

Association Between District Racial Composition and Stopping Rate of Minorities

Links

- Fixed Effects in R
- Paper outlining the logic behind fixed effects and clustering
- What to do when you have time invariant predictors?
- Brief tutorial for Fixed Effects Regression
- Ideas for how to improve the interpretation of fixed effects models
- Another article helping to interpret Fixed Effects

Crime

- Violent crime encompasses murder, assault, rape, and robbery.
- Property crime encompasses burglary, larceny, motor vehicle theft, and arson.

Generate the demographic composition of each police unit in each month

- Calculate the total number of officers and the mean amount of work experience in each unit for each month.
- Calculate the number and percentage of Black, Hispanic, White, and Asian American/Pacific Islander officers for each unit for each month.
- Calculate the number and percentage of male and female officers for each unit for each month.

“The CPD currently subdivides Chicago into 22 policing districts which correspond to CPD units, in which the majority of police officers work. A typical district covers roughly ten square miles. There were 25 districts (numbered 1–25) until 2012, at which time 3 smaller districts were eliminated and merged with other districts. Districts 23 and 21 and District 13 were eliminated and absorbed into neighboring districts in March and December of 2012, respectively. While District 23 was mostly absorbed by District 19 and most of District 13 was absorbed by District 12, significant parts of District 21 were absorbed by Districts 1, 2, and 9. page 11 of Appendix Section S2.2 in Ba et al. 2021

- To ensure an apples-to-apples comparison, I will be only comparing the current 22 police districts using data from 2013 - 2015.

Read in census data

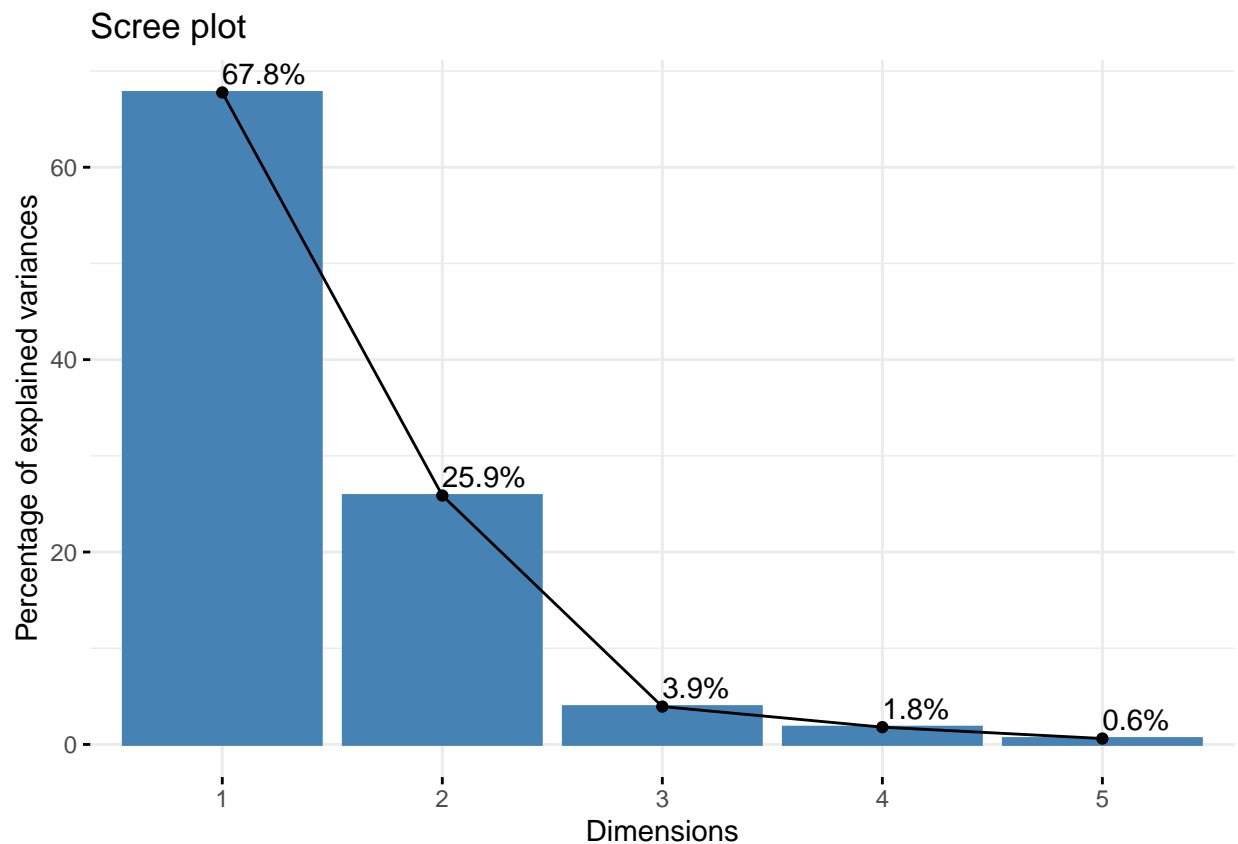
- Percentage of individuals in poverty
- Percentage of households which are single-parent families (male or female headed)
- Percentage of individuals 25 and older without a high school education
- Percentage of individuals who are foreign born
- Percentage of individuals who are not in the labor force
- I collect census tracts and use the centroid of each census tract to assign each census tract to a police district. I then aggregate the raw totals of each census tract to the police district level to come up with a summed total for the police district. I then calculate the percentage outcome variables.
- **ISSUE:** Because the period of time I am covering is only from 2013-2015, I use the 2011-2015 5-year ACS data. However, this means there is only **one** time period observation for each of these variables

which leads to issues of collinearity. More importantly, it defeats the purpose of fixed effects which is theoretically supposed to capture the effect of time-invariant variables (which in this case could technically be argued is the disadvantage of these neighborhoods).

- I also estimated a PCA model to deal with the underlying correlation in these variables, and it seems like these variables can be decomposed into two factors. However, this is a bit of a moot point at the moment because of the above issue.

```
## Reading layer `tl_2011_17_tract' from data source
##   `/home/joe/Documents/policing/clean-and-process/input/census-tract_shape'
##   using driver `ESRI Shapefile'
## Simple feature collection with 3123 features and 12 fields
## Geometry type: POLYGON
## Dimension:      XY
## Bounding box:   xmin: -91.51308 ymin: 36.9703 xmax: -87.01993 ymax: 42.50834
## Geodetic CRS:   NAD83

## Warning in st_centroid.sf(.): st_centroid assumes attributes are constant over
## geometries of x
```



Estimate models

```
##
## Call:
## MASS::glm.nb(formula = black_stops ~ black_diff + violent_cr_capita +
##   property_cr_capita + log_total_officers + unit + year + offset(log(black)),
##   data = final_vars, init.theta = 11.42523913, link = log)
##
```

```

## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -3.10477  -0.78190  -0.08716   0.52675   2.96800
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -8.82481    1.94444  -4.538 5.67e-06 ***
## black_diff      -3.77905    0.78396  -4.820 1.43e-06 ***
## violent_cr_capita -0.08168    0.05759  -1.418 0.156090
## property_cr_capita -0.05488    0.02346  -2.340 0.019298 *
## log_total_officers  0.99084    0.32636   3.036 0.002397 **
## unit10          -1.50172    0.31612  -4.751 2.03e-06 ***
## unit11          -2.84420    0.59420  -4.787 1.70e-06 ***
## unit12          -0.60469    0.17905  -3.377 0.000732 ***
## unit14          -0.77387    0.17018  -4.547 5.43e-06 ***
## unit15          -2.96667    0.63834  -4.647 3.36e-06 ***
## unit16           0.03460    0.19593   0.177 0.859844
## unit17          -0.74431    0.20554  -3.621 0.000293 ***
## unit18           0.50699    0.12387   4.093 4.26e-05 ***
## unit19          -0.13239    0.15616  -0.848 0.396536
## unit2           -0.50307    0.18140  -2.773 0.005549 **
## unit20          -0.42600    0.24016  -1.774 0.076098 .
## unit22          -1.79239    0.32891  -5.449 5.05e-08 ***
## unit24          -1.17855    0.24349  -4.840 1.30e-06 ***
## unit25          -1.80707    0.21380  -8.452 < 2e-16 ***
## unit3           -1.49161    0.36296  -4.110 3.96e-05 ***
## unit4           -2.21462    0.30760  -7.200 6.04e-13 ***
## unit5           -2.28580    0.38851  -5.883 4.02e-09 ***
## unit6           -2.60254    0.43725  -5.952 2.65e-09 ***
## unit7           -2.60754    0.56915  -4.581 4.62e-06 ***
## unit8           -1.72703    0.22521  -7.668 1.74e-14 ***
## unit9           -0.47694    0.18090  -2.637 0.008376 **
## year2014         -0.01656    0.02892  -0.573 0.566866
## year2015         -0.25432    0.03162  -8.042 8.85e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(11.4252) family taken to be 1)
##
##      Null deviance: 2301.5  on 791  degrees of freedom
## Residual deviance:  803.9  on 764  degrees of freedom
## AIC: 10521
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  11.425
##             Std. Err.:  0.586
##
## 2 x log-likelihood: -10462.638
##
## Call:
## MASS::glm.nb(formula = black_stops ~ black_diff + violent_cr_capita +

```

```

##      property_cr_capita + log_total_officers + disadvantage +
##      year + offset(log(black)), data = final_vars, init.theta = 4.814336582,
##      link = log)
##
## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -2.6498  -0.8726  -0.1759   0.5549   2.5883
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -9.02219    0.56502  -15.968  < 2e-16 ***
## black_diff       0.24649    0.15687   1.571  0.116111
## violent_cr_capita  0.23519    0.07002   3.359  0.000783 ***
## property_cr_capita -0.05238    0.01501  -3.489  0.000485 ***
## log_total_officers  0.96547    0.10287   9.386  < 2e-16 ***
## disadvantage    -1.61311    0.51613  -3.125  0.001776 **
## year2014         0.02820    0.04057   0.695  0.487083
## year2015        -0.21098    0.04106  -5.139  2.77e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(4.8143) family taken to be 1)
##
##      Null deviance: 987.66  on 791  degrees of freedom
## Residual deviance: 819.26  on 784  degrees of freedom
## AIC: 11180
##
## Number of Fisher Scoring iterations: 1
##
##              Theta:  4.814
##      Std. Err.:  0.237
##
## 2 x log-likelihood:  -11161.850

```

