bw_100_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
```

```
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
```

```
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
```

```
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
```

```
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
```

```
########################
##### Distances #####

#Distance to Mysore-Bombay Border
rd10.mb=data[which(data$NEAR_DIST_border1<50000),]

table(rd10.mb$border1)
```

```
##
##    0    1
## 1858 2635
```

```
#Distance to Hyderabad-Bombay Border
rd10.hb=data[which(data$NEAR_DIST_border2<50000),]

table(rd10.hb$border2)
```

```
##
##    0    1
## 1540 1905
```

```
#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
## -0.65419 -0.25583 -0.16036 -0.04489  0.95772
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -2.205e+00  1.323e+00  -1.666   0.0958 .
## border1     -1.765e-02  2.039e-02  -0.866   0.3867
## TOT_POP      7.122e-05  8.049e-06   8.848  < 2e-16 ***
## TOT_SC       1.422e-04  2.200e-05   6.464 1.13e-10 ***
## TOT_ST       2.789e-04  2.895e-05   9.635  < 2e-16 ***
## Slope        5.528e-04  4.990e-04   1.108   0.2680
## TerrainRug  -5.162e-03  3.137e-03  -1.646   0.0999 .
## Latitude     3.560e-02  2.701e-02   1.318   0.1875
## Longitude    2.291e-02  1.828e-02   1.253   0.2101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.406 on 4484 degrees of freedom
## Multiple R-squared:  0.0807, Adjusted R-squared:  0.07906
## F-statistic:  49.2 on 8 and 4484 DF,  p-value: < 2.2e-16
```

```
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
## -1.17209  0.00211  0.15076  0.22782  0.43162
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.064e+00  1.204e+00   0.884   0.3767
## border1     -1.126e-01  1.855e-02  -6.071 1.37e-09 ***
## TOT_POP      4.863e-05  7.321e-06   6.642 3.46e-11 ***
## TOT_SC       1.589e-04  2.001e-05   7.942 2.49e-15 ***
## TOT_ST       2.179e-04  2.633e-05   8.274  < 2e-16 ***
## Slope        8.480e-04  4.539e-04   1.868   0.0618 .
## TerrainRug  -5.792e-03  2.853e-03  -2.030   0.0424 *
## Latitude    -2.336e-02  2.457e-02  -0.951   0.3417
## Longitude   -6.193e-04  1.663e-02  -0.037   0.9703
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3693 on 4484 degrees of freedom
## Multiple R-squared:  0.08673,    Adjusted R-squared:  0.0851
## F-statistic: 53.23 on 8 and 4484 DF,  p-value: < 2.2e-16
```

```
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
## -0.8240 -0.3120 -0.1817  0.4333  1.0690
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.994e+00  2.370e+00   1.685  0.09204 .
## border2     -4.791e-02  2.353e-02  -2.036  0.04180 *
## TOT_POP      1.090e-04  8.147e-06  13.377  < 2e-16 ***
## TOT_SC       2.047e-04  2.258e-05   9.067  < 2e-16 ***
## TOT_ST       1.798e-04  2.897e-05   6.207 6.05e-10 ***
## Slope       -1.901e-04  4.904e-04  -0.388  0.69829
## TerrainRug  -3.766e-03  7.402e-03  -0.509  0.61096
## Latitude     4.412e-02  1.434e-02   3.077  0.00211 **
## Longitude   -6.116e-02  3.271e-02  -1.870  0.06156 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4274 on 3436 degrees of freedom
## Multiple R-squared:  0.1274, Adjusted R-squared:  0.1254
## F-statistic: 62.71 on 8 and 3436 DF,  p-value: < 2.2e-16
```

```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
## -1.07848  0.03561  0.11352  0.17240  0.33819
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.453e+00  1.824e+00   5.182 2.32e-07 ***
## border2      2.168e-02  1.811e-02   1.197    0.231
```

```
## TOT_POP      4.636e-05  6.270e-06   7.394 1.78e-13 ***
## TOT_SC       9.126e-05  1.738e-05   5.252 1.60e-07 ***
## TOT_ST       1.069e-04  2.230e-05   4.794 1.71e-06 ***
## Slope        8.488e-05  3.775e-04   0.225    0.822
## TerrainRug   5.349e-04  5.697e-03   0.094    0.925
## Latitude     7.645e-02  1.104e-02   6.926 5.13e-12 ***
## Longitude   -1.310e-01  2.517e-02  -5.203 2.08e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3289 on 3436 degrees of freedom
## Multiple R-squared:  0.06376,    Adjusted R-squared:  0.06158
## F-statistic: 29.25 on 8 and 3436 DF,  p-value: < 2.2e-16
```

```
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 - 15:04:45
## \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) & $-$0.018 & $-$0.113$^{***}$ &  &  \\
##   & (0.030) & (0.038) &  &  \\
##   & & & & \\
##  Indirect Rule (Hyderabad) &  &  & $-$0.048$^{**}$ & 0.022 \\
##   &  &  & (0.021) & (0.050) \\
##   & & & & \\
##  Constant & $-$2.205 & 1.064 & 3.994 & 9.453$^{**}$ \\
##   & (2.200) & (3.190) & (2.455) & (4.491) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Controls & \checkmark & \checkmark & \checkmark & \checkmark \\
## Observations & 4,493 & 4,493 & 3,445 & 3,445 \\
## \hline
```

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