

Economics of Development: Problem Set 1

S M Sajid Al Sanai

April 20, 2019

Contents

1	Question II	2
1.1	Source	2
1.2	Output	4
1.3	Graphs	11

1 Question II

1.1 Source

```
* Development Economics: Problem Set 1
* S M Sajid Al Sanai

* Import dataset
use burkina731.dta, clear

* Identifier variables do not use {0} as indexing value
//tab village //v [1-6]
//tab hhn //h [1-32]
//tab year //t [1-3]
//tab fcrop //c [1-44] discount.
//tab plotnum //i [1-999] discount.

* Generate fixed effects
** Fixed effects are
gen fe_vtc = (village * 1000) + (fcrop * 10) + year
label var fe_vtc "vtc Fixed Effects"
gen fe_vhtc = (fe_vtc * 100) + hhn
label var fe_vhtc "vhtc Fixed Effects"

* Dependent variable is log of yield (vhtci)
** {yield = tot_value / tot_area}
** => ln_yield = ln( tot_value / tot_area )
** => ln_yield = ln( tot_value ) - ln( tot_area )
rename lnvalue ln_value
rename lnarea ln_area
gen ln_yield = ln_value - ln_area

* Fixed Effects Regression
** {xtreg, fe i(.)} Command format
** {topo* soil* loc* ln_area} Covariates matrix
** {totarea lntarea ln_hhsize} Exclude
xtreg ln_yield topo* soil* loc* ln_area, fe i(fe_vtc)
predict residuals_v, e
label var residuals_v "v Residuals"

xtreg ln_yield topo* soil* loc* ln_area, fe i(fe_vhtc)
predict residuals_h, e
label var residuals_h "h Residuals"

* Estimation of Density
kdensity residuals_v, normal gen(est_pts_vtc density_vtc)
label var density_vtc "vtc Density"
label var est_pts_vtc "vtc Pred. Dev."
graph export graph1.png, replace

kdensity residuals_h, at(est_pts_vtc) gen(est_pts_vhtc density_vhtc)
label var density_vhtc "vhtc Density"
label var est_pts_vhtc "vhtc Pred. Dev."
graph export graph2.png, replace

graph twoway connected density_vtc density_vhtc est_pts_vtc, l("Kernel Density Estimate")
graph export graph3.png, replace

* Modifying dataset structure and cleaning up
```

```

** Append residuals for village level and household level regressions
** into a single column by saving with one constructed index and then
** importing from temporary dataset.
keep residuals_v residuals_h
* Create the index
gen is_village = 1
* Save temporary dataset
save temp_burkina73, replace
* Change the index
replace is_village = 0
* Append using old index
append using temp_burkina73
* Verify
tab is_village
* Construct single variable residuals using index
gen residuals = .
replace residuals = residuals_v if is_village == 1
replace residuals = residuals_h if is_village == 0
keep residuals is_village
* Save temporary dataset
save temp_burkina73, replace

* Testing for equality of distribution
ksmirnov residuals, by(is_village)

* Reuse dataset prior to temporary modifications
use burkina731.dta, clear

* Generate fixed effects
** Fixed effects are
gen fe_vtc = (village * 1000) + (fcrop * 10) + year
label var fe_vtc "vtc Fixed Effects"
gen fe_vhtc = (fe_vtc * 100) + hhn
label var fe_vhtc "vhtc Fixed Effects"

* Dependent variable is log of yield (vhtci)
** {yield = tot_value / tot_area}
** => ln_yield = ln( tot_value / tot_area )
** => ln_yield = ln( tot_value ) - ln( tot_area )
rename lnvalue ln_value
rename lnarea ln_area
gen ln_yield = ln_value - ln_area

* Fixed Effects Regression w/ Additional Household Characteristics
** {xtreg, fe i(.)} Command format
** {topo* soil* loc* ln_area} Covariates matrix
** {Intarea ln_hhsize} Additional Characteristics
** {totarea} Exclude
xtreg ln_yield topo* soil* loc* ln_area ln_hhsize, fe i(fe_vtc)
predict residuals_v_hs, e
label var residuals_v_hs "v Residuals w/ Log of Household Size"

xtreg ln_yield topo* soil* loc* ln_area Intarea, fe i(fe_vtc)
predict residuals_v_hta, e
label var residuals_v_hta "v Residuals w/ Log of Total Area Cultivated by Household Head"

```

1.2 Output

```

    name: <unnamed>
    log: C:\Users\Dell\Documents\Graduate – Economics\Development Economics\PS1_SMSajidAISana
    log type: smcl
    opened on: 23 Apr 2019, 14:28:00

. do "C:\Users\Dell\AppData\Local\Temp\STD01000000.tmp"

. * Development Economics: Problem Set 1
. * S M Sajid Al Sanai
.
. * Import dataset
. use burkina731.dta, clear

.
. * Identifier variables do not use {0} as indexing value
. //tab village //v      [1–6]
. //tab hhn //h      [1–32]
. //tab year //t      [1–3]
. //tab fcrop //c      [1–44]  discount.
. //tab plotnum //i      [1–999]  discount.
.
. * Generate fixed effects
. ** Fixed effects are
. gen fe_vtc = (village * 1000) + (fcrop * 10) + year

. label var fe_vtc "vtc Fixed Effects"

. gen fe_vhtc = (fe_vtc * 100) + hhn

. label var fe_vhtc "vhtc Fixed Effects"

.
. * Dependent variable is log of yield (vhtci)
. ** {yield = tot_value / tot_area}
. ** => ln_yield = ln( tot_value / tot_area )
. ** => ln_yield = ln( tot_value ) – ln( tot_area )
. rename lnvalue ln_value

. rename lnarea ln_area

. gen ln_yield = ln_value – ln_area

.
. * Fixed Effects Regression
. ** {xtreg, fe i(.)}
. ** {topo* soil* loc* ln_area} Covariates matrix
. ** {totarea lnarea ln_hhsize} Exclude
. xtreg ln_yield topo* soil* loc* ln_area, fe i(fe_vtc)
warning: existing panel variable is not fe_vtc

Fixed-effects (within) regression      Number of obs      =      2576
Group variable: fe_vtc                Number of groups    =      162

R-sq:  within  = 0.0422                Obs per group: min =          1
      between  = 0.1326                  avg   =      15.9
      overall   = 0.1224                  max   =      56

```

corr(u_i, Xb) = 0.1829 F(24,2390) = 4.38
 Prob > F = 0.0000

ln_yield	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
topo1	-.4173048	.1552245	-2.69	0.007	-.7216934	-.1129162
topo2	-.2770841	.1379051	-2.01	0.045	-.54751	-.0066581
topo3	-.2596383	.1350329	-1.92	0.055	-.5244321	.0051555
topo4	-.2068287	.1326425	-1.56	0.119	-.466935	.0532775
soil7	-.1693681	.1374232	-1.23	0.218	-.438849	.1001129
soil21	.2189651	.1928816	1.14	0.256	-.1592675	.5971976
soil31	.0639357	.1454792	0.44	0.660	-.2213427	.3492141
soil32	-.0029926	.1493357	-0.02	0.984	-.2958336	.2898484
soil33	.2108274	.1877727	1.12	0.262	-.1573867	.5790415
soil37	.1287846	.1982255	0.65	0.516	-.259927	.5174963
soil35	.1403218	.2155674	0.65	0.515	-.2823967	.5630402
soil45	.0298251	.1640325	0.18	0.856	-.2918356	.3514858
soil51	.1275817	.0972852	1.31	0.190	-.0631904	.3183539
soil1	.1579305	.4604687	0.34	0.732	-.7450289	1.06089
soil3	-.1055179	.1332205	-0.79	0.428	-.3667576	.1557217
soil11	-.3610024	.2326668	-1.55	0.121	-.8172519	.0952472
soil12	-.0667989	.2188587	-0.31	0.760	-.4959714	.3623735
soil13	.1708714	.2667383	0.64	0.522	-.3521909	.6939337
soil34	-.1531976	.2647133	-0.58	0.563	-.6722891	.3658938
soil46	-.418554	.2377307	-1.76	0.078	-.8847337	.0476257
soil53	.3928609	.2157001	1.82	0.069	-.0301177	.8158396
loc1	.2106108	.0753538	2.79	0.005	.0628452	.3583763
loc2	.118715	.0584738	2.03	0.042	.0040504	.2333796
ln_area	-.1610271	.0247182	-6.51	0.000	-.2094984	-.1125558
_cons	3.415241	.1392904	24.52	0.000	3.142099	3.688384
sigma_u	1.1032508					
sigma_e	1.091883					
rho	.50517851	(fraction of variance due to u_i)				

F test that all u_i=0: F(161, 2390) = 7.46 Prob > F = 0.0000

. predict residuals_v , e

. label var residuals_v "v Residuals"

. xtreg ln_yield topo* soil* loc* ln_area , fe i(fe_vhtc)
 warning: existing panel variable is not fe_vhtc

Fixed-effects (within) regression Number of obs = 2576
 Group variable: fe_vhtc Number of groups = 1846

R-sq: within = 0.0761 Obs per group: min = 1
 between = 0.0529 avg = 1.4
 overall = 0.0584 max = 6

corr(u_i, Xb) = -0.0659 F(24,706) = 2.42
 Prob > F = 0.0002

ln_yield	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
----------	-------	-----------	---	------	----------------------	--

topo1	-.4709716	.267833	-1.76	0.079	-.9968161	.054873
topo2	-.146441	.2340169	-0.63	0.532	-.6058934	.3130113
topo3	-.1741458	.2262705	-0.77	0.442	-.6183895	.2700979
topo4	-.1131443	.2229497	-0.51	0.612	-.5508682	.3245796
soil7	-.2937165	.2778046	-1.06	0.291	-.8391386	.2517057
soil21	.2639428	.3082066	0.86	0.392	-.3411685	.8690541
soil31	-.3307579	.2284215	-1.45	0.148	-.7792246	.1177089
soil32	-.3485796	.2463119	-1.42	0.157	-.8321712	.135012
soil33	-.1971536	.3566942	-0.55	0.581	-.8974619	.5031547
soil37	.0668419	.3283176	0.20	0.839	-.5777538	.7114377
soil35	-.1496067	.4514536	-0.33	0.740	-1.035959	.7367456
soil45	-.1476147	.2500988	-0.59	0.555	-.6386411	.3434117
soil51	.1632733	.2259021	0.72	0.470	-.2802471	.6067937
soil1	1.022489	.7351434	1.39	0.165	-.4208398	2.465818
soil3	.2367755	.2669585	0.89	0.375	-.287352	.760903
soil11	-.6923676	.4213079	-1.64	0.101	-1.519534	.1347988
soil12	.424006	.3692	1.15	0.251	-.3008553	1.148867
soil13	1.082068	.6215119	1.74	0.082	-.1381647	2.302301
soil34	-.5078006	.4040409	-1.26	0.209	-1.301066	.2854649
soil46	-.8469797	.3576017	-2.37	0.018	-1.54907	-.1448896
soil53	.3945319	.4681921	0.84	0.400	-.5246835	1.313747
loc1	.1868941	.1147844	1.63	0.104	-.0384656	.4122537
loc2	.0960644	.0976	0.98	0.325	-.0955565	.2876853
ln_area	-.1537155	.0403257	-3.81	0.000	-.2328881	-.0745428
_cons	3.417972	.228832	14.94	0.000	2.968699	3.867244
<hr/>						
sigma_u	1.3621766					
sigma_e	.95713414					
rho	.66947051	(fraction of variance due to u_i)				

F test that all u_i=0: F(1845, 706) = 2.15 Prob > F = 0.0000

```
. predict residuals_h , e

. label var residuals_h "h Residuals"

.

. * Estimation of Density
. kdensity residuals_v , normal gen(est_pts_vtc density_vtc)

. label var density_vtc "vtc Density"

. label var est_pts_vtc "vtc Pred. Dev."

. graph export graph1.png, replace
(file graph1.png written in PNG format)

.

. kdensity residuals_h , at(est_pts_vtc) gen(est_pts_vhtc density_vhtc)

. label var density_vhtc "vhtc Density"

. label var est_pts_vhtc "vhtc Pred. Dev."

. graph export graph2.png, replace
(file graph2.png written in PNG format)

.

. graph twoway connected density_vtc density_vhtc est_pts_vtc , l("Kernel Density Estimate")
```

```

. graph export graph3.png, replace
(file graph3.png written in PNG format)

.
. * Modifying dataset structure and cleaning up
. ** Append residuals for village level and household level regressions
. ** into a single column by saving with one constructed index and then
. ** importing from temporary dataset.
. keep residuals_v residuals_h

. * Create the indexx
. gen is_village = 1

. * Save temporary dataset
. save temp_burkina73, replace
file temp_burkina73.dta saved

. * Change the index
. replace is_village = 0
(2576 real changes made)

. * Append using old index
. append using temp_burkina73

. * Verify
. tab is_village

```

is_village	Freq.	Percent	Cum.
0	2,576	50.00	50.00
1	2,576	50.00	100.00
Total	5,152	100.00	

```

. * Construct single variable residuals using index
. gen residuals = .
(5152 missing values generated)

. replace residuals = residuals_v if is_village == 1
(2576 real changes made)

. replace residuals = residuals_h if is_village == 0
(2576 real changes made)

. keep residuals is_village

. * Save temporary dataset
. save temp_burkina73, replace
file temp_burkina73.dta saved

.
. * Testing for equality of distribution
. ksmirnov residuals, by(is_village)

```

Two-sample Kolmogorov–Smirnov test for equality of distribution functions

Smaller group	D	P-value	Corrected
---------------	---	---------	-----------

```

0:                0.3090    0.000
1:               -0.1836    0.000
Combined K-S:      0.3090    0.000    0.000

```

Note: ties exist in combined dataset;
there are 3847 unique values out of 5152 observations.

```

.
. * Reuse dataset prior to temporary modifications
. use burkina731.dta, clear

.
. * Generate fixed effects
. ** Fixed effects are
. gen fe_vtc = (village * 1000) + (fcrop * 10) + year

. label var fe_vtc "vtc Fixed Effects"

. gen fe_vhtc = (fe_vtc * 100) + hhn

. label var fe_vhtc "vhtc Fixed Effects"

.
. * Dependent variable is log of yield (vhtci)
. ** {yield = tot_value / tot_area}
. ** => ln_yield = ln( tot_value / tot_area )
. ** => ln_yield = ln( tot_value ) - ln( tot_area )
. rename lnvalue ln_value

. rename lnarea ln_area

. gen ln_yield = ln_value - ln_area

.
. * Fixed Effects Regression w/ Additional Household Characteristics
. ** {xtreg, fe i(.)} Command format
. ** {topo* soil* loc* ln_area} Covariates matrix
. ** {lntarea ln_hhsize} Additional Characteristics
. ** {totarea} Exclude
. xtreg ln_yield topo* soil* loc* ln_area ln_hhsize, fe i(fe_vtc)
warning: existing panel variable is not fe_vtc

```

```

Fixed-effects (within) regression              Number of obs   =      2270
Group variable: fe_vtc                        Number of groups =      152

R-sq:  within = 0.0560                        Obs per group:  min =         1
        between = 0.1603                        avg   =      14.9
        overall  = 0.1255                        max   =         56

                                                F(25,2093)      =       4.97
corr(u_i, Xb) = 0.1226                        Prob > F         =      0.0000

```

ln_yield	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
topo1	-.4751642	.1662653	-2.86	0.004	-.8012268	-.1491017
topo2	-.3160177	.1499737	-2.11	0.035	-.6101309	-.0219046
topo3	-.2824825	.1473026	-1.92	0.055	-.5713574	.0063923
topo4	-.233111	.1454243	-1.60	0.109	-.5183023	.0520804

soil37	.1431607	.1981965	0.72	0.470	-.2454944	.5318157
soil35	.1129503	.215735	0.52	0.601	-.310097	.5359977
soil45	.030253	.1639322	0.18	0.854	-.2912112	.3517172
soil51	.1071003	.0976226	1.10	0.273	-.0843335	.2985341
soil1	.1405978	.460244	0.31	0.760	-.7619214	1.043117
soil3	-.1157721	.1332917	-0.87	0.385	-.3771515	.1456073
soil11	-.338055	.2327611	-1.45	0.147	-.7944899	.1183799
soil12	-.0647688	.2187546	-0.30	0.767	-.4937374	.3641999
soil13	.1637517	.2665937	0.61	0.539	-.3590274	.6865307
soil34	-.1508433	.2645525	-0.57	0.569	-.6696197	.3679331
soil46	-.4207864	.2375957	-1.77	0.077	-.8867016	.0451288
soil53	.3912891	.2155743	1.82	0.070	-.0314432	.8140213
loc1	.2016891	.0755108	2.67	0.008	.0536156	.3497627
loc2	.108903	.0585879	1.86	0.063	-.0059853	.2237913
ln_area	-.1661606	.0248036	-6.70	0.000	-.2147993	-.1175219
lntarea	.0654094	.0273768	2.39	0.017	.0117246	.1190942
_cons	3.341878	.1426763	23.42	0.000	3.062096	3.621661
<hr/>						
sigma_u	1.1002254					
sigma_e	1.0912119					
rho	.50411295	(fraction of variance due to u_i)				

F test that all u_i=0: F(161, 2387) = 7.46 Prob > F = 0.0000

. predict residuals_v_hta , e
(2 missing values generated)

. label var residuals_v_hta "v Residuals w/ Log of Total Area Cultivated by Household Head"

end of do-file

. log close
name: <unnamed>
log: C:\Users\Dell\Documents\Graduate – Economics\Development Economics\PS1_SMSajidAISana
log type: smcl
closed on: 23 Apr 2019, 14:28:27

1.3 Graphs





