bw_10_upd.R

datalab

2023-06-16

```
#############################################################
#################       Linear Polynomial        ############
################      October 10, 2018     ##################
####  Rerun: December 15, 2022 - checks out with the paper
# coefs and se validated in the replication file table1.R

rm(list=ls())
library(foreign)
library(plyr)
library(readstata13)
library(multiwayvcov)
library(sandwich)
library(lmtest)
library(stargazer)

data=read.csv("~/Dropbox/Personal Research 2017/replications/karn_nov16.csv")
names(data)
```

```
##   [1] "X.1"              "dist_name"        "vilname91"        "v1"               "dist_code"
##  [15] "phc_cntr"         "hc_cntr"          "fpc_cntr"         "tb_cntr"          "nh_cntr"
##  [29] "st_town"          "agri_land"        "near_town"        "circl_code"       "m_pop"
##  [43] "m_sc"             "f_sc"             "tot_st"           "m_st"             "f_st"
##  [57] "ngmf_char"        "ngmfprwmedd"      "ngmfprwothd"      "ngmfprwnod"       "ngmftrpr"
##  [71] "taptr"            "tapuntr"          "hp"               "covwell"          "uncovwell"
##  [85] "phs_cntr"         "prhsc"            "stname"           "stname1991"       "d_name"
##  [99] "all_hosp"         "area_na_cu"       "ayu_disp"         "ayu_hosp"         "canal_govt"
## [113] "ind_sch"          "lake"             "m_home"           "m_sch"            "nw_fac"
## [127] "power_oth"        "power_supl"       "p_sch"            "p_t_fac"          "rang_mcw"
## [141] "s_sch"            "s_s_sch"          "tot_exp"          "tot_inc"          "tr_sch"
## [155] "gov_ps_n"         "pr_ps_n"          "gov_ms_n"         "pr_ms_n"          "gov_secs_n"
## [169] "nviltmsna"        "nviltsecs"        "nviltsecsna"      "power"            "hplost"
## [183] "pnt_fac"          "power_ea"         "power_eag"        "power_edea"       "power_eo"
## [197] "medfac"           "rangmed"          "tot_hh"           "pucca_binary"     "kucha_binary"
## [211] "dist_fr_town"     "tbcl"             "tank"             "tap"              "X"
## [225] "M_POP"            "F_POP"            "TOT_L6"           "M_L6"             "F_L6"
## [239] "TOT_ILLT"         "M_ILLT"           "F_ILLT"           "TOT_W"            "M_W"
## [253] "F_AGLB"           "TOT_MFHH"         "M_MFHH"           "F_MFHH"           "TOT_OTH_W"
## [267] "M_MRG_AGLB"       "F_MRG_AGLB"       "T_MRG_HH"         "M_MRG_HH"         "F_MRG_HH"
## [281] "NEAR_DIST_border1" "NEAR_ANGLE"       "temp_av"          "wc2010mt_1"       "TerrainRug"
```

```
summary(data$Latitude)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   13.49   14.28   14.96   15.25   16.21   17.75     138
```

```r
summary(data$Longitude)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   74.12   75.26   75.89   75.90   76.48   77.67     138
```

```r
summary(data$border1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   0.000   1.000   0.599   1.000   1.000    5146
```

```r
summary(data$border2)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   0.000   0.000   1.000   0.569   1.000   1.000    6425
```

```r
summary(data$Slope)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    0.00   89.98   89.99   87.69   89.99   90.00     138
```

```r
########################
##### Distances #####

#Distance to Mysore-Bombay Border
rd10.mb=data[which(data$NEAR_DIST_border1<5000),]

table(rd10.mb$border1)
```

```
##
##   0   1
## 301 325
```

```r
#Distance to Hyderabad-Bombay Border
rd10.hb=data[which(data$NEAR_DIST_border2<5000),]

table(rd10.hb$border2)
```

```
##
##   0   1
## 260 318
```

```r
#Mysore
#outcome-health centers
health.mys=lm(health_binary~border1+TOT_POP+
        TOT_SC+TOT_ST+Slope+TerrainRug+Latitude+Longitude, data=rd10.mb) #OLS estimation
summary(health.mys)
```

```
##
## Call:
## lm(formula = health_binary ~ border1 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + Latitude + Longitude, data = rd10.mb)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.60225 -0.26334 -0.16729  0.02189  1.02026
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.414e+00  4.425e+00  -1.223  0.22163
```

```
## border1      4.499e-02  3.555e-02   1.265  0.20621
## TOT_POP      5.760e-05  2.100e-05   2.743  0.00626 **
## TOT_SC       1.441e-04  6.447e-05   2.235  0.02574 *
## TOT_ST       2.342e-04  7.843e-05   2.986  0.00294 **
## Slope        8.958e-04  1.416e-03   0.633  0.52711
## TerrainRug  -1.899e-02  9.173e-03  -2.071  0.03881 *
## Latitude    -2.469e-01  1.188e-01  -2.078  0.03813 *
## Longitude    1.199e-01  5.974e-02   2.008  0.04511 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4114 on 617 degrees of freedom
## Multiple R-squared:  0.08451,    Adjusted R-squared:  0.07264
## F-statistic: 7.119 on 8 and 617 DF,  p-value: 4.958e-09
```

```
health.mys.cl=cluster.vcov(health.mys, rd10.mb$dist_name)
health.mys.se=sqrt(diag(health.mys.cl)) #cluster standard errors

#outcome - paved roads
pucca.mys=lm(pucca_binary~border1+TOT_POP+
       TOT_SC+TOT_ST+Slope+TerrainRug+Latitude+Longitude, data=rd10.mb) #OLS estimation
summary(pucca.mys)
```

```
##
## Call:
## lm(formula = pucca_binary ~ border1 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + Latitude + Longitude, data = rd10.mb)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20850 -0.00392  0.14164  0.21281  0.36255
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.723e+00  3.793e+00   0.982  0.32663
## border1     -9.200e-02  3.047e-02  -3.019  0.00264 **
## TOT_POP      4.846e-05  1.800e-05   2.693  0.00728 **
## TOT_SC       1.596e-04  5.526e-05   2.888  0.00401 **
## TOT_ST       2.208e-04  6.722e-05   3.284  0.00108 **
## Slope        2.154e-03  1.213e-03   1.775  0.07636 .
## TerrainRug  -4.559e-03  7.862e-03  -0.580  0.56220
## Latitude    -1.005e-01  1.018e-01  -0.987  0.32396
## Longitude   -2.258e-02  5.120e-02  -0.441  0.65930
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3526 on 617 degrees of freedom
## Multiple R-squared:  0.09439,    Adjusted R-squared:  0.08264
## F-statistic: 8.038 on 8 and 617 DF,  p-value: 2.403e-10
```

```
pucca.mys.cl=cluster.vcov(pucca.mys, rd10.mb$dist_name)
pucca.mys.se=sqrt(diag(pucca.mys.cl)) #clustered standard errors
```

```r
#Hyderabad
#outcome-health centers
health.hyd=lm(health_binary~border2+TOT_POP+
        TOT_SC+TOT_ST+Slope+TerrainRug+Latitude+Longitude, data=rd10.hb) #ols estimation
summary(health.hyd)
```

```
##
## Call:
## lm(formula = health_binary ~ border2 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + Latitude + Longitude, data = rd10.hb)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5806 -0.2884 -0.1919  0.4365  0.9338
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.098e+01  8.896e+00   1.234 0.217779
## border2     -9.707e-02  3.774e-02  -2.572 0.010363 *
## TOT_POP      7.288e-05  2.059e-05   3.540 0.000434 ***
## TOT_SC       1.442e-04  6.048e-05   2.384 0.017466 *
## TOT_ST       1.105e-04  7.614e-05   1.452 0.147066
## Slope       -3.915e-04  1.232e-03  -0.318 0.750803
## TerrainRug   3.292e-02  2.319e-02   1.419 0.156413
## Latitude     8.465e-02  4.973e-02   1.702 0.089230 .
## Longitude   -1.604e-01  1.241e-01  -1.292 0.196734
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4268 on 569 degrees of freedom
## Multiple R-squared:  0.08258,    Adjusted R-squared:  0.06968
## F-statistic: 6.402 on 8 and 569 DF,  p-value: 5.574e-08
```

```r
health.hyd.cl=cluster.vcov(health.hyd, rd10.hb$dist_name)
health.hyd.se=sqrt(diag(health.hyd.cl)) #cluster standard errors

#outcome - paved roads
pucca.hyd=lm(pucca_binary~border2+TOT_POP+
        TOT_SC+TOT_ST+Slope+TerrainRug+Latitude+Longitude, data=rd10.hb) #ols estimation
summary(pucca.hyd)
```

```
##
## Call:
## lm(formula = pucca_binary ~ border2 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + Latitude + Longitude, data = rd10.hb)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05408  0.02461  0.10209  0.16243  0.32487
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.341e-01  6.601e+00  -0.081   0.9355
## border2     -1.256e-02  2.801e-02  -0.449   0.6539
```

```
## TOT_POP      6.818e-05  1.528e-05   4.463 9.77e-06 ***
## TOT_SC       7.693e-05  4.488e-05   1.714   0.0870 .
## TOT_ST       1.040e-04  5.650e-05   1.841   0.0661 .
## Slope        1.914e-04  9.143e-04   0.209   0.8343
## TerrainRug   2.347e-02  1.721e-02   1.364   0.1731
## Latitude     4.732e-02  3.690e-02   1.282   0.2002
## Longitude    6.019e-03  9.207e-02   0.065   0.9479
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3167 on 569 degrees of freedom
## Multiple R-squared:  0.0724, Adjusted R-squared:  0.05936
## F-statistic: 5.552 on 8 and 569 DF,  p-value: 8.824e-07
```

```
pucca.hyd.cl=cluster.vcov(pucca.hyd, rd10.hb$dist_name)
pucca.hyd.se=sqrt(diag(pucca.hyd.cl)) #clustered standard errors
```

```
stargazer(health.mys, pucca.mys, health.hyd, pucca.hyd, se=list(health.mys.se, pucca.mys.se, health.hyd
          omit=c("TOT_POP", "TOT_SC", "TOT_ST", "Slope", "TerrainRug", "Latitude", "Longitude"),
          dep.var.labels=c("Health Centers", "Paved Roads", "Health Centers", "Paved Roads"),
          covariate.labels = c("Indirect Rule (Mysore)", "Indirect Rule (Hyderabad)", "Constant"),
          add.lines = list(c("Controls","\\checkmark","\\checkmark","\\checkmark","\\checkmark")),
          omit.stat = c("rsq", "f", "adj.rsq", "ser"))
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harva
## % Date and time: Fri, Jun 16, 2023 - 15:04:20
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{4}{c}{\textit{Dependent variable:}} \\
## \cline{2-5}
## \\[-1.8ex] & Health Centers & Paved Roads & Health Centers & Paved Roads \\
## \\[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \\[-1.8ex]
##  Indirect Rule (Mysore) & 0.045$^{***}$ & $-$0.092$^{*}$ &  &  \\
##   & (0.008) & (0.048) &  &  \\
##   & & & & \\
##  Indirect Rule (Hyderabad) &  &  & $-$0.097$^{***}$ & $-$0.013 \\
##   &  &  & (0.012) & (0.061) \\
##   & & & & \\
##  Constant & $-$5.414 & 3.723 & 10.976 & $-$0.534 \\
##   & (6.881) & (6.487) & (8.069) & (12.294) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Controls & \checkmark & \checkmark & \checkmark & \checkmark \\
## Observations & 626 & 626 & 578 & 578 \\
## \hline
```

```
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```