & \\ & \ddots & \\ & & \Sigma \end{pmatrix}$$

$$\hat{\beta}_{RE} = (X'\Omega^{-1}X)^{-1}X'\Omega Y$$

Random effects estimates are more effective than fixed effects models. But α_i *and* X may correlate.

Random effects model vs Pooled regression

$$H_0 : \sigma_u^2 \equiv 0,$$

$$H_1 : \sigma_u^2 \neq 0$$

$$Breusch - Pagan \quad test$$

$$F = \frac{\hat{\sigma}_B^2}{\hat{\sigma}_W^2} \sim F(N - k, NT - N - k)$$

$$or \quad LM$$

Random effects model vs Pooled regression

$$LM = \frac{NT}{2(T-1)}\left[\frac{T^2\sum_{i=1}^{N} e_{i.}^2}{\sum_{i=1}^{N}\sum_{t=1}^{T} e_{it}^2} - 1\right]^2,$$

$$e_{it} - residuals \quad of \quad pooled \quad regression$$

Fixed effects model vs Random effects model

$$H_0 : RE \Leftrightarrow corr(\alpha_{i,} X_{it}) = 0$$

$$\Rightarrow RE\ and\ FE \quad are \quad both \quad consistent.$$

$$their \quad difference \quad is \quad small.$$

$$H_1 : FE \Leftrightarrow corr(\alpha_{i,} X_{it}) \neq 0$$

$$FE \quad is \quad consistent,$$

$$RE \quad is \quad inconsistent.$$

$$Difference\ of\ \ RE\ and\ \ FE\ is\ l\arg e.$$

Hausman test

$$m = \hat{q}'\,\mathrm{var}^{-1}(\hat{q})\,\hat{q} \sim \chi_k^2,$$

$$\hat{q} = \hat{\beta}_{FE} - \hat{\beta}_{RE}$$

Example

The data are taken from the National Longitudinal Survey (NLS Youth Sample) and contain observations on 545 males for the years 1980-1987.

Variables:

NR	**Observations number**
YEAR	**Year of observation**
School	**Years of schooling**
Exper	**Age-6-School**
Exper2	**Experience Squared**
LogExper	**Log(1+Experience)**
Union	**Wage set by collective bargaining**
Mar	**Married**
Black	**Black**
Hisp	**Hispanic**
Health	**Has health disability**
Rural	**Lives in rural area**
NE	**Lives in North East**
NC	**Lives in Northern Central**
S	**Lives in south**
Wage	**Log of hourly wage**

Example

Industry Dummies:

AG	Agricultural
MIN	Mining
CON	Construction
TRAD	Trade
TRA	Transportation
FIN	Finance
BUS	Business & Repair Service
PER	Personal Service
ENT	Entertainment
MAN	Manufacturing
PRO	Professional & Related Service
PUB	Public Administration

Occupational Dummies:

OCC1	Professional, Technical and kindred
OCC2	Managers, Officials and Proprietors
OCC3	Sales Workers
OCC4	Clerical and kindred
OCC5	Craftsmen, Foremen and kindred
OCC6	Operatives and kindred
OCC7	Laborers and farmers
OCC8	Farm Laborers and Foreman
OCC9	Service Workers

Example

```
. xtreg  WAGE SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB, fe
note: SCHOOL omitted because of collinearity
note: BLACK omitted because of collinearity
note: HISP omitted because of collinearity

Fixed-effects (within) regression               Number of obs      =      4360
Group variable: NR                              Number of groups   =       545

R-sq:  within  = 0.1782                         Obs per group: min =         8
       between = 0.0006                                        avg =       8.0
       overall = 0.0642                                        max =         8

                                                F(5,3810)          =    165.26
corr(u_i, Xb)  = -0.1130                        Prob > F           =    0.0000

------------------------------------------------------------------------------
        WAGE |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      SCHOOL |  (omitted)
       EXPER |    .116457   .0084309    13.81   0.000     .0999275    .1329865
      EXPER2 |  -.0042886   .0006054    -7.08   0.000    -.0054756   -.0031015
       UNION |    .081203   .0193159     4.20   0.000     .0433325    .1190736
         MAR |   .0451061   .0183114     2.46   0.014      .009205    .0810072
       BLACK |  (omitted)
        HISP |  (omitted)
         PUB |   .0349267   .0386082     0.90   0.366     -.040768    .1106214
       _cons |   1.065698   .0266766    39.95   0.000     1.013396       1.118
-------------+----------------------------------------------------------------
     sigma_u |  .39989822
     sigma_e |  .35126372
         rho |  .56447541   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(544, 3810) =     7.98           Prob > F = 0.0000
```

Example

```
. xtreg  WAGE SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB, re

Random-effects GLS regression                   Number of obs      =      4360
Group variable: NR                              Number of groups   =       545

R-sq:  within  = 0.1776                         Obs per group: min =         8
       between = 0.1835                                        avg =       8.0
       overall = 0.1808                                        max =         8

Random effects u_i ~ Gaussian                   Wald chi2(8)       =    944.56
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
        WAGE |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      SCHOOL |   .1010237   .0089219    11.32   0.000     .0835372    .1185103
       EXPER |   .1117851   .0082709    13.52   0.000     .0955744    .1279959
      EXPER2 |  -.0040575    .000592    -6.85   0.000    -.0052177   -.0028972
       UNION |   .1064134   .0178669     5.96   0.000     .0713949    .1414319
         MAR |   .0625465   .0167762     3.73   0.000     .0296658    .0954272
       BLACK |  -.1440026   .0476439    -3.02   0.003     -.237383   -.0506223
        HISP |   .0197269   .0426303     0.46   0.644    -.0638269    .1032807
         PUB |   .0301555   .0364671     0.83   0.408    -.0413187    .1016296
       _cons |  -.1043113    .110834    -0.94   0.347    -.3215421    .1129194
-------------+----------------------------------------------------------------
     sigma_u |  .32482045
     sigma_e |  .35126372
         rho |  .46094736   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Example

```
. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

        WAGE[NR,t] = Xb + u[NR] + e[NR,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                    WAGE |   .2836728       .5326094
                       e |   .1233862       .3512637
                       u |   .1055083       .3248205

        Test:   Var(u) = 0
                             chi2(1) =  3217.14
                         Prob > chi2 =     0.0000
```

Example

```
.
.       hausman fixed

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed          .          Difference          S.E.
-------------+----------------------------------------------------------------
       EXPER |    .116457     .1117851        .0046718        .0016345
      EXPER2 |  -.0042886    -.0040575       -.0002311        .0001269
       UNION |    .081203     .1064134       -.0252104        .0073402
         MAR |   .0451061     .0625465       -.0174403        .0073395
         PUB |   .0349267     .0301555        .0047713        .0126785
-------------+----------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(5) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       31.75
                Prob>chi2 =      0.0000
```