Data_Analysi

Yeshwant Chillakuru

4/20/2017

Load Data

Descriptive

Regressions

Base Specification

Any Disease

```
plm.1f <- plm(data=df, mbAnyDisease ~ roadPaved, index = c("id", "year"), model="within")
plm.2f <- plm(data=df, mbAnyDisease ~ roadPaved + income, index = c("id", "year"), model="within")
plm.3f <- plm(data=df, mbAnyDisease ~ roadPaved + literate, index = c("id", "year"), model="within")
plm.4f <- plm(data=df, mbAnyDisease ~ roadPaved + smokeTobacco, index = c("id", "year"), model="within")
plm.5f <- plm(data=df, mbAnyDisease ~ roadPaved + ownToilet, index = c("id", "year"), model="within")
plm.6f <- plm(data=df, mbAnyDisease ~ roadPaved + electricity, index = c("id", "year"), model="within")
plm.7f <- plm(data=df, mbAnyDisease ~ roadPaved + seenDoctor, index = c("id", "year"), model="within")
plm.8f <- plm(data=df, mbAnyDisease ~ roadPaved + income + smokeTobacco, index = c("id", "year"), model=


r.se.1 <- sqrt(diag(vcovHC(plm.1f, type = "HC1")))
r.se.2 <- sqrt(diag(vcovHC(plm.2f, type = "HC1")))
r.se.3 <- sqrt(diag(vcovHC(plm.3f, type = "HC1")))
r.se.4 <- sqrt(diag(vcovHC(plm.4f, type = "HC1")))
r.se.5 <- sqrt(diag(vcovHC(plm.5f, type = "HC1")))
r.se.6 <- sqrt(diag(vcovHC(plm.6f, type = "HC1")))
r.se.7 <- sqrt(diag(vcovHC(plm.7f, type = "HC1")))
r.se.8 <- sqrt(diag(vcovHC(plm.8f, type = "HC1")))

stargazer(plm.1f, plm.2f, plm.3f, plm.4f, plm.5f, plm.6f, plm.7f, plm.8f, se=list(r.se.1, r.se.2, r.se.
```

```
##
## Road Paved and Incidence of Any Disease: OLS with Village and Time-Fixed Effects
## ================================================================================================
##                                          Dependent variable:
##              ------------------------------------------------------------------------
##                                              mbAnyDisease
##                 (1)       (2)       (3)       (4)       (5)       (6)       (7)       (8)
## ------------------------------------------------------------------------------------------------
## roadPaved    1.584*** 0.926*** 1.178*** 0.538**  1.263*** 1.210*** 1.295***   0.476*
##              (0.275)  (0.280)  (0.280)  (0.272)  (0.278)  (0.287)   (0.283)  (0.273)
##
## income                1.764***                                                 0.837***
##                       (0.171)                                                  (0.220)
```

```
##
## literate                        9.272***
##                                 (1.247)
##
## smokeTobacco                                1.390***                                1.058***
##                                             (0.116)                                 (0.155)
##
## ownToilet                                             0.048***
##                                                       (0.006)
##
## electricity                                                     0.030***
##                                                                 (0.005)
##
## seenDoctor                                                                -0.060***
##                                                                            (0.011)
##
## ------------------------------------------------------------------------------------
## Observations  2,544    2,544    2,544    2,544    2,544    2,544     2,544    2,544
## R2            0.030    0.110    0.076    0.140    0.093    0.059     0.049    0.151
## Adjusted R2   -1.284   -1.098   -1.178   -1.028   -1.137   -1.219   -1.241    -1.002
## ====================================================================================
## Note:                                                    *p<0.1; **p<0.05; ***p<0.01
```

```
plm.1f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd, index = c("id", "year"), model="within")
plm.2f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd + income, index = c("id", "year"), model="wit
plm.3f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd + literate, index = c("id", "year"), model="w
plm.4f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd + smokeTobacco, index = c("id", "year"), mode
plm.5f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd + ownToilet, index = c("id", "year"), model="
plm.6f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd + electricity, index = c("id", "year"), model
plm.7f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd + seenDoctor, index = c("id", "year"), model=
plm.8f <- plm(data=dfNoRoad, mbAnyDisease ~ lnDistToPvdRd + income + smokeTobacco, index = c("id", "yea


r.se.1 <- sqrt(diag(vcovHC(plm.1f, type = "HC1")))
r.se.2 <- sqrt(diag(vcovHC(plm.2f, type = "HC1")))
r.se.3 <- sqrt(diag(vcovHC(plm.3f, type = "HC1")))
r.se.4 <- sqrt(diag(vcovHC(plm.4f, type = "HC1")))
r.se.5 <- sqrt(diag(vcovHC(plm.5f, type = "HC1")))
r.se.6 <- sqrt(diag(vcovHC(plm.6f, type = "HC1")))
r.se.7 <- sqrt(diag(vcovHC(plm.7f, type = "HC1")))
r.se.8 <- sqrt(diag(vcovHC(plm.8f, type = "HC1")))

stargazer(plm.1f, plm.2f, plm.3f, plm.4f, plm.5f, plm.6f, plm.7f, plm.8f, se=list(r.se.1, r.se.2, r.se.
```

```
##
## Road Paved and Incidence of Any Disease: OLS with Village and Time-Fixed Effects
## ================================================================================
##                                      Dependent variable:
##                -----------------------------------------------------------------
##                                          mbAnyDisease
##                  (1)      (2)      (3)      (4)      (5)      (6)       (7)      (8)
## --------------------------------------------------------------------------------
## lnDistToPvdRd -0.835* -0.813*  -0.530   -0.175  -0.609   -0.687   -0.599    -0.102
##               (0.442) (0.429)  (0.447) (0.406)  (0.429) (0.487)   (0.454)  (0.427)
##
```

```
## income                3.039***                                                -0.836
##                       (0.924)                                                 (1.272)
##
## literate                        8.468*
##                                 (4.369)
##
## smokeTobacco                            2.238***                            2.503***
##                                         (0.306)                             (0.436)
##
## ownToilet                                       0.058**
##                                                 (0.026)
##
## electricity                                             0.046***
##                                                         (0.014)
##
## seenDoctor                                                      -0.072***
##                                                                  (0.024)
##
## ------------------------------------------------------------------------------------
## Observations    599     599      599     599      599     599      599      599
## R2             0.022   0.134    0.059   0.319    0.107   0.114    0.059    0.323
## Adjusted R2   -4.570  -3.978  -4.412   -2.916  -4.134  -4.096   -4.411   -2.929
## ====================================================================================
## Note:                                                 *p<0.1; **p<0.05; ***p<0.01
```

Communicable Disease

Non-Communicable Disease

```
plm.1f <- plm(data=df, mbNonComDisease ~ roadPaved, index = c("id", "year"), model="within")
plm.2f <- plm(data=df, mbNonComDisease ~ roadPaved + income, index = c("id", "year"), model="within")
plm.3f <- plm(data=df, mbNonComDisease ~ roadPaved + literate, index = c("id", "year"), model="within")
plm.4f <- plm(data=df, mbNonComDisease ~ roadPaved + smokeTobacco, index = c("id", "year"), model="with
plm.5f <- plm(data=df, mbNonComDisease ~ roadPaved + ownToilet, index = c("id", "year"), model="within"
plm.6f <- plm(data=df, mbNonComDisease ~ roadPaved + electricity, index = c("id", "year"), model="withi
plm.7f <- plm(data=df, mbNonComDisease ~ roadPaved + seenDoctor, index = c("id", "year"), model="within
plm.8f <- plm(data=df, mbNonComDisease ~ roadPaved + income + smokeTobacco, index = c("id", "year"), mo


r.se.1 <- sqrt(diag(vcovHC(plm.1f, type = "HC1")))
r.se.2 <- sqrt(diag(vcovHC(plm.2f, type = "HC1")))
r.se.3 <- sqrt(diag(vcovHC(plm.3f, type = "HC1")))
r.se.4 <- sqrt(diag(vcovHC(plm.4f, type = "HC1")))
r.se.5 <- sqrt(diag(vcovHC(plm.5f, type = "HC1")))
r.se.6 <- sqrt(diag(vcovHC(plm.6f, type = "HC1")))
r.se.7 <- sqrt(diag(vcovHC(plm.7f, type = "HC1")))
r.se.8 <- sqrt(diag(vcovHC(plm.8f, type = "HC1")))

stargazer(plm.1f, plm.2f, plm.3f, plm.4f, plm.5f, plm.6f, plm.7f, plm.8f, se=list(r.se.1, r.se.2, r.se.
```

```
##
## Road Paved and Incidence of Non Communicable Disease: OLS with Village and Time-Fixed Effects
## ====================================================================================
```

```
##                                    Dependent variable:
##              ---------------------------------------------------------------------------
##                                        mbNonComDisease
##                (1)      (2)      (3)      (4)      (5)      (6)       (7)      (8)
## ---------------------------------------------------------------------------------------
## roadPaved    0.153***  0.089*  0.108**   0.049   0.130*** 0.117**   0.115**   0.044
##              (0.045)  (0.047)  (0.047)  (0.050)  (0.045)  (0.048)   (0.049)  (0.050)
##
## income                0.170***                                                 0.076
##                       (0.036)                                                 (0.050)
##
## literate                       1.027***
##                                (0.298)
##
## smokeTobacco                             0.138***                           0.107***
##                                          (0.027)                            (0.036)
##
## ownToilet                                         0.003***
##                                                   (0.001)
##
## electricity                                                 0.003***
##                                                             (0.001)
##
## seenDoctor                                                           -0.008***
##                                                                       (0.003)
##
## ---------------------------------------------------------------------------------------
## Observations  2,544    2,544    2,544    2,544    2,544    2,544     2,544    2,544
## R2            0.007    0.027    0.022    0.036    0.016    0.014     0.016    0.038
## Adjusted R2   -1.337   -1.293   -1.305   -1.273   -1.319   -1.323   -1.320    -1.269
## =======================================================================================
## Note:                                                    *p<0.1; **p<0.05; ***p<0.01
```

```
plm.1f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd, index = c("id", "year"), model="within")
plm.2f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd + income, index = c("id", "year"), model="
plm.3f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd + literate, index = c("id", "year"), model=
plm.4f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd + smokeTobacco, index = c("id", "year"), mo
plm.5f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd + ownToilet, index = c("id", "year"), model
plm.6f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd + electricity, index = c("id", "year"), mod
plm.7f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd + seenDoctor, index = c("id", "year"), mode
plm.8f <- plm(data=dfNoRoad, mbNonComDisease ~ lnDistToPvdRd + income + smokeTobacco, index = c("id", "


r.se.1 <- sqrt(diag(vcovHC(plm.1f, type = "HC1")))
r.se.2 <- sqrt(diag(vcovHC(plm.2f, type = "HC1")))
r.se.3 <- sqrt(diag(vcovHC(plm.3f, type = "HC1")))
r.se.4 <- sqrt(diag(vcovHC(plm.4f, type = "HC1")))
r.se.5 <- sqrt(diag(vcovHC(plm.5f, type = "HC1")))
r.se.6 <- sqrt(diag(vcovHC(plm.6f, type = "HC1")))
r.se.7 <- sqrt(diag(vcovHC(plm.7f, type = "HC1")))
r.se.8 <- sqrt(diag(vcovHC(plm.8f, type = "HC1")))

stargazer(plm.1f, plm.2f, plm.3f, plm.4f, plm.5f, plm.6f, plm.7f, plm.8f, se=list(r.se.1, r.se.2, r.se.
```

```
##
```

```
## Road Paved and Incidence of Non-Communicable Disease: OLS with Village and Time-Fixed Effects
## ================================================================================
##                                    Dependent variable:
##               ---------------------------------------------------------------
##                                       mbNonComDisease
##                 (1)      (2)      (3)      (4)      (5)      (6)      (7)      (8)
## --------------------------------------------------------------------------------
## lnDistToPvdRd -0.068  -0.067  -0.048  -0.050  -0.059  -0.063  -0.058  -0.058
##               (0.047) (0.047) (0.042) (0.045) (0.049) (0.047) (0.042) (0.046)
##
## income                 0.150                                            0.102
##                       (0.097)                                          (0.133)
##
## literate                       0.570
##                               (0.437)
##
## smokeTobacco                            0.063                           0.031
##                                        (0.059)                         (0.082)
##
## ownToilet                                       0.002
##                                                (0.002)
##
## electricity                                             0.002
##                                                        (0.001)
##
## seenDoctor                                                      -0.003
##                                                                 (0.006)
##
## --------------------------------------------------------------------------------
## Observations    599      599      599      599      599      599      599      599
## R2             0.009   0.025   0.019   0.023   0.017   0.016   0.013   0.027
## Adjusted R2   -4.645  -4.604  -4.642  -4.617  -4.651  -4.656  -4.673  -4.649
## ================================================================================
## Note:                                          *p<0.1; **p<0.05; ***p<0.01
```

STD/HIV

Table 1: 2005 Summary Statistics: Household and Village Characteristics

Statistic	N	Mean	St. Dev.	Min	Max
roadPaved	1,266	0.660	0.474	0	1
income	1,266	0.417	0.265	0.005	2.387
literate	1,266	0.583	0.164	0.080	1.000
smokeTobacco	1,266	1.242	0.593	0.000	2.000
ownToilet	1,266	30.334	32.300	0.000	100.000
electricity	1,266	69.855	32.460	0.000	100.000
seenDoctor	1,266	100.000	0.000	100	100
ImmuniCamps	1,266	3.299	1.711	0	7
PctPipedWater	1,266	33.441	38.998	0.000	100.000
healthSubCntr	1,266	0.535	0.837	0	12
primaryHealthCntr	1,266	0.173	0.412	0	6
commHealthCntr	1,266	0.029	0.169	0	1
pvtClinic	1,266	0.494	1.336	0	20
pvtHospital	1,266	0.090	0.463	0	6
distanceToPavedRoad	430	4.965	5.858	1	50

Table 2: 2012 Summary Statistics: Household and Village Characteristics

Statistic	N	Mean	St. Dev.	Min	Max
roadPaved	1,278	0.868	0.339	0	1
income	1,278	0.997	0.694	0.117	7.545
literate	1,278	0.632	0.139	0.059	0.950
smokeTobacco	1,278	2.180	0.832	1.000	4.000
ownToilet	1,278	40.041	31.965	0.000	100.000
electricity	1,278	82.096	25.478	0.000	100.000
seenDoctor	1,278	94.514	9.924	28.571	100.000
ImmuniCamps	1,278	4.703	5.242	0	90
PctPipedWater	1,278	37.102	39.667	0.000	100.000
healthSubCntr	1,278	0.538	0.767	0	8
primaryHealthCntr	1,278	0.131	0.343	0	2
commHealthCntr	1,278	0.033	0.207	0	4
pvtClinic	1,278	0.688	1.601	0	20
pvtHospital	1,278	0.087	0.478	0	8
distanceToPavedRoad	169	3.858	4.662	1	35

roadPaved: 1 if village has a paved road, 0 if it does not.
distanceToPavedRoad: the distance to the nearest paved road in villages without a paved road.
income: average income of village in hundred thousand Indian Rupee units.
literate: percentage of villagers in a village that can read and write.
smokeTobacco: percentage of villagers in a village that smoke tobacco.
ownToilet: the percentage of villagers in a village that own a toilet with plumbing inside their home.
electricity: percentage of villagers in a village that have access to electricity.
seenDoctor: percentage of villagers in a village that have seen a doctor for minor illness in the past five years.
PctPipedWater: percentage of villagers in a village that have access to a piped water supply
and plumbing inside their home.
healthSubCntr, primaryHealthCntr, commHealthCntr, pvtClinic, pvtHospital:
number of those respective health facilities in each village.