In [1]:
```
import numpy as np
import pandas as pd
from scipy import stats
# set up panel data
pdata = pd.read_csv("Bank-5.csv", index_col = ["ID", "Year"],
                    parse_dates = True)
pdata
# alternatively
# data = pd.read_csv("http://web.pdx.edu/~crkl/ceR/data/airline.txt", sep='\s+',
# Set data as panel data
# pdata = data.set_index(['I', 'T'], inplace=True)
```

Out[1]:

		SL	Market Capitalization	Market Cap Growth	de	df	roe	roa	roic	ta	
ID	**Year**										
	2004-01-01	1	2,889,069,788	0.30743	1.94645	8.43194	0.135	0.013	0.017	0.062697	0.050
	2005-01-01	1	3,023,672,153	0.04659	1.74653	9.09780	0.135	0.013	0.018	0.037701	0.038
BOKF	**2006-01-01**	1	3,676,927,974	0.21605	1.88263	14.24817	0.130	0.012	0.017	0.026666	0.032
	2007-01-01	1	3,470,470,088	-0.05615	2.40312	16.25542	0.118	0.011	0.015	0.026347	0.013
	2008-01-01	1	2,724,327,015	-0.21500	2.67886	40.67754	0.079	0.007	0.008	0.920015	0.067
...	**...**	...	...	...	...	...	...	...	...	...	
	2017-01-01	20	2,535,730,288	-0.02221	0.06899	1.00060	0.077	0.011	0.018	-0.003037	0.067
	2018-01-01	20	2,658,524,655	0.04843	0.24394	2.05074	0.087	0.014	0.018	-0.085148	0.061
COLB	**2019-01-01**	20	2,934,863,006	0.10394	0.48759	5.13913	0.091	0.014	0.018	0.007901	0.049
	2020-01-01	20	2,570,642,368	-0.12410	0.04957	0.62387	0.068	0.010	0.013	0.575604	0.143
	2021-01-01	20	2,569,284,274	-0.00053	0.04391	0.46435	0.085	0.011	0.015	0.175592	0.071

360 rows × 18 columns

In [2]:
```
pdata.describe()
```

Out[2]:

	SL	Market Cap Growth	de	df	roe	roa	roic	
count	360.000000	360.000000	360.000000	360.000000	360.000000	360.000000	360.000000	360.0
mean	10.500000	0.144246	0.897066	10.915104	0.079881	0.008744	0.012639	0.1
std	5.774307	0.410973	0.680102	52.405240	0.084881	0.008488	0.009891	0.2
min	1.000000	-0.656710	0.026000	-109.443000	-0.692000	-0.069000	-0.060000	-0.0
25%	5.750000	-0.081395	0.376320	2.167350	0.069750	0.008000	0.011000	-0.0
50%	10.500000	0.093700	0.732990	5.801075	0.091000	0.010000	0.015000	0.0
75%	15.250000	0.299200	1.242723	10.150288	0.114250	0.012000	0.017000	0.1
max	20.000000	3.729530	3.601420	678.411760	0.340000	0.036000	0.029000	0.9

In [3]:
```
df = pdata
# Descriptive statistics
df.describe()
```

Out[3]:

	SL	Market Cap Growth	de	df	roe	roa	roic	
count	360.000000	360.000000	360.000000	360.000000	360.000000	360.000000	360.000000	360.0
mean	10.500000	0.144246	0.897066	10.915104	0.079881	0.008744	0.012639	0.1
std	5.774307	0.410973	0.680102	52.405240	0.084881	0.008488	0.009891	0.2
min	1.000000	-0.656710	0.026000	-109.443000	-0.692000	-0.069000	-0.060000	-0.0
25%	5.750000	-0.081395	0.376320	2.167350	0.069750	0.008000	0.011000	-0.0
50%	10.500000	0.093700	0.732990	5.801075	0.091000	0.010000	0.015000	0.0
75%	15.250000	0.299200	1.242723	10.150288	0.114250	0.012000	0.017000	0.1
max	20.000000	3.729530	3.601420	678.411760	0.340000	0.036000	0.029000	0.9

In [7]:
```
# Pooled OLS estimator
from linearmodels import PooledOLS
pooled = PooledOLS.from_formula('roe ~ 1 + ta + ffr +cc +lf', df).fit()
print(pooled)
```

```
                          PooledOLS Estimation Summary
================================================================================
=
Dep. Variable:                    roe   R-squared:                         0.117
7
Estimator:                   PooledOLS   R-squared (Between):               0.000
0
No. Observations:                 360   R-squared (Within):                0.130
7
Date:                Tue, Jul 12 2022   R-squared (Overall):               0.117
7
Time:                        15:19:02   Log-likelihood                     400.1
7
Cov. Estimator:            Unadjusted
                                        F-statistic:                       11.84
3
Entities:                          20   P-value                            0.000
0
Avg Obs:                       18.000   Distribution:                   F(4,35
5)
Min Obs:                       18.000
Max Obs:                       18.000   F-statistic (robust):              11.84
3
                                        P-value                            0.000
0
Time periods:                      18   Distribution:                   F(4,35
5)
Avg Obs:                       20.000
Min Obs:                       20.000
Max Obs:                       20.000

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
Intercept      0.0996     0.0116     8.5915     0.0000      0.0768      0.1223
cc            -0.4812     0.1899    -2.5347     0.0117     -0.8546     -0.1078
ffr            0.0199     0.0061     3.2517     0.0013      0.0078      0.0319
lf             0.0023     0.0006     4.1000     0.0001      0.0012      0.0034
ta             0.0750     0.0334     2.2439     0.0255      0.0093      0.1408
==============================================================================
```

In [9]:
```
# Fixed effects or within estimator
# with constant inclued or not, will have the same results
# with constant term surpressed
from linearmodels import PanelOLS
fixed = PanelOLS.from_formula('roe ~ ta + ffr + cc + lf + EntityEffects', df).fit
print(fixed)
```

```
                          PanelOLS Estimation Summary
================================================================================
=
Dep. Variable:                    roe   R-squared:                        0.130
7
Estimator:                   PanelOLS   R-squared (Between):             -0.497
9
No. Observations:                 360   R-squared (Within):               0.130
7
Date:                Tue, Jul 12 2022   R-squared (Overall):             -0.197
9
Time:                        15:21:28   Log-likelihood                    421.5
7
Cov. Estimator:            Unadjusted
                                        F-statistic:                      12.62
4
Entities:                          20   P-value                           0.000
0
Avg Obs:                       18.000   Distribution:                  F(4,33
6)
Min Obs:                       18.000
Max Obs:                       18.000   F-statistic (robust):             12.62
4
                                        P-value                           0.000
0
Time periods:                      18   Distribution:                  F(4,33
6)
Avg Obs:                       20.000
Min Obs:                       20.000
Max Obs:                       20.000

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
cc            -0.4812     0.1839    -2.6169     0.0093     -0.8430     -0.1195
ffr            0.0199     0.0059     3.3572     0.0009      0.0082      0.0315
lf             0.0023     0.0005     4.2330     0.0000      0.0012      0.0033
ta             0.0750     0.0324     2.3167     0.0211      0.0113      0.1388
==============================================================================

F-test for Poolability: 2.2319
P-value: 0.0025
Distribution: F(19,336)

Included effects: Entity
```

In [10]:
```
# extract fixed effects
fixed.estimated_effects
fixed_effects = fixed.estimated_effects.unstack(level=0).values[0]
print(fixed_effects)
# F test for fixed effects versus OLS
print(fixed.f_pooled)
```

```
[0.08434473 0.09490029 0.09845585 0.1375114  0.05606696 0.12506696
 0.1065114  0.11295585 0.13823362 0.11434473 0.09612251 0.1050114
 0.10634473 0.06212251 0.10534473 0.11556696 0.11662251 0.11556696
 0.03006696 0.0700114 ]
Pooled F-statistic
H0: Effects are zero
Statistic: 2.2319
P-value: 0.0025
Distributed: F(19,336)
```

In [11]:
```
# with constant term included
fixed1 = PanelOLS.from_formula('roe ~ 1 + ta + ffr + cc + lf + EntityEffects', df
print(fixed1)
# extract fixed effects
fixed1.estimated_effects
fixed1_effects = fixed1.params.Intercept + fixed1.estimated_effects.unstack(level
print(fixed1_effects)
```

```
                          PanelOLS Estimation Summary
================================================================================
=
Dep. Variable:                    roe   R-squared:                        0.130
7
Estimator:                   PanelOLS   R-squared (Between):              0.000
0
No. Observations:                 360   R-squared (Within):               0.130
7
Date:                Tue, Jul 12 2022   R-squared (Overall):              0.117
7
Time:                        15:23:28   Log-likelihood                    421.5
7
Cov. Estimator:            Unadjusted
                                        F-statistic:                     12.62
4
Entities:                          20   P-value                           0.000
0
Avg Obs:                       18.000   Distribution:                  F(4,33
6)
Min Obs:                       18.000
Max Obs:                       18.000   F-statistic (robust):            12.62
4
                                        P-value                           0.000
0
Time periods:                      18   Distribution:                  F(4,33
6)
Avg Obs:                       20.000
Min Obs:                       20.000
Max Obs:                       20.000

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
Intercept      0.0996     0.0112     8.8702     0.0000      0.0775      0.1216
cc            -0.4812     0.1839    -2.6169     0.0093     -0.8430     -0.1195
ffr            0.0199     0.0059     3.3572     0.0009      0.0082      0.0315
lf             0.0023     0.0005     4.2330     0.0000      0.0012      0.0033
ta             0.0750     0.0324     2.3167     0.0211      0.0113      0.1388
==============================================================================

F-test for Poolability: 2.2319
P-value: 0.0025
Distribution: F(19,336)

Included effects: Entity
[0.08434473 0.09490029 0.09845585 0.1375114  0.05606696 0.12506696
 0.1065114  0.11295585 0.13823362 0.11434473 0.09612251 0.1050114
```

```
0.10634473 0.06212251 0.10534473 0.11556696 0.11662251 0.11556696
0.03006696 0.0700114 ]
```

In [12]:
```
# F test for fixed effects versus OLS
print(fixed1.f_pooled)
```

```
Pooled F-statistic
H0: Effects are zero
Statistic: 2.2319
P-value: 0.0025
Distributed: F(19,336)
```

In [13]:
```
# Random effects estimator, constant term must be included
# should not have EntityEffects or TimeEffects in the formula
from linearmodels import RandomEffects
random = RandomEffects.from_formula('roe ~ 1 + ta + ffr + cc + lf', df).fit()
print(random)
# extract fixed effects
random.estimated_effects
random_effects = random.params.Intercept + random.estimated_effects.unstack(level
print(random_effects)
print(random.variance_decomposition)
```

```
                          RandomEffects Estimation Summary
================================================================================
=
Dep. Variable:                     roe   R-squared:                         0.125
8
Estimator:              RandomEffects   R-squared (Between):           -2.22e-1
6
No. Observations:                 360   R-squared (Within):                0.130
7
Date:                Tue, Jul 12 2022   R-squared (Overall):               0.117
7
Time:                        15:24:46   Log-likelihood                    413.7
1
Cov. Estimator:            Unadjusted
                                        F-statistic:                       12.76
8
Entities:                          20   P-value                           0.000
0
Avg Obs:                       18.000   Distribution:                   F(4,35
5)
Min Obs:                       18.000
Max Obs:                       18.000   F-statistic (robust):              12.76
8
                                        P-value                           0.000
0
Time periods:                      18   Distribution:                   F(4,35
5)
Avg Obs:                       20.000
Min Obs:                       20.000
Max Obs:                       20.000

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
Intercept      0.0996     0.0124     8.0014     0.0000      0.0751      0.1240
cc            -0.4812     0.1829    -2.6318     0.0089     -0.8409     -0.1216
ffr            0.0199     0.0059     3.3763     0.0008      0.0083      0.0314
lf             0.0023     0.0005     4.2571     0.0000      0.0012      0.0033
ta             0.0750     0.0322     2.3298     0.0204      0.0117      0.1384
==============================================================================
[0.09339304 0.09767079 0.09911171 0.11493938 0.08193317 0.10989614
 0.10237631 0.10498799 0.11523207 0.10555085 0.0981661  0.10176842
 0.10230877 0.08438725 0.10190351 0.10604617 0.10647394 0.10604617
 0.0713964  0.0875843 ]
Effects                  0.000612
```

```
Residual                  0.006031
Percent due to Effects    0.092153
Name: Variance Decomposition, dtype: float64
```

In [15]:
```
from linearmodels.panel.results import compare
```

In [16]:
```
# compare fixed effects and random effects models
res2 = {'Pooled':pooled,'Fixed+1':fixed1,'Fixed':fixed,'Random':random}
print(compare(res2))

effects = pd.DataFrame({'Fixed Effects':fixed_effects,'Random Effects':random_eff
                        index=pdata.index.levels[0])
print(effects)
```

```
                                Model Comparison
================================================================================
=========
                                Pooled        Fixed+1          Fixed
Random
--------------------------------------------------------------------------------
---------
Dep. Variable                      roe            roe            roe
roe
Estimator                   PooledOLS       PanelOLS       PanelOLS     Rand
omEffects
No. Observations                   360            360            360
360
Cov. Est.                  Unadjusted     Unadjusted     Unadjusted        U
nadjusted
R-squared                      0.1177         0.1307         0.1307
0.1258
R-Squared (Within)             0.1307         0.1307         0.1307
0.1307
R-Squared (Between)            0.0000         0.0000        -0.4979
-2.22e-16
R-Squared (Overall)            0.1177         0.1177        -0.1979
0.1177
F-statistic                    11.843         12.624         12.624
12.768
P-value (F-stat)               0.0000         0.0000         0.0000
0.0000
====================     ============   ============   ============   ======
=========
Intercept                      0.0996         0.0996
0.0996
                              (8.5915)       (8.8702)
(8.0014)
cc                            -0.4812        -0.4812        -0.4812
-0.4812
                              (-2.5347)      (-2.6169)      (-2.6169)
(-2.6318)
ffr                            0.0199         0.0199         0.0199
0.0199
                              (3.2517)       (3.3572)       (3.3572)
(3.3763)
lf                             0.0023         0.0023         0.0023
0.0023
                              (4.1000)       (4.2330)       (4.2330)
(4.2571)
ta                             0.0750         0.0750         0.0750
0.0750
                              (2.2439)       (2.3167)       (2.3167)
(2.3298)
```

```
====================== ============== ============== ============== ========
=========
Effects                                                 Entity         Entity
------------------------------------------------------------------------------
---------

T-stats reported in parentheses
      Fixed Effects  Random Effects
ID
ABCB       0.084345        0.093393
ASB        0.094900        0.097671
AUB        0.098456        0.099112
BANF       0.137511        0.114939
BANR       0.056067        0.081933
BOKF       0.125067        0.109896
CATY       0.106511        0.102376
CBU        0.112956        0.104988
CFR        0.138234        0.115232
CMA        0.114345        0.105551
COLB       0.096123        0.098166
SASR       0.105011        0.101768
SFNC       0.106345        0.102309
SNV        0.062123        0.084387
SSB        0.105345        0.101904
TCBI       0.115567        0.106046
TRMK       0.116623        0.106474
UBSI       0.115567        0.106046
UCBI       0.030067        0.071396
ZION       0.070011        0.087584
```

In [17]:
```
# LM test for random effects versus OLS
n = pdata.index.levels[0].size
T = pdata.index.levels[1].size
D = np.kron(np.eye(n), np.ones(T)).T
e = pooled.resids
LM = (e.dot(D).dot(D.T).dot(e) / e.dot(e) - 1) ** 2 * n * T / 2 / (T - 1)
LM_pvalue = stats.chi2(1).sf(LM)
print("LM Test: chisq = {0}, df = 1, p-value = {1}".format(LM, LM_pvalue))
```

```
LM Test: chisq = 10.955797257255929, df = 1, p-value = 0.0009331121542716446
```

In [18]:
```
# Hausman test for fixed versus random effects model
# null hypothesis: random effects model
psi = fixed.cov - random.cov.iloc[1:,1:]
diff = fixed.params - random.params[1:]
# psi = fixed1.cov.iloc[1:,1:] - random.cov.iloc[1:,1:]
# diff = fixed1.params[1:] - random.params[1:]
W = diff.dot(np.linalg.inv(psi)).dot(diff)
dof = random.params.size -1
pvalue = stats.chi2(dof).sf(W)
print("Hausman Test: chisq = {0}, df = {1}, p-value = {2}".format(W, dof, pvalue)
```

```
Hausman Test: chisq = 1.0752048114488193e-27, df = 4, p-value = 1.0
```

In [20]:
```
# panel robust hetero cov
fixed_robust = PanelOLS.from_formula('roe ~ 1 + ta + ffr + cc + lf + EntityEffect
print(fixed_robust)
random_robust = RandomEffects.from_formula('roe ~ 1 + ta + ffr + cc + lf', df).fi
print(random_robust)
```

```
                          PanelOLS Estimation Summary
================================================================================
=
Dep. Variable:                     roe   R-squared:                        0.130
7
Estimator:                   PanelOLS   R-squared (Between):              0.000
0
No. Observations:                 360   R-squared (Within):               0.130
7
Date:                Tue, Jul 12 2022   R-squared (Overall):              0.117
7
Time:                        15:29:28   Log-likelihood                    421.5
7
Cov. Estimator:            Clustered
                                        F-statistic:                     12.62
4
Entities:                          20   P-value                           0.000
0
Avg Obs:                       18.000   Distribution:                  F(4,33
6)
Min Obs:                       18.000
Max Obs:                       18.000   F-statistic (robust):             31.87
0
                                        P-value                           0.000
0
Time periods:                      18   Distribution:                  F(4,33
6)
Avg Obs:                       20.000
Min Obs:                       20.000
Max Obs:                       20.000

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
Intercept      0.0996     0.0048     20.901     0.0000      0.0902      0.1089
cc            -0.4812     0.1047    -4.5959     0.0000     -0.6872     -0.2753
ffr            0.0199     0.0033     6.0110     0.0000      0.0134      0.0264
lf             0.0023     0.0006     3.8273     0.0002      0.0011      0.0034
ta             0.0750     0.0222     3.3799     0.0008      0.0314      0.1187
==============================================================================

F-test for Poolability: 2.2319
P-value: 0.0025
Distribution: F(19,336)

Included effects: Entity
                        RandomEffects Estimation Summary
================================================================================
=
```

Dep. Variable:	roe	R-squared:	0.1258
Estimator:	RandomEffects	R-squared (Between):	-2.22e-16
No. Observations:	360	R-squared (Within):	0.1307
Date:	Tue, Jul 12 2022	R-squared (Overall):	0.1177
Time:	15:29:28	Log-likelihood	413.71
Cov. Estimator:	Clustered		
		F-statistic:	12.768
Entities:	20	P-value	0.0000
Avg Obs:	18.000	Distribution:	F(4,355)
Min Obs:	18.000		
Max Obs:	18.000	F-statistic (robust):	31.870
		P-value	0.0000
Time periods:	18	Distribution:	F(4,355)
Avg Obs:	20.000		
Min Obs:	20.000		
Max Obs:	20.000		

Parameter Estimates

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
Intercept	0.0996	0.0093	10.654	0.0000	0.0812	0.1179
cc	-0.4812	0.1047	-4.5959	0.0000	-0.6872	-0.2753
ffr	0.0199	0.0033	6.0110	0.0000	0.0134	0.0264
lf	0.0023	0.0006	3.8273	0.0002	0.0011	0.0034
ta	0.0750	0.0222	3.3799	0.0008	0.0314	0.1187

In [21]:
```
# compare fixed effects and random effects models
res3 = {'Fixed (Panel-Robust)':fixed_robust,'Random (Panel-Robust)':random_robust
print(compare(res3))
```

```
                          Model Comparison
=========================================================================
                        Fixed (Panel-Robust) Random (Panel-Robust)
-------------------------------------------------------------------------
Dep. Variable                            roe                       roe
Estimator                          PanelOLS               RandomEffects
No. Observations                        360                       360
Cov. Est.                         Clustered                 Clustered
R-squared                            0.1307                    0.1258
R-Squared (Within)                   0.1307                    0.1307
R-Squared (Between)                  0.0000                -2.22e-16
R-Squared (Overall)                  0.1177                    0.1177
F-statistic                          12.624                    12.768
P-value (F-stat)                     0.0000                    0.0000
=====================           ===========           ===============
Intercept                            0.0996                    0.0996
                                    (20.901)                  (10.654)
cc                                  -0.4812                   -0.4812
                                   (-4.5959)                 (-4.5959)
ffr                                  0.0199                    0.0199
                                    (6.0110)                  (6.0110)
lf                                   0.0023                    0.0023
                                    (3.8273)                  (3.8273)
ta                                   0.0750                    0.0750
                                    (3.3799)                  (3.3799)
=======================         =============     =================
Effects                                Entity
-------------------------------------------------------------------------

T-stats reported in parentheses
```

In [ ]:
```
```