UNIVERSITY OF ESSEX
Session: 2011/12

DEPARTMENT OF ECONOMICS
João Santos Silva

EC501 Econometric Methods and Applications

Problem Set 10: Sketch Solutions

Panel Data Models

1. (a) We begin by stacking the observations over t for each i:

$$y_i = X_i\beta + 1_T\alpha_i + \epsilon_i, \quad i = 1, \ldots, n,$$

where y_i is $T \times 1$, X_i is $T \times K$, α_i is a scalar and 1_T denotes a $T \times 1$ vector of ones. We then stack these equations to obtain

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \beta + \begin{pmatrix} 1_T & 0 & \ldots & 0 \\ 0 & 1_T & \ldots & 0 \\ & & \vdots & \\ 0 & 0 & \ldots & 1_T \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$y = X\beta + D\alpha + \epsilon,$$

where y is $nT \times 1$, X is $nT \times K$, $D = I_n \otimes 1_T$ is $nT \times n$, α is $n \times 1$ and ϵ is $NT \times 1$.

(b) Let $M_D = I - D(D'D)^{-1}D'$ and note that $M_D D = 0$. Then the OLS estimator of β can be obtained by pre-multiplying the model by M_D:

$$M_D y = M_D X\beta + M_D\epsilon \quad \text{(because } M_D D = 0\text{)}$$

$$\Rightarrow \quad y_* = X_*\beta + \epsilon_*$$

$$\Rightarrow \quad b = (X_*'X_*)^{-1}X_*'y_* = (X'M_D X)^{-1}X'M_D y.$$

The estimator of α can be obtained by noting that the regression residuals, e, are orthogonal to both X and D (i.e. $X'e = 0$ and $D'e = 0$). Pre-multiplying the residuals $e = y - Xb - Da$ by D' results in

$$D'e = D'y - D'Xb - D'Da = 0 \quad \Rightarrow \quad D'Da = D'(y - Xb)$$

$$\Rightarrow \quad a = (D'D)^{-1}D'(y - Xb).$$

(c) In the expression for b make the substitution $y = X\beta + D\alpha + \epsilon$ to obtain

$$b = (X'M_D X)^{-1}X'M_D(X\beta + D\alpha + \epsilon) = \beta + (X'M_D X)^{-1}X'M_D\epsilon,$$

the term involving α cancelling because $M_D D = 0$. It follows that

$$\begin{aligned} E[(b-\beta)(b-\beta)'] &= (X'M_D X)^{-1}X'M_D E(\epsilon\epsilon')M_D X(X'M_D X)^{-1} \\ &= \sigma^2(X'M_D X)^{-1} \end{aligned}$$

because $E(\epsilon\epsilon'|x_{it}) = \sigma^2 I_{nT}$ and $M_D^2 = M_D$.

(a) The OLS estimates are as follows:

Source	SS	df	MS
Model	112.705452	3	37.5684839
Residual	1.33544153	86	.01552839
Total	114.040893	89	1.28135835

Number of obs	90
F(3, 86)	2419.34
Prob > F	0.0000
R-squared	0.9883
Adj R-squared	0.9879
Root MSE	.12461

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x1	.8827385	.0132545	66.60	0.000	.8563895	.9090876
x2	.453977	.0203042	22.36	0.000	.4136136	.4943404
lf	-1.62751	.345302	-4.71	0.000	-2.313948	-.9410727
_cons	9.516923	.2292445	41.51	0.000	9.0612	9.972645

(b) The OLS estimates of the fixed effects model are:

Source	SS	df	MS
Model	113.74827	8	14.2185338
Residual	.292622872	81	.003612628
Total	114.040893	89	1.28135835

Number of obs	90
F(8, 81)	3935.79
Prob > F	0.0000
R-squared	0.9974
Adj R-squared	0.9972
Root MSE	.06011

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x1	.9192846	.0298901	30.76	0.000	.8598126	.9787565
x2	.4174918	.0151991	27.47	0.000	.3872503	.4477333
lf	-1.070396	.20169	-5.31	0.000	-1.471696	-.6690963
d1	-.0870617	.0841995	-1.03	0.304	-.2545924	.080469
d2	-.1282976	.0757281	-1.69	0.094	-.2789728	.0223776
d3	-.2959828	.0500231	-5.92	0.000	-.395513	-.1964526
d4	.097494	.0330093	2.95	0.004	.0318159	.1631721
d5	-.063007	.0238919	-2.64	0.010	-.1105443	-.0154697
_cons	9.793004	.2636622	37.14	0.000	9.268399	10.31761

To test the null that all six $\alpha_i = \alpha$ (i.e. that they are all equal to the same constant value) we have 5 restrictions in all. There are $K = 3$ regressors, and based on the above estimations we find

$$F = \frac{S_{Pooled} - S_{LSDV}}{S_{LSDV}} \cdot \frac{nT - n - K}{n - 1} = \frac{1.33544153 - 0.292622872}{0.292622872} \cdot \frac{90 - 6 - 3}{5} = 57.732$$

which is distributed as $F_{5,81}$ under the null. The 5% critical value is approximately 2.34 and hence we reject the null that the α_i are constant across i at the 5% level. The same results can be obtained with: `xtreg y x* lf, fe`.

(c) The estimates of the random effects model and the associated LM test are as follows:

```
Random-effects GLS regression                   Number of obs      =        90
Group variable (i): i                           Number of groups   =         6

R-sq:  within  = 0.9925                         Obs per group: min =        15
       between = 0.9856                                        avg =      15.0
       overall = 0.9876                                        max =        15

Random effects u_i ~Gaussian                   Wald chi2(3)       =  11091.33
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .9066805    .025625    35.38   0.000     .8564565    .9569045
          x2 |   .4227784   .0140248    30.15   0.000     .3952904    .4502665
          lf |  -1.064499   .2000703    -5.32   0.000    -1.456629    -.672368
       _cons |   9.627909    .210164    45.81   0.000     9.215995    10.03982
-------------+----------------------------------------------------------------
     sigma_u |  .12488859
     sigma_e |  .06010514
         rho |  .81193816   (fraction of variance due to u_i)
------------------------------------------------------------------------------

Breusch and Pagan Lagrangian multiplier test for random effects:

        y[i,t] = Xb + u[i] + e[i,t]

        Estimated results:
                        |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                       y |   1.281358       1.131971
                       e |   .0036126       .0601051
                       u |   .0155972       .1248886

        Test:   Var(u) = 0
                             chi2(1) =   334.85
                         Prob > chi2 =     0.0000
```

The null hypothesis that $\sigma_u^2 = 0$ is strongly rejected. The additional assumption that is required for this estimator to be consistent is: $E(u_i|x_{it}) = 0$.