bw_50_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<25000),]

table(rd10.mb$border1)
```

```
##
##    0    1
## 1196 1329
```

```r
#Distance to Hyderabad-Bombay Border
rd10.hb=data[which(data$NEAR_DIST_border2<25000),]

table(rd10.hb$border2)
```

```
##
##   0   1
## 866 983
```

```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
## -0.69297 -0.25658 -0.15203 -0.01871  0.99728
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.162e+00  2.019e+00  -3.052  0.00230 **
```

```
## border1     -2.379e-02  2.212e-02  -1.075  0.28240
## TOT_POP      7.435e-05  1.052e-05   7.069 2.01e-12 ***
## TOT_SC       1.376e-04  2.910e-05   4.730 2.37e-06 ***
## TOT_ST       3.116e-04  3.776e-05   8.252 2.48e-16 ***
## Slope        8.473e-04  6.921e-04   1.224  0.22100
## TerrainRug  -4.496e-03  4.310e-03  -1.043  0.29696
## Latitude     3.251e-02  4.532e-02   0.717  0.47328
## Longitude    7.568e-02  2.813e-02   2.690  0.00718 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4029 on 2516 degrees of freedom
## Multiple R-squared:  0.0997, Adjusted R-squared:  0.09684
## F-statistic: 34.83 on 8 and 2516 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
## -1.08733  0.00461  0.13875  0.21466  0.39055
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.865e+00  1.796e+00  -1.596    0.111
## border1     -1.190e-01  1.967e-02  -6.047 1.69e-09 ***
## TOT_POP      4.527e-05  9.353e-06   4.840 1.38e-06 ***
## TOT_SC       1.089e-04  2.588e-05   4.208 2.66e-05 ***
## TOT_ST       1.439e-04  3.358e-05   4.287 1.88e-05 ***
## Slope        9.077e-04  6.155e-04   1.475    0.140
## TerrainRug  -6.238e-03  3.832e-03  -1.628    0.104
## Latitude     4.171e-02  4.030e-02   1.035    0.301
## Longitude    3.966e-02  2.501e-02   1.585    0.113
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3583 on 2516 degrees of freedom
## Multiple R-squared:  0.08091,    Adjusted R-squared:  0.07799
## F-statistic: 27.69 on 8 and 2516 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
## -0.8378 -0.3077 -0.1913  0.4380  1.0027
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.196e+00  4.124e+00   1.745  0.08114 .
## border2     -6.136e-02  2.665e-02  -2.302  0.02143 *
## TOT_POP      9.989e-05  1.110e-05   9.002  < 2e-16 ***
## TOT_SC       1.930e-04  3.166e-05   6.095 1.33e-09 ***
## TOT_ST       2.000e-04  4.060e-05   4.926 9.15e-07 ***
## Slope       -4.159e-04  6.874e-04  -0.605  0.54521
## TerrainRug  -1.052e-02  1.138e-02  -0.924  0.35545
## Latitude     6.300e-02  2.386e-02   2.640  0.00836 **
## Longitude   -1.066e-01  5.741e-02  -1.856  0.06355 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4294 on 1840 degrees of freedom
## Multiple R-squared:  0.1146, Adjusted R-squared:  0.1108
## F-statistic: 29.78 on 8 and 1840 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
## -1.05908  0.04214  0.10383  0.15138  0.27763
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.311e+00  3.005e+00   2.433  0.01506 *
## border2      8.474e-03  1.942e-02   0.436  0.66262
```

```
## TOT_POP      4.568e-05  8.086e-06   5.650 1.86e-08 ***
## TOT_SC       9.377e-05  2.307e-05   4.064 5.02e-05 ***
## TOT_ST       9.138e-05  2.959e-05   3.088  0.00204 **
## Slope       -8.432e-04  5.009e-04  -1.683  0.09247 .
## TerrainRug   3.451e-03  8.293e-03   0.416  0.67735
## Latitude     4.084e-02  1.739e-02   2.349  0.01894 *
## Longitude   -9.388e-02  4.184e-02  -2.244  0.02495 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3129 on 1840 degrees of freedom
## Multiple R-squared:  0.04817,    Adjusted R-squared:  0.04403
## F-statistic: 11.64 on 8 and 1840 DF,  p-value: 2.93e-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 - 15:04:37
## \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) & $-$0.024 & $-$0.119$^{***}$ &  &  \\
##   & (0.028) & (0.036) &  &  \\
##   & & & & \\
##  Indirect Rule (Hyderabad) &  &  & $-$0.061$^{***}$ & 0.008 \\
##   &  &  & (0.022) & (0.061) \\
##   & & & & \\
##  Constant & $-$6.162$^{**}$ & $-$2.865 & 7.196$^{*}$ & 7.311 \\
##   & (2.876) & (2.147) & (4.344) & (4.494) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Controls & \checkmark & \checkmark & \checkmark & \checkmark \\
## Observations & 2,525 & 2,525 & 1,849 & 1,849 \\
## \hline
```

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