Replication of a Research Claim from Fielding-Miller et al. (2020), from medRxiv

2020-11-22

Replication Team: *Kent Jason Cheng* and *Radoslaw Panczak*

Research Scientist: *Nick Fox*

Action Editor: *Gustav Nilsonne*

Project ID: Fielding-Miller_covid_R3pV - Cheng/Panczak - Data Analytic Replication - 615k

OSF project: https://osf.io/w7gdb/

Preregistration: https://osf.io/2kd4e

For stage 1 data collection when the dataset extended beyond the time of the preprint analyses was used, we found no significant effect of the `uninsured` variable. Table below reports full details of the regression model and number of observa tions used for the analysis.

At stage 1, the replication was **unsuccessful** according to SCORE criteria.

```
      . qui: spmatrix create contiguity W, replace

      . spregress deaths nonenglish farmwork uninsured poverty older pop_dens time_case1 time_case100,
gs2sls dvarlag(W)
        (2864 observations)
        (2864 observations (places) used)
        (weighting matrix defines 2864 places)

      Spatial autoregressive model                    Number of obs     =      2,864
      GS2SLS estimates                                Wald chi2(9)      =    1305.88
                                                      Prob > chi2       =     0.0000
                                                      Pseudo R2         =     0.2913

      ------------------------------------------------------------------------------
            deaths |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
      -------------+----------------------------------------------------------------
      deaths       |
        nonenglish |   2.014099    .354975     5.67   0.000     1.318361    2.709838
           farmwork |   .4451357   .2615213     1.70   0.089    -.0674367    .9577081
          uninsured |  -.1243106   .0962492    -1.29   0.197    -.3129555    .0643343
            poverty |    .605348   .1539303     3.93   0.000     .3036501    .9070459
              older |   .5243047   .2283545     2.30   0.022     .0767382    .9718713
           pop_dens |   .1850662   .0071575    25.86   0.000     .1710378    .1990946
         time_case1 |   .1397902   .0472536     2.96   0.003     .0471748    .2324056
      time_case100 |  -.1702398   .0260448    -6.54   0.000    -.2212867   -.1191929
             _cons |  -37.73724   8.030567    -4.70   0.000    -53.47686   -21.99762
      -------------+----------------------------------------------------------------
      W            |
            deaths |   .3099279   .0677016     4.58   0.000     .1772351    .4426207
      ------------------------------------------------------------------------------
      Wald test of spatial terms:          chi2(1) = 20.96      Prob > chi2 = 0.0000
```

This finding remained when sensitivity analysis using alternative definition of neighbourhoods (“rook”) was used.

```
      . qui: spmatrix create contiguity W, rook replace

      . spregress deaths nonenglish farmwork uninsured poverty older pop_dens time_case1 time_case100,
gs2sls dvarlag(W)
```

```
  (2864 observations)
  (2864 observations (places) used)
  (weighting matrix defines 2864 places)

Spatial autoregressive model                    Number of obs     =      2,864
GS2SLS estimates                                Wald chi2(9)      =    1317.60
                                                Prob > chi2       =     0.0000
                                                Pseudo R2         =     0.2918

------------------------------------------------------------------------------
      deaths |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
deaths       |
  nonenglish |   1.990345   .3537315     5.63   0.000     1.297044    2.683646
     farmwork |   .4550382   .2606638     1.75   0.081    -.0558536    .9659299
    uninsured |  -.1127116   .0961405    -1.17   0.241    -.3011434    .0757203
      poverty |   .6032849   .1533827     3.93   0.000     .3026603    .9039095
        older |   .5242044   .2273288     2.31   0.021     .0786482    .9697605
     pop_dens |   .1841953   .0071539    25.75   0.000     .1701738    .1982168
   time_case1 |   .1373079   .0470823     2.92   0.004     .0450284    .2295875
time_case100 |  -.1691834   .0259523    -6.52   0.000     -.220049   -.1183178
       _cons |   -37.8308   7.994161    -4.73   0.000    -53.49907   -22.16253
-------------+----------------------------------------------------------------
W            |
      deaths |   .3045639   .0629688     4.84   0.000     .1811474    .4279804
------------------------------------------------------------------------------
Wald test of spatial terms:          chi2(1) = 23.39      Prob > chi2 = 0.0000
```

Stage 2: analyses using “original” dataset

For stage 2 data collection when the dataset created suing specifications from the preprint was used, we found significant effect of the `uninsured` variable. The effect was observed in the same direction, but it was weaker and the p-value hig her than that reported in the preprint. Table below reports full details of the regression model and number of observations used for the analysis.

At stage 1, the replication was **unsuccessful** according to SCORE criteria.

```
. qui: spmatrix create contiguity W, replace

. spregress deaths nonenglish farmwork uninsured poverty older pop_dens time_case1 time_case100,
gs2sls dvarlag(W)
  (2590 observations)
  (2590 observations (places) used)
  (weighting matrix defines 2590 places)

Spatial autoregressive model                    Number of obs     =      2,590
GS2SLS estimates                                Wald chi2(9)      =    1024.49
                                                Prob > chi2       =     0.0000
                                                Pseudo R2         =     0.2473

------------------------------------------------------------------------------
      deaths |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
deaths       |
  nonenglish |   .3975335   .1174548     3.38   0.001     .1673262    .6277408
     farmwork |   .2062333   .0913142     2.26   0.024     .0272608    .3852059
    uninsured |  -.0679844   .0325508    -2.09   0.037    -.1317828    -.004186
      poverty |   .1494817   .0508166     2.94   0.003     .0498829    .2490804
        older |   .1731376   .0745076     2.32   0.020     .0271055    .3191697
     pop_dens |   .0511042   .0023071    22.15   0.000     .0465823    .0556261
   time_case1 |   .1359991   .0349891     3.89   0.000     .0674218    .2045765
time_case100 |  -.2391546   .0456659    -5.24   0.000    -.3286581   -.1496511
       _cons |  -11.31111   2.165447    -5.22   0.000     -15.5553   -7.066907
-------------+----------------------------------------------------------------
W            |
      deaths |   .4228964   .0700925     6.03   0.000     .2855175    .5602752
------------------------------------------------------------------------------
Wald test of spatial terms:          chi2(1) = 36.40      Prob > chi2 = 0.0000
```

Once again the analysis was not affected by alternative definition of neighbourhoods.

```
. qui: spmatrix create contiguity W, rook replace

. spregress deaths nonenglish farmwork uninsured poverty older pop_dens time_case1 time_case100,
gs2sls dvarlag(W)
  (2590 observations)
  (2590 observations (places) used)
  (weighting matrix defines 2590 places)

Spatial autoregressive model                    Number of obs     =      2,590
GS2SLS estimates                                Wald chi2(9)      =    1033.33
                                                Prob > chi2       =     0.0000
                                                Pseudo R2         =     0.2485

------------------------------------------------------------------------------
      deaths |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
deaths       |
  nonenglish |   .3971046   .1170367     3.39   0.001     .1677168    .6264924
     farmwork |   .2075147   .0910597     2.28   0.023     .0290409    .3859884
    uninsured |  -.0638184   .0325205    -1.96   0.050    -.1275574   -.0000795
      poverty |   .1499786    .050674     2.96   0.003     .0506594    .2492977
        older |   .1744984   .0742499     2.35   0.019     .0289713    .3200255
     pop_dens |   .0509472   .0023021    22.13   0.000     .0464352    .0554592
   time_case1 |   .1333811   .0348994     3.82   0.000     .0649795    .2017826
time_case100 |  -.2421952   .0455636    -5.32   0.000    -.3314982   -.1528921
        _cons |  -11.37807   2.158068    -5.27   0.000     -15.6078   -7.148333
-------------+----------------------------------------------------------------
W            |
      deaths |   .4298715   .0682701     6.30   0.000     .2960645    .5636785
------------------------------------------------------------------------------
Wald test of spatial terms:          chi2(1) = 39.65      Prob > chi2 = 0.0000
```

Deviations from preregistration

There were no deviations from preregistration during the analysis.

Description of materials provided.

The following materials are publicly available on the OSF project site :

- The raw datafile saved as Stata file: `merged_covid_usa_v2.dta`
- The analytic datafiles saved as Stata files: `merged_covid_usa_prepared_original.dta` and `merged_covid_usa_prepared_original.dta`
- The raw spatial datafile saved as shape file: `cb_2018_us_county_20m.zip`
- The analytic spatial datafile saved as shape file: `cb_2018_us_county_20m_prep.zip`
- The data prepraration files saved as literate programing markdown for Stata: `02_data-preparation-extended.stmd` and `04_data-preparation-original.stmd`
- The full data analysis script, provided as a Stata markdown document: `06_analysys-final-report.do` with the pdf output file being this report.

Citation

Fielding-Miller RK, Sundaram ME, Brouwer K (2020) Social determinants of COVID-19 mortality at the county level. *medRxiv* 2020.05.03.20089698; doi: 10.1101/2020.05.03.20089698