appendix_altspec_cubicpoly.R

datalab

2023-06-16

```
#############################################################
#########    Altenrtive Specification ############
##################   Cubic Polynomial        #############
###############      October 10, 2018    ##################
#######  Check rerun: December 16, 2022 #######

rm(list=ls())
library(foreign)
library(plyr)
library(multiwayvcov)
library(sandwich)
library(lmtest)
library(stargazer)
library(ggplot2)


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

```
##   [1] "X.1"              "dist_name"        "vilname91"        "v1"
##   [5] "dist_code"        "thsil_code"       "block_code"       "vill_code"
##   [9] "loc_code"         "area"             "h_cntr"           "mcw_cntr"
##  [13] "mh_cntr"          "cwc_cntr"         "phc_cntr"         "hc_cntr"
##  [17] "fpc_cntr"         "tb_cntr"          "nh_cntr"          "chw_cntr"
##  [21] "rp_cntr"          "smp_cntr"         "oth_cntr"         "well"
##  [25] "canal"            "loc_town"         "st_code"          "talu_code"
##  [29] "st_town"          "agri_land"        "near_town"        "circl_code"
##  [33] "m_pop"            "f_pop"            "village_name"     "st_name"
##  [37] "subdist_n"        "cdblock_c"        "cdblock_n"        "neastt_n"
##  [41] "neastt_d"         "tot_sc"           "m_sc"             "f_sc"
##  [45] "tot_st"           "m_st"             "f_st"             "com_hctr"
##  [49] "hosall"           "hosalt"           "dispens"          "vethosp"
##  [53] "mobclin"          "famwel"           "ngmf_op"          "ngmf_iop"
##  [57] "ngmf_char"        "ngmfprwmedd"      "ngmfprwothd"      "ngmfprwnod"
##  [61] "ngmftrpr"         "ngmfmedshop"      "ngmfoth"          "nonagrarea"
##  [65] "permpastr"        "fallland"         "curfalland"       "netarea"
##  [69] "roadall"          "fp"               "taptr"            "tapuntr"
##  [73] "hp"               "covwell"          "uncovwell"        "tw"
##  [77] "power_binary"     "mchhc"            "prhc"             "primsch"
##  [81] "midsch"           "secsch"           "sensecsch"        "disp"
##  [85] "phs_cntr"         "prhsc"            "stname"           "stname1991"
##  [89] "d_name"           "distname91"       "year"             "primary_binary"
##  [93] "middle_binary"    "totalhealth"      "health_binary"    "high_binary"
```

```
##  [97] "tap_binary"        "all_disp"          "all_hosp"          "area_na_cu"
## [101] "ayu_disp"          "ayu_hosp"          "canal_govt"        "canal_pvt"
## [105] "college"           "crsoc_fac"         "dist_town"         "edu_fac"
## [109] "fwc_cntr"          "handpump"          "hom_disp"          "hom_hosp"
## [113] "ind_sch"           "lake"              "m_home"            "m_sch"
## [117] "nw_fac"            "n_home"            "other"             "other_soc"
## [121] "oth_sch"           "phs_cnt"           "ph_cntr"           "power_agr"
## [125] "power_all"         "power_dom"         "power_oth"         "power_supl"
## [129] "p_sch"             "p_t_fac"           "rang_mcw"          "rang_m_sch"
## [133] "rang_nac"          "rang_nw"           "rang_oth"          "rang_phc"
## [137] "rang_p_sch"        "rang_spcl"         "sou_summ"          "sp_cl_fac"
## [141] "s_sch"             "s_s_sch"           "tot_exp"           "tot_inc"
## [145] "tr_sch"            "tubewell"          "un_disp"           "un_hosp"
## [149] "lost"              "st_c"              "dist_c"            "subdist_c"
## [153] "vill_c"            "gp_n"              "gov_ps_n"          "pr_ps_n"
## [157] "gov_ms_n"          "pr_ms_n"           "gov_secs_n"        "pr_secs_n"
## [161] "gov_sens_n"        "pr_sens_n"         "gov_oth_n"         "pr_oth_n"
## [165] "prim_hctr"         "prim_hsubctr"      "macwf"             "nviltms"
## [169] "nviltmsna"         "nviltsecs"         "nviltsecsna"       "power"
## [173] "hplost"            "lostmandi"         "lostdata2011"      "disp_cntr"
## [177] "eudu_inst"         "gr_coll"           "hand_pump"         "h_sch"
## [181] "oth_fac"           "ph_fac"            "pnt_fac"           "power_ea"
## [185] "power_eag"         "power_edea"        "power_eo"          "pu_coll"
## [189] "state_code"        "state_name"        "tube_well"         "thsil_name"
## [193] "vill_name"         "lost91"            "lost01"            "vn91"
## [197] "medfac"            "rangmed"           "tot_hh"            "pucca_binary"
## [201] "kucha_binary"      "tot_pop"           "uncult"            "tot_irr"
## [205] "cultwaste"         "tot_unir"          "land_forest"       "drink_wat"
## [209] "rang_water"        "river"             "dist_fr_town"      "tbcl"
## [213] "tank"              "tap"               "X"                 "VILLAGE_ID"
## [217] "NAME"              "SUB_DISTRI"        "DISTRICT"          "STATE_UT"
## [221] "C_CODE01"          "LEVEL"             "TOT_NM_HH"         "TOT_POP"
## [225] "M_POP"             "F_POP"             "TOT_L6"            "M_L6"
## [229] "F_L6"              "TOT_SC"            "M_SC"              "F_SC"
## [233] "TOT_ST"            "M_ST"              "F_ST"              "TOT_LIT"
## [237] "M_LIT"             "F_LIT"             "TOT_ILLT"          "M_ILLT"
## [241] "F_ILLT"            "TOT_W"             "M_W"               "F_W"
## [245] "TOT_MNW"           "M_MNW"             "F_MNW"             "TOT_CULT"
## [249] "M_CULT"            "F_CULT"            "TOT_AGLB"          "M_AGLB"
## [253] "F_AGLB"            "TOT_MFHH"          "M_MFHH"            "F_MFHH"
## [257] "TOT_OTH_W"         "M_OTH_W"           "F_OTH_W"           "TOT_MRW"
## [261] "M_MRW"             "F_MRW"             "T_MRG_CULT"        "M_MRG_CULT"
## [265] "F_MRG_CULT"        "T_MRG_AGLB"        "M_MRG_AGLB"        "F_MRG_AGLB"
## [269] "T_MRG_HH"          "M_MRG_HH"          "F_MRG_HH"          "T_MRG_OTH"
## [273] "M_MRG_OTH"         "F_MRG_OTH"         "TOT_NNW"           "M_NNW"
## [277] "F_NNW"             "Latitude"          "Longitude"         "NEAR_FID_border1"
## [281] "NEAR_DIST_border1" "NEAR_ANGLE"        "temp_av"           "wc2010mt_1"
## [285] "TerrainRug"        "Slope"             "border1"           "NEAR_FID_border2"
## [289] "NEAR_DIST_border2" "border2"
```

```
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<10000),] #20 km


table(rd10.mb$border1)
```

```
##
##   0   1
## 559 599
```

```r
#Distance to Hyderabad-Bombay Border
rd10.hb=data[which(data$NEAR_DIST_border2<10000),] #20 km

table(rd10.hb$border2)
```

```
##
##   0   1
## 447 493
```

```r
# mysore
rd10.mb$cubic=polym(rd10.mb$Latitude, rd10.mb$Longitude, degree=3, raw=TRUE) #cubic poly
summary(rd10.mb$cubic)
```

```
##       1.0              2.0              3.0             0.1              1.1
##  Min.   :13.93   Min.   :194.2   Min.   :2705   Min.   :74.54   Min.   :1040
##  1st Qu.:14.33   1st Qu.:205.3   1st Qu.:2942   1st Qu.:74.93   1st Qu.:1078
##  Median :14.46   Median :209.2   Median :3026   Median :75.14   Median :1088
##  Mean   :14.43   Mean   :208.4   Mean   :3009   Mean   :75.21   Mean   :1086
##  3rd Qu.:14.56   3rd Qu.:211.9   3rd Qu.:3084   3rd Qu.:75.49   3rd Qu.:1096
##  Max.   :14.78   Max.   :218.3   Max.   :3226   Max.   :75.90   Max.   :1120
##       2.1              0.2             1.2             0.3
##  Min.   :14511   Min.   :5556   Min.   :77614   Min.   :414100
##  1st Qu.:15441   1st Qu.:5615   1st Qu.:81076   1st Qu.:420760
##  Median :15739   Median :5645   Median :81825   Median :424177
##  Mean   :15673   Mean   :5656   Mean   :81650   Mean   :425424
##  3rd Qu.:15935   3rd Qu.:5698   3rd Qu.:82507   3rd Qu.:430128
```

```
##  Max.   :16543   Max.   :5761   Max.   :84933   Max.   :437250
```

```
health.mys=lm(health_binary~border1+TOT_POP+
        TOT_SC+TOT_ST+Slope+TerrainRug+rd10.mb$cubic, data=rd10.mb)
summary(health.mys)
```

```
##
## Call:
## lm(formula = health_binary ~ border1 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + rd10.mb$cubic, data = rd10.mb)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6373 -0.2528 -0.1535 -0.0333  0.9736
##
## Coefficients: (1 not defined because of singularities)
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -4.631e+04  6.951e+04  -0.666 0.505460
## border1          -1.375e-03  2.857e-02  -0.048 0.961613
## TOT_POP           4.952e-05  1.534e-05   3.228 0.001281 **
## TOT_SC            1.684e-04  4.515e-05   3.730 0.000201 ***
## TOT_ST            2.612e-04  5.870e-05   4.450 9.43e-06 ***
## Slope             7.825e-04  1.116e-03   0.701 0.483221
## TerrainRug       -1.001e-02  7.058e-03  -1.419 0.156269
## rd10.mb$cubic1.0  1.775e+03  4.257e+03   0.417 0.676865
## rd10.mb$cubic2.0  3.766e+01  1.117e+02   0.337 0.736113
## rd10.mb$cubic3.0  4.632e+00  2.532e+00   1.830 0.067551 .
## rd10.mb$cubic0.1  1.524e+03  2.027e+03   0.752 0.452123
## rd10.mb$cubic1.1 -6.184e+01  1.297e+02  -0.477 0.633640
## rd10.mb$cubic2.1 -3.171e+00  2.096e+00  -1.513 0.130522
## rd10.mb$cubic0.2 -1.455e+01  1.499e+01  -0.971 0.331829
## rd10.mb$cubic1.2  1.021e+00  1.038e+00   0.984 0.325359
## rd10.mb$cubic0.3         NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.405 on 1143 degrees of freedom
## Multiple R-squared:  0.08998,    Adjusted R-squared:  0.07884
## F-statistic: 8.073 on 14 and 1143 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


pucca.mys=lm(pucca_binary~border1+TOT_POP+
        TOT_SC+TOT_ST+Slope+TerrainRug+rd10.mb$cubic, data=rd10.mb)
summary(pucca.mys)
```

```
##
## Call:
## lm(formula = pucca_binary ~ border1 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + rd10.mb$cubic, data = rd10.mb)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -1.11777 -0.01317  0.12724  0.21085  0.36300
##
## Coefficients: (1 not defined because of singularities)
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -4.947e+04  5.935e+04  -0.834 0.404691
## border1          -1.256e-01  2.439e-02  -5.151 3.05e-07 ***
## TOT_POP           4.935e-05  1.310e-05   3.768 0.000173 ***
## TOT_SC            1.099e-04  3.855e-05   2.851 0.004431 **
## TOT_ST            1.562e-04  5.012e-05   3.117 0.001873 **
## Slope             1.765e-03  9.525e-04   1.853 0.064206 .
## TerrainRug       -3.111e-03  6.026e-03  -0.516 0.605798
## rd10.mb$cubic1.0  3.409e+03  3.635e+03   0.938 0.348487
## rd10.mb$cubic2.0 -1.851e+01  9.537e+01  -0.194 0.846164
## rd10.mb$cubic3.0  9.402e-01  2.161e+00   0.435 0.663645
## rd10.mb$cubic0.1  1.301e+03  1.730e+03   0.752 0.452135
## rd10.mb$cubic1.1 -8.381e+01  1.107e+02  -0.757 0.449295
## rd10.mb$cubic2.1 -2.744e-01  1.789e+00  -0.153 0.878109
## rd10.mb$cubic0.2 -8.998e+00  1.280e+01  -0.703 0.482090
## rd10.mb$cubic1.2  6.076e-01  8.863e-01   0.686 0.493166
## rd10.mb$cubic0.3         NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3457 on 1143 degrees of freedom
## Multiple R-squared:  0.1013, Adjusted R-squared:  0.0903
## F-statistic: 9.203 on 14 and 1143 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)) #cluster standard errors


#
# stargazer(mcl2, mcl3, digits=3,
#           column.labels = c("Health Facilities",
#                             "Pucca Roads"))
#


# hyderabad
rd10.hb$cubic=polym(rd10.hb$Latitude, rd10.hb$Longitude, degree=3, raw=TRUE) #cubic poly
summary(rd10.hb$cubic)
```

```
##       1.0             2.0             3.0            0.1             1.1
##  Min.   :15.08   Min.   :227.4   Min.   :3429   Min.   :75.69   Min.   :1144
##  1st Qu.:15.85   1st Qu.:251.2   1st Qu.:3980   1st Qu.:75.99   1st Qu.:1204
##  Median :16.07   Median :258.2   Median :4149   Median :76.21   Median :1225
##  Mean   :16.21   Mean   :262.9   Mean   :4271   Mean   :76.18   Mean   :1235
##  3rd Qu.:16.62   3rd Qu.:276.1   3rd Qu.:4587   3rd Qu.:76.37   3rd Qu.:1271
##  Max.   :17.33   Max.   :300.5   Max.   :5208   Max.   :76.55   Max.   :1319
##       2.1             0.2            1.2             0.3
##  Min.   :17256   Min.   :5729   Min.   : 86831   Min.   :433576
##  1st Qu.:19100   1st Qu.:5774   1st Qu.: 91409   1st Qu.:438729
##  Median :19683   Median :5808   Median : 93410   Median :442634
##  Mean   :20031   Mean   :5803   Mean   : 94053   Mean   :442046
##  3rd Qu.:21116   3rd Qu.:5832   3rd Qu.: 97126   3rd Qu.:445342
```

```
##  Max.   :22852   Max.   :5860   Max.   :100656   Max.   :448616
```

```
health.hyd=lm(health_binary~border2+TOT_POP+
        TOT_SC+TOT_ST+Slope+TerrainRug+rd10.hb$cubic, data=rd10.hb)
summary(health.hyd)
```

```
##
## Call:
## lm(formula = health_binary ~ border2 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + rd10.hb$cubic, data = rd10.hb)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6284 -0.2971 -0.1909  0.4557  0.9903
##
## Coefficients: (1 not defined because of singularities)
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -9.643e+04  1.025e+05  -0.941   0.3470
## border2          -9.267e-02  3.747e-02  -2.473   0.0136 *
## TOT_POP           7.848e-05  1.589e-05   4.938 9.37e-07 ***
## TOT_SC            1.998e-04  4.585e-05   4.358 1.46e-05 ***
## TOT_ST            1.466e-04  5.993e-05   2.447   0.0146 *
## Slope            -1.581e-04  1.076e-03  -0.147   0.8832
## TerrainRug        2.056e-02  1.754e-02   1.172   0.2415
## rd10.hb$cubic1.0  4.758e+03  6.095e+03   0.781   0.4352
## rd10.hb$cubic2.0  5.376e+01  3.470e+01   1.549   0.1216
## rd10.hb$cubic3.0  2.171e-01  1.628e-01   1.333   0.1828
## rd10.hb$cubic0.1  2.740e+03  2.753e+03   0.995   0.3198
## rd10.hb$cubic1.1 -1.485e+02  1.664e+02  -0.892   0.3726
## rd10.hb$cubic2.1 -8.452e-01  5.210e-01  -1.622   0.1051
## rd10.hb$cubic0.2 -1.952e+01  1.854e+01  -1.053   0.2927
## rd10.hb$cubic1.2  1.159e+00  1.146e+00   1.011   0.3121
## rd10.hb$cubic0.3         NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4302 on 925 degrees of freedom
## Multiple R-squared:  0.09629,    Adjusted R-squared:  0.08261
## F-statistic:  7.04 on 14 and 925 DF,  p-value: 5.936e-14
```

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

pucca.hyd=lm(pucca_binary~border2+TOT_POP+
        TOT_SC+TOT_ST+Slope+TerrainRug+rd10.hb$cubic, data=rd10.hb)
summary(pucca.hyd)
```

```
##
## Call:
## lm(formula = pucca_binary ~ border2 + TOT_POP + TOT_SC + TOT_ST +
##     Slope + TerrainRug + rd10.hb$cubic, data = rd10.hb)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05543  0.02724  0.10297  0.16244  0.33972
```

```
##
## Coefficients: (1 not defined because of singularities)
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -5.775e+04  7.623e+04  -0.758   0.4489
## border2          -9.310e-03  2.787e-02  -0.334   0.7384
## TOT_POP           5.580e-05  1.182e-05   4.720 2.73e-06 ***
## TOT_SC            7.705e-05  3.410e-05   2.259   0.0241 *
## TOT_ST            9.919e-05  4.457e-05   2.225   0.0263 *
## Slope            -3.241e-04  8.005e-04  -0.405   0.6856
## TerrainRug        2.249e-02  1.305e-02   1.724   0.0850 .
## rd10.hb$cubic1.0  3.865e+03  4.534e+03   0.852   0.3942
## rd10.hb$cubic2.0 -1.212e+01  2.581e+01  -0.470   0.6388
## rd10.hb$cubic3.0  1.085e-02  1.211e-01   0.090   0.9286
## rd10.hb$cubic0.1  1.474e+03  2.048e+03   0.720   0.4718
## rd10.hb$cubic1.1 -9.625e+01  1.238e+02  -0.778   0.4370
## rd10.hb$cubic2.1  1.526e-01  3.875e-01   0.394   0.6938
## rd10.hb$cubic0.2 -9.404e+00  1.379e+01  -0.682   0.4953
## rd10.hb$cubic1.2  5.988e-01  8.523e-01   0.703   0.4825
## rd10.hb$cubic0.3         NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.32 on 925 degrees of freedom
## Multiple R-squared:  0.0618, Adjusted R-squared:  0.0476
## F-statistic: 4.352 on 14 and 925 DF,  p-value: 1.501e-07
```

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



# stargazer(mcl2, mcl3, digits=3,
#           column.labels = c("Health Facilities",
#                             "Pucca Roads"))




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",
                 "cubic1.0", "cubic2.0", "cubic0.1", "cubic1.1",
                 "cubic2.1", "cubic0.2", "cubic1.2", "cubic0.3", "cubic3.0"),
          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 - 14:48:12
## \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) & $-$0.001 & $-$0.126$^{***}$ &  &  \\
##   & (0.011) & (0.035) &  &  \\
##   & & & & \\
##  Indirect Rule (Hyderabad) &  &  & $-$0.093$^{***}$ & $-$0.009 \\
##   &  &  & (0.021) & (0.073) \\
##   & & & & \\
##  Constant & $-$46,305.380 & $-$49,469.980 & $-$96,428.040 & $-$57,746.760 \\
##   & (47,332.330) & (40,972.960) & (60,642.290) & (53,923.800) \\
##   & & & & \\
## \hline \\[-1.8ex]
## Controls & \checkmark & \checkmark & \checkmark & \checkmark \\
## Observations & 1,158 & 1,158 & 940 & 940 \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{4}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```