Infrastructure Gap

Paul Mukoki

8/6/2021

Contents

Preliminary: Read in the data

Here I load the dataset

```
mukoki_gap_data <- readxl::read_xlsx("dataset_40_countries.xlsx") %>%

  janitor::clean_names() %>%

  data.frame()
```

Dependent and Independent Variables

I select the dependent and independent variables

```
modelling_data <- mukoki_gap_data %>%

  select(countries, country, state, year,

         income_group_2009, income_2009,

         overall_index, infra_index,

         transport_gap, electricity_gap,

         ict_gap, water_gap, overall_index,

         transport_invest, electricity_invest,

         ict_invest, water_invest, water_invest,

         stock_mkt_cap, human_di, exchange_rate,

         rinterestrate, percent_change_in_gfcf_public,
```

```
        fdi_gdp, govenance_index) %>%

  tibble()
```

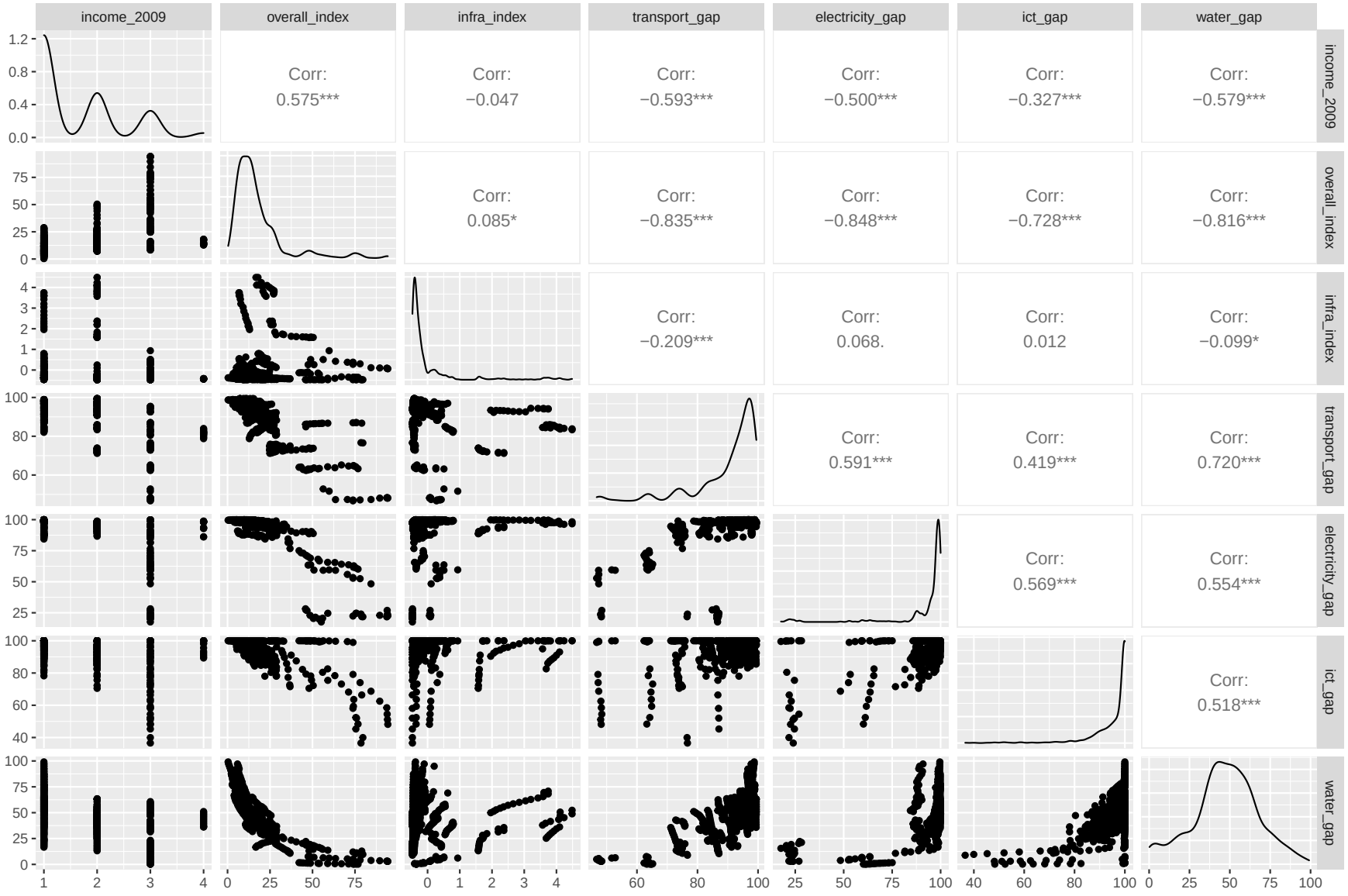
Table 1: Independent Variables

skim_variable	numeric.mean	numeric.sd	numeric.p0	numeric.p25	numeric.p50	numeric.p75	numeric.p100
income_2009	1.62500	0.8279167	1.0000000	1.000000	1.00000	2.00000	4.00000
overall_index	17.86327	15.9435392	0.3687802	8.439691	13.49782	20.30228	94.32366
transport_gap	90.59428	10.2512712	46.6914350	87.967433	94.29789	97.38841	99.44557
electricity_gap	92.88688	15.0433290	17.6244140	94.715203	98.05586	99.17087	100.00000
ict_gap	95.26291	8.7731510	36.5554970	93.874929	99.55578	99.99659	99.99999
water_gap	47.30374	20.1042116	0.2118731	36.646024	48.16911	60.48719	99.10913

Dependent Variables

```
modelling_data %>%
  select(income_2009,
         overall_index,
         transport_gap, electricity_gap,
         ict_gap, water_gap) %>%
  na.omit() %>%
  skim_without_charts() %>%
  select(-n_missing, -complete_rate, -skim_type) %>%
  kbl(., booktabs = TRUE, caption = "Independent Variables") %>%
  kable_classic(full_width = TRUE)
```

```
modelling_data %>%
  select(income_2009,
         overall_index, infra_index,
         transport_gap, electricity_gap,
         ict_gap, water_gap) %>%
  na.omit() %>%
  ggpairs()
```


income_2009
overall_index
infra_index
transport_gap
electricity_gap
ict_gap
water_gap
Corr: 0.575***
Corr: −0.047
Corr: −0.593***
Corr: −0.500***
Corr: −0.327***
Corr: −0.579***
Corr: 0.085*
Corr: −0.835***
Corr: −0.848***
Corr: −0.728***
Corr: −0.816***
Corr: −0.209***
Corr: 0.068.
Corr: 0.012
Corr: −0.099*
Corr: 0.591***
Corr: 0.419***
Corr: 0.720***
Corr: 0.569***
Corr: 0.554***
Corr: 0.518***

Independent Variables

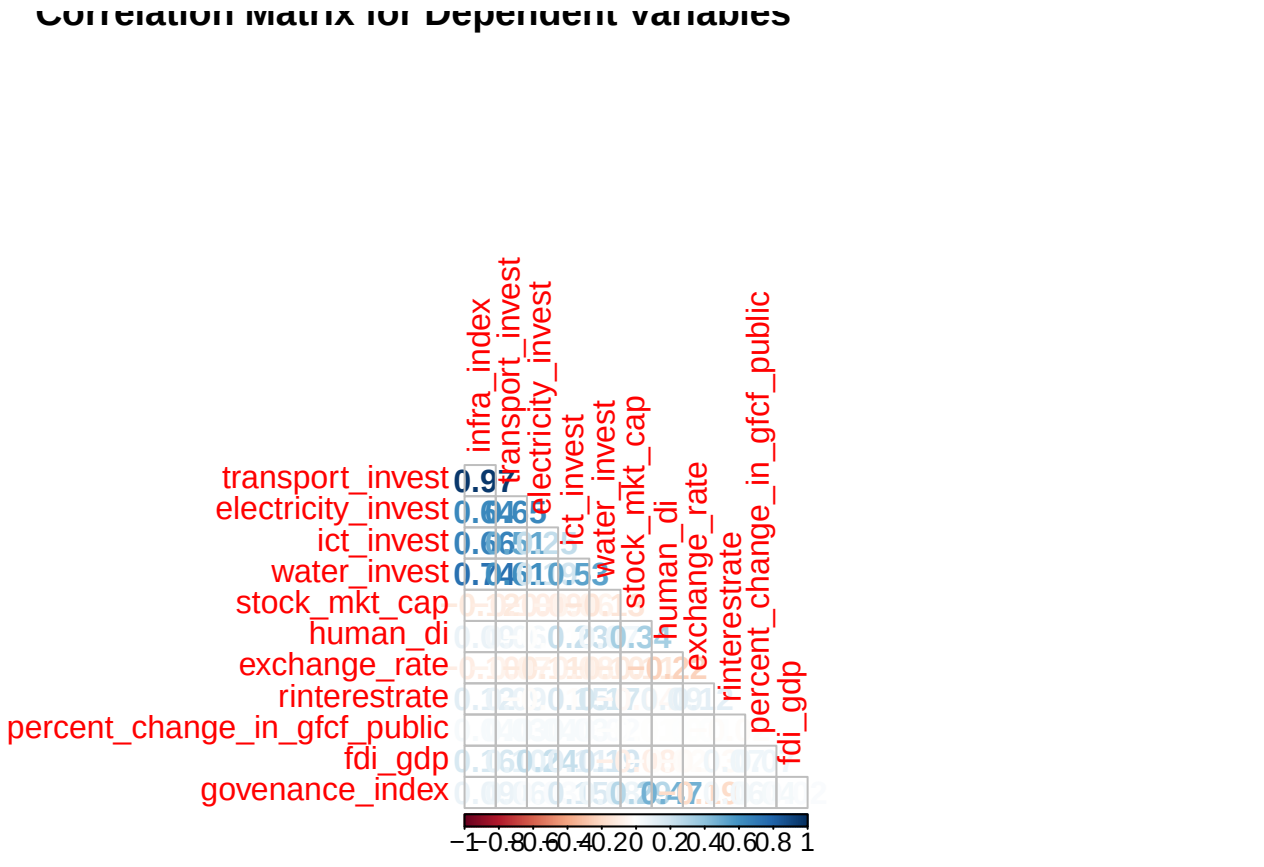
We examine the independent variables. Please confirm that these are correct before I go ahead and run the regressions. I am especially keen on overall index and its meaning. What variable in the data stands for Public-Private Partnerships (PPP) investment in Aggregate/overall infrastructure? I dont think its the overall index.

```
modelling_data %>%

  select(infra_index,

         transport_invest, electricity_invest,

         ict_invest, water_invest, water_invest,

         stock_mkt_cap, human_di, exchange_rate,

         rinterestrate, percent_change_in_gfcf_public,

         fdi_gdp, govenance_index) %>%

  skim_without_charts() %>%

  select(-n_missing, -complete_rate, -skim_type) %>%

  kbl(., booktabs = TRUE, caption = "Independent Variables") %>%

  kable_classic(full_width = TRUE)
```

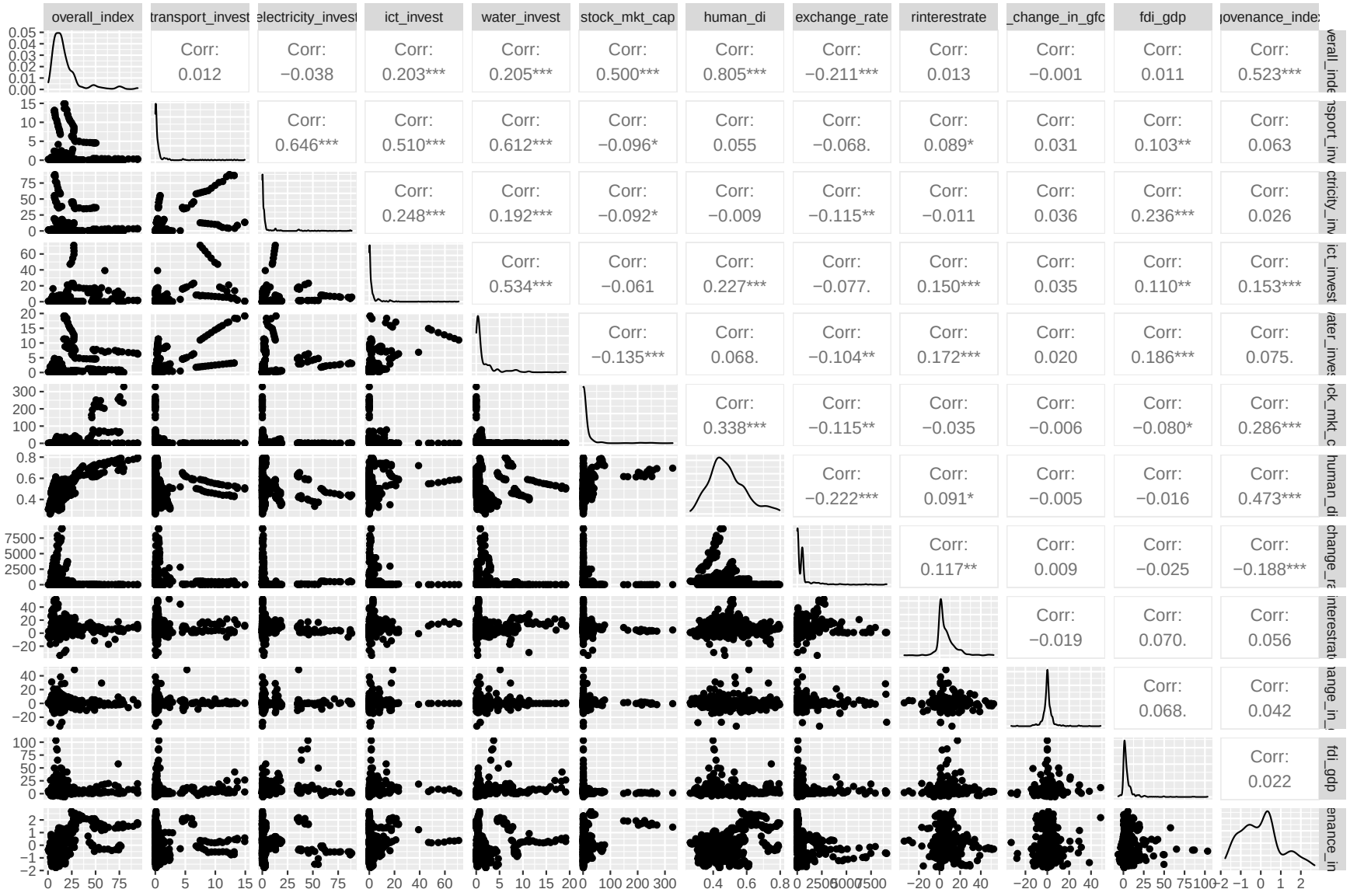
Table 2: Independent Variables

skim_variable	numeric.mean	numeric.sd	numeric.p0	numeric.p25	numeric.p50	numeric.p75	numeric.p100
infra_index	0.0000000	0.9085647	-0.4701808	-0.4064351	-0.3251464	-0.1158849	4.485926
transport_invest	0.9829726	2.4573516	0.0000908	0.0520468	0.1884312	0.5139178	14.905687
electricity_invest	5.6575235	14.3728215	0.0013178	0.1495541	0.6089139	2.3695731	88.429515
ict_invest	2.8558644	7.2741104	0.0001807	0.1146541	0.6695169	2.1571728	70.985354
water_invest	1.9211378	3.1850498	0.0146976	0.2785261	0.5765471	2.0650996	19.154820
stock_mkt_cap	10.2802022	36.0435167	0.7241040	1.0000000	1.0000000	1.0000000	328.361000
human_di	0.4838609	0.1074758	0.2630000	0.4120000	0.4680000	0.5485000	0.797000
exchange_rate	652.2826577	1260.5619533	0.8667643	17.6843045	370.4251100	555.4500000	9088.320000
rinterestrate	6.1573402	9.0830600	-34.2103090	1.0000000	3.8272198	9.2988998	52.436821
percent_change_in_gfcf	0.3170261	5.5788054	-33.0171620	-1.3669003	0.0000000	1.5717645	48.789712
fdi_gdp	4.9825751	9.3207813	-6.3698773	1.0000000	2.5635350	5.3448228	103.337390
govenance_index	0.0000000	0.9842590	-1.7782380	-0.7445558	-0.0166953	0.4886982	2.688182

```
modelling_data %>%
  select(infra_index,
         transport_invest, electricity_invest,
         ict_invest, water_invest, water_invest,
         stock_mkt_cap, human_di, exchange_rate,
         rinterestrate, percent_change_in_gfcf_public,
         fdi_gdp, govenance_index) %>%
  drop_na(percent_change_in_gfcf_public) %>%
  cor() %>%
  corrplot(method = "number", type = "lower",
           diag = FALSE, number.digits = 2,
           title = "Correlation Matrix for Dependent Variables")
```



```
modelling_data %>%
  select(overall_index,
         transport_invest, electricity_invest,
         ict_invest, water_invest, water_invest,
         stock_mkt_cap, human_di, exchange_rate,
         rinterestrate, percent_change_in_gfcf_public,
         fdi_gdp, govenance_index) %>%
  drop_na(percent_change_in_gfcf_public) %>%
  ggpairs()
```


overall_index
transport_invest
electricity_invest
ict_invest
water_invest
stock_mkt_cap
human_di
exchange_rate
rinterestrate
_change_in_gfc
fdi_gdp
provenance_index
Corr: 0.012
Corr: −0.038
Corr: 0.203***
Corr: 0.205***
Corr: 0.500***
Corr: 0.805***
Corr: −0.211***
Corr: 0.013
Corr: −0.001
Corr: 0.011
Corr: 0.523***
Corr: 0.646***
Corr: 0.510***
Corr: 0.612***
Corr: −0.096*
Corr: 0.055
Corr: −0.068.
Corr: 0.089*
Corr: 0.031
Corr: 0.103**
Corr: 0.063
Corr: 0.248***
Corr: 0.192***
Corr: −0.092*
Corr: −0.009
Corr: −0.115**
Corr: −0.011
Corr: 0.036
Corr: 0.236***
Corr: 0.026
Corr: 0.534***
Corr: −0.061
Corr: 0.227***
Corr: −0.077.
Corr: 0.150***
Corr: 0.035
Corr: 0.110**
Corr: 0.153***
Corr: −0.135***
Corr: 0.068.
Corr: −0.104**
Corr: 0.172***
Corr: 0.020
Corr: 0.186***
Corr: 0.075.
Corr: 0.338***
Corr: −0.115**
Corr: −0.035
Corr: −0.006
Corr: −0.080*
Corr: 0.286***
Corr: −0.222***
Corr: 0.091*
Corr: −0.005
Corr: −0.016
Corr: 0.473***
Corr: 0.117**
Corr: 0.009
Corr: −0.025
Corr: −0.188***
Corr: −0.019
Corr: 0.070.
Corr: 0.056
Corr: 0.068.
Corr: 0.042
Corr: 0.022

The Fixed Effects Model

Write a regression function to run fixed and random effects models

```
model_fixed <- plm(overall_index ~ transport_invest + electricity_invest +

        ict_invest + water_invest + water_invest +

        stock_mkt_cap + human_di + exchange_rate +

        rinterestrate + percent_change_in_gfcf_public +

        fdi_gdp + govenance_index, data = modelling_data,

    effect = "twoways", model = "within", index = c("country", "year"))

summary(model_fixed)
```

```
## Twoways effects Within Model
##
## Call:
## plm(formula = overall_index ~ transport_invest + electricity_invest +
##     ict_invest + water_invest + water_invest + stock_mkt_cap +
##     human_di + exchange_rate + rinterestrate + percent_change_in_gfcf_public +
##     fdi_gdp + govenance_index, data = modelling_data, effect = "twoways",
##     model = "within", index = c("country", "year"))
##
## Unbalanced Panel: n = 40, T = 15-16, N = 638
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -20.683049  -1.525893   0.073119   1.348739  15.303867
##
## Coefficients:
##                                 Estimate  Std. Error t-value
## transport_invest              -0.35325707  0.38003978 -0.9295
## electricity_invest             0.21568781  0.06320574  3.4125
## ict_invest                    -0.36417626  0.04459065 -8.1671
## water_invest                  -2.41149004  0.41918167 -5.7529
## stock_mkt_cap                  0.01172604  0.01122126  1.0450
## human_di                      -1.77500572 13.29340967 -0.1335
## exchange_rate                 -0.00168294  0.00033569 -5.0134
## rinterestrate                  0.01380202  0.02280432  0.6052
## percent_change_in_gfcf_public  0.00173114  0.02468044  0.0701
## fdi_gdp                        0.00319925  0.01882435  0.1700
## govenance_index                0.29464569  0.67132286  0.4389
##                                  Pr(>|t|)
## transport_invest                 0.3530084
## electricity_invest               0.0006893 ***
## ict_invest                    0.00000000000000203 ***
## water_invest                  0.00000001430627886 ***
## stock_mkt_cap                    0.2964719
## human_di                         0.8938250
## exchange_rate                 0.00000071442260501 ***
## rinterestrate                    0.5452615
```

```
## percent_change_in_gfcf_public           0.9441050
## fdi_gdp                                  0.8651075
## govenance_index                          0.6608975
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    7537.7
## Residual Sum of Squares: 6439.9
## R-Squared:      0.14564
## Adj. R-Squared: 0.048552
## F-statistic: 8.86419 on 11 and 572 DF, p-value: 0.000000000000011985
```

The Random Effects Model

```
model_random <- plm(overall_index ~ transport_invest + electricity_invest +

        ict_invest + water_invest + water_invest +

        stock_mkt_cap + human_di + exchange_rate +

        rinterestrate + percent_change_in_gfcf_public +

        fdi_gdp + govenance_index, data = modelling_data,

    effect = "twoways", model = "random", index = c("country", "year"))

summary(model_random)
```

```
## Twoways effects Random Effect Model
##    (Swamy-Arora's transformation)
##
## Call:
## plm(formula = overall_index ~ transport_invest + electricity_invest +
##     ict_invest + water_invest + water_invest + stock_mkt_cap +
##     human_di + exchange_rate + rinterestrate + percent_change_in_gfcf_public +
##     fdi_gdp + govenance_index, data = modelling_data, effect = "twoways",
##     model = "random", index = c("country", "year"))
##
## Unbalanced Panel: n = 40, T = 15-16, N = 638
##
## Effects:
##                  var std.dev share
## idiosyncratic 11.259   3.355 0.217
## individual    40.689   6.379 0.783
## time           0.000   0.000 0.000
## theta:
##          Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## id    0.865418 0.8696177 0.8696177 0.8694203 0.8696177 0.8696177
## time  0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
## total 0.000000 0.0000000 0.0000000 0.0000000 0.0000000 0.0000000
##
## Residuals:
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -24.366  -6.697  -1.373  -0.015   6.074  45.035
```

```
##
## Coefficients:
##                                 Estimate    Std. Error  z-value
## (Intercept)                  -32.402214761   0.683121951 -47.4325
## transport_invest              -0.391223792   0.098438367  -3.9743
## electricity_invest             0.079203227   0.015440069   5.1297
## ict_invest                    -0.182765283   0.010503074 -17.4011
## water_invest                  -0.561587855   0.082234318  -6.8291
## stock_mkt_cap                  0.025544717   0.003035691   8.4148
## human_di                     107.469091339   1.169534367  91.8905
## exchange_rate                 -0.001025873   0.000090426 -11.3449
## rinterestrate                  0.062362431   0.006468734   9.6406
## percent_change_in_gfcf_public -0.015441537   0.007255643  -2.1282
## fdi_gdp                       -0.034249089   0.005406611  -6.3347
## govenance_index                0.091614655   0.161133645   0.5686
##                                   Pr(>|z|)
## (Intercept)                  < 0.00000000000000022 ***
## transport_invest               0.000070585969750 ***
## electricity_invest             0.000000290173714 ***
## ict_invest                   < 0.00000000000000022 ***
## water_invest                   0.000000000008544 ***
## stock_mkt_cap                < 0.00000000000000022 ***
## human_di                     < 0.00000000000000022 ***
## exchange_rate                < 0.00000000000000022 ***
## rinterestrate                < 0.00000000000000022 ***
## percent_change_in_gfcf_public                0.03332 *
## fdi_gdp                        0.000000000237852 ***
## govenance_index                          0.56965
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    162300
## Residual Sum of Squares: 67518
## R-Squared:      0.58401
## Adj. R-Squared: 0.5767
## Chisq: 10321 on 11 DF, p-value: < 0.000000000000000222
```

The Hausmann Test

Given that the test is significant, then we should use the fixed effects model .

```
phtest(model_fixed, model_random)
```

```
##
##  Hausman Test
##
## data:  overall_index ~ transport_invest + electricity_invest + ict_invest +  ...
## chisq = 115.01, df = 11, p-value < 0.00000000000000022
## alternative hypothesis: one model is inconsistent
```

Cross-Sectional Dependence

The model shows the presence of cross sectional dependence. Thus, the unit roots tests have to consider the presence of cross sectional dependence in the panels.

```
pcdtest(model_fixed, method = "cd")
```

```
##
##  Pesaran CD test for cross-sectional dependence in panels
##
## data: overall_index ~ transport_invest + electricity_invest + ict_invest +    water_invest + water_invest + stock_mkt
## z = 4.7227, p-value = 0.000002327
## alternative hypothesis: cross-sectional dependence
```

Panel unit root tests in the Presence of Cross-Sectional Dependence

The main unit root tests implemented in R's plm and punitroots, namely the tests proposed in Maddala and Wu (1999), Choi (2001), Levin, Lin, and Chu (2002), Im, Pesaran, and Shin (2003) (plm); Demetrescu, Hassler, and Tarcolea (2006), Hanck (2008), Costantini and Lupi (2011) (punitroots). The first four tests assume that the series under scrutiny are cross-sectionally independent: the last three allow for cross-dependence across the panel units.