Car Jacking MPLS - Tract

Ryan Larson

2024-03-15

Spatial Data

```
#MN tracts
tracts <- get_acs(geography = "tract",
                   state = "MN",
                 variables = "B01001_001E",
                 output = "wide",
                 survey = "acs5",
                 year = 2020,
                 geometry = T)
```

```
##   |                                                                              |
```

```
#Minneapolis Shapefile
mpls <- st_read("Data/mpls_city-shp/16cdbbfa-ad10-493c-afaf-52b61f2e76e42020329-1-180h9ap.whbo.shp") %>%
   st_transform(st_crs(tracts))
```

```
## Reading layer `16cdbbfa-ad10-493c-afaf-52b61f2e76e42020329-1-180h9ap.whbo' from data source `C:\User
##   using driver `ESRI Shapefile'
## Simple feature collection with 1 feature and 4 fields
## Geometry type: POLYGON
## Dimension:     XY
## Bounding box:  xmin: -93.32911 ymin: 44.89059 xmax: -93.19433 ymax: 45.05125
## Geodetic CRS:  WGS 84
```

```
mpls_tract <- tracts %>%
  st_filter(mpls, .predicate = st_intersects) %>%
  mutate(GEOID = as.numeric(GEOID),
         tract_area = as.numeric(st_area(.)),
         tract_area_sqkm = tract_area*.000001,
         tract_area_sqmi = tract_area_sqkm*.386102,
         intersection_area = as.numeric(st_area(st_intersection(., mpls))),
         perc_intersection = intersection_area/tract_area*100) %>%
  filter(perc_intersection >= 2) %>%
  select(-"B01001_001M")
```

ACS Covariates and Denominators

```
acs_17 <- get_acs(
  geography = "tract",
  variables = c("B01001_001E"),
  year = 2017,
```

```r
  state = "MN",
  county = "Hennepin",
  geometry = TRUE) %>%
  select(estimate)
```

```
## Getting data from the 2013-2017 5-year ACS

## Downloading feature geometry from the Census website.  To cache shapefiles for use in future sessions

##   |                                                                              |
```

```r
acs_18 <- get_acs(
  geography = "tract",
  variables = c("B01001_001E"),
  year = 2018,
  state = "MN",
  county = "Hennepin",
  geometry = TRUE
) %>%
select(estimate)
```

```
## Getting data from the 2014-2018 5-year ACS
## Downloading feature geometry from the Census website.  To cache shapefiles for use in future sessions

##   |                                                                              |
```

```r
acs_19 <- get_acs(
  geography = "tract",
  variables = c("B01001_001E"),
  year = 2019,
  state = "MN",
  county = "Hennepin",
  geometry = TRUE
) %>%
select(estimate)
```

```
## Getting data from the 2015-2019 5-year ACS
## Downloading feature geometry from the Census website.  To cache shapefiles for use in future sessions

##   |                                                                              |
```

```r
acs_20 <- get_acs(
  geography = "tract",
  variables = c("B01001_001E"),
  year = 2020,
  state = "MN",
  county = "Hennepin",
  geometry = TRUE
 )
```

```
## Getting data from the 2016-2020 5-year ACS
## Downloading feature geometry from the Census website.  To cache shapefiles for use in future sessions
```

```r
library(tigris)
```

```
## To enable caching of data, set `options(tigris_use_cache = TRUE)`
## in your R script or .Rprofile.
```

```r
hennepin_blocks <- blocks(
  "MN",
  "Hennepin",
  year = 2020
)
```

```
##   |                                                                              |
```

```r
#2017 ACS interpolation
acs_1720 <- interpolate_pw(
  from = acs_17,
  to = acs_20,
  to_id = "GEOID",
  weights = hennepin_blocks,
  weight_column = "POP20",
  crs = 26993,
  extensive = TRUE) %>%
  mutate(year = 2017)

#2018 ACS interpolation
acs_1820 <- interpolate_pw(
  from = acs_18,
  to = acs_20,
  to_id = "GEOID",
  weights = hennepin_blocks,
  weight_column = "POP20",
  crs = 26993,
  extensive = TRUE) %>%
  mutate(year = 2018)


#2019 ACS interpolation
acs_1920 <- interpolate_pw(
  from = acs_19,
  to = acs_20,
  to_id = "GEOID",
  weights = hennepin_blocks,
  weight_column = "POP20",
  crs = 26993,
  extensive = TRUE) %>%
  mutate(year = 2019)


pop_denoms <- acs_20 %>%
  st_transform(crs = 26993) %>%
  mutate(year = 2020) %>%
  select(-moe, -variable, -NAME) %>%
  rbind(acs_1720, acs_1820, acs_1920) %>%
  filter(GEOID %in% mpls_tract$GEOID)

#2021+2022: LOCF
pop_denom_21 <- pop_denoms %>%
  filter(year==2020) %>%
  select(GEOID, year, estimate) %>%
```

```
  mutate(year = 2021)

pop_denom_22 <- pop_denoms %>%
  filter(year==2020) %>%
  select(GEOID, year, estimate) %>%
  mutate(year = 2022)

pop_denom_locf <- pop_denoms %>%
  rbind(pop_denom_21, pop_denom_22) %>%
  rename(total_pop = estimate) %>%
  mutate(GEOID = as.numeric(GEOID)) %>%
  st_drop_geometry()

#ACS 2020 L-2 covariates
acs_2020 <- get_acs(geography = "tract",
            state = "MN",
               variables = c("B01001_001E", "B03002_003E", "B03002_004E", "B03002_005E",
                             "B03002_006E", "B03002_007E", "B03002_008E", "B03002_009E",
                             "B03002_010E", "B03002_011E", "B03002_012E", "B23025_002E",
                             "B23025_005E", "B17001_002E", "B19057_002E", "B11003_015E",
                             "B06009_002E", "B06009_005E", "C24010_001E", "C24010_003E",
                             "C24010_039E", "B11001_003E", "B01001_002E", "B05001_006E",
                             "B01001_003E", "B01001_004E", "B01001_005E", "B01001_006E",
                             "B01001_007E", "B01001_008E", "B01001_009E", "B01001_010E",
                             "B01001_011E", "B01001_012E", "B01001_013E", "B01001_014E",
                             "B01001_015E", "B01001_016E", "B01001_017E", "B01001_018E",
                             "B01001_019E", "B01001_020E", "B01001_021E", "B01001_022E",
                             "B01001_023E", "B01001_024E", "B01001_025E", "B01001_027E",
                             "B01001_028E", "B01001_029E", "B01001_030E", "B01001_031E",
                             "B01001_032E", "B01001_033E", "B01001_034E", "B01001_035E",
                             "B01001_036E", "B01001_037E", "B01001_038E", "B01001_039E",
                             "B01001_040E", "B01001_041E", "B01001_042E", "B01001_043E",
                             "B01001_044E", "B01001_045E", "B01001_046E", "B01001_047E",
                             "B01001_048E", "B01001_049E", "B07001_017E", "B25003_002E",
                             "B05002_013E", "B19013_001E"),
               output = "wide",
               survey = "acs5",
               year = 2020) %>%
  select(-ends_with("M", ignore.case = F)) %>%
  rename(total_pop = B01001_001E, white_pop = B03002_003E, black_pop = B03002_004E,
         na_pop = B03002_005E, asian_pop = B03002_006E, hpi_pop = B03002_007E,
         other_pop = B03002_008E, biracial_pop = B03002_009E, biracial_other_pop = B03002_010E,
         biracial_three_pop = B03002_011E, hisp_pop = B03002_012E, total_ilf = B23025_002E,
         unemp = B23025_005E, povlevel = B17001_002E, pub_assist = B19057_002E,
         female_hh = B11003_015E, no_hs_dip = B06009_002E, bach_degree = B06009_005E,
         total_employed = C24010_001E, employed_mbsa_male = C24010_003E,
         employed_mbsa_female = C24010_039E, mar_fam = B11001_003E, male = B01001_002E,
         noncitizen = B05001_006E,
        age_m_5_under = B01001_003E, age_m_5_9 = B01001_004E, age_m_10_14 = B01001_005E,
        age_m_15_17 = B01001_006E, age_m_18_19 = B01001_007E, age_m_20 = B01001_008E,
        age_m_21 = B01001_009E, age_m_22_24 = B01001_010E, age_m_25_29 = B01001_011E,
        age_m_30_34 = B01001_012E, age_m_35_39 = B01001_013E, age_m_40_44 = B01001_014E,
        age_m_45_49 = B01001_015E, age_m_50_54 = B01001_016E, age_m_55_59 = B01001_017E,
```

```
        age_m_60_61 = B01001_018E, age_m_62_64 = B01001_019E, age_m_65_66 = B01001_020E,
        age_m_67_69 = B01001_021E, age_m_70_74 = B01001_022E, age_m_75_79 = B01001_023E,
        age_m_80_84 = B01001_024E, age_m_85_plus = B01001_025E, age_f_5_under = B01001_027E,
        age_f_5_9 = B01001_028E, age_f_10_14 = B01001_029E, age_f_15_17 = B01001_030E,
        age_f_18_19 = B01001_031E, age_f_20 = B01001_032E, age_f_21 = B01001_033E,
        age_f_22_24 = B01001_034E, age_f_25_29 = B01001_035E, age_f_30_34 = B01001_036E,
        age_f_35_39 = B01001_037E, age_f_40_44 = B01001_038E, age_f_45_49 = B01001_039E,
        age_f_50_54 = B01001_040E, age_f_55_59 = B01001_041E, age_f_60_61 = B01001_042E,
        age_f_62_64 = B01001_043E, age_f_65_66 = B01001_044E, age_f_67_69 = B01001_045E,
        age_f_70_74 = B01001_046E, age_f_75_79 = B01001_047E, age_f_80_84 = B01001_048E,
        age_f_85_plus = B01001_049E, res_mob = B07001_017E,
        own_hh = B25003_002E, foreign = B05002_013E,
        med_hh_inc = B19013_001E) %>%
  mutate(white_prop = white_pop/total_pop,
         black_prop = black_pop/total_pop,
         na_prop = na_pop/total_pop,
         asian_prop = asian_pop/total_pop,
         hpi_prop = hpi_pop/total_pop,
         other_prop = other_pop/total_pop,
         biracial_prop = (biracial_pop+biracial_other_pop+biracial_three_pop)/total_pop,
         hisp_prop = hisp_pop/total_pop,
         white_perc = 100*white_pop/total_pop,
         black_perc = 100*black_pop/total_pop,
         na_perc = 100*na_pop/total_pop,
         asian_perc = 100*asian_pop/total_pop,
         hpi_perc = 100*hpi_pop/total_pop,
         other_perc = 100*other_pop/total_pop,
         biracial_perc = 100*(biracial_pop+biracial_other_pop+biracial_three_pop)/total_pop,
         hisp_perc = 100*hisp_pop/total_pop,
         unemp_rate = 100*unemp/total_ilf,
         pov_rate = 100*povlevel/total_pop,
         pub_assist_rate = 100*pub_assist/total_pop,
         female_hh_rate = 100*female_hh/total_pop,
         no_hs_dip_rate = 100*no_hs_dip/total_pop,
         bach_degree_rate = 100*bach_degree/total_pop,
         employed_mbsa = employed_mbsa_male+employed_mbsa_female,
         employed_mbsa_rate = 100*employed_mbsa/total_employed,
         mar_fam_rate = 100*mar_fam/total_pop,
         male_rate = 100*male/total_pop,
         noncitizen_rate = 100*noncitizen/total_pop,
         race_eth_hetero  = 1-(white_prop^2+black_prop^2+na_prop^2+asian_prop^2+
                       hpi_prop^2+other_prop^2+other_prop^2+biracial_prop^2+hisp_prop^2),
         age_below_18_perc = 100*(age_m_5_under+age_f_5_under+age_m_5_9+
                                  age_f_5_9+age_m_10_14+age_f_10_14+age_m_15_17+
                                  age_f_15_17)/total_pop,
         age_19_29_perc = 100*(age_m_18_19+age_f_18_19+age_m_20+age_f_20+age_m_21+age_f_21+
                          age_m_22_24+age_f_22_24+age_m_25_29+age_f_25_29)/total_pop,
         age_30_49_perc = 100*(age_m_30_34+age_f_30_34+age_m_35_39+age_f_35_39+
                               age_m_40_44+age_f_40_44+age_m_45_49+age_f_45_49)/total_pop,
         age_50_69_perc = 100*(age_m_50_54+age_f_50_54+age_m_55_59+age_f_55_59+
                               age_m_60_61+age_f_60_61+age_m_62_64+age_f_62_64+
                               age_m_65_66+age_f_65_66+age_m_67_69+age_f_67_69)/total_pop,
         age_70_plus_perc = 100*(age_m_70_74+age_f_70_74+age_m_75_79+age_f_75_79+
```

```
                                        age_m_80_84+age_f_80_84+age_m_85_plus+age_f_85_plus)/total_pop,
         res_mob_rate = 100-100*res_mob/total_pop,
         own_hh_rate = 100*own_hh/total_pop,
         foreign_rate = 100*foreign/total_pop)
```

```
## Getting data from the 2016-2020 5-year ACS
```

```
# 2020 Census 18+ Denominator
dc2020 <- get_decennial(
  geography = "tract",
  variables = c("P3_001N"),
  year = 2020,
  state = "MN",
  county = "Hennepin",
  geometry = F) %>%
  mutate(GEOID = is.numeric(GEOID)) %>%
  rename(total_pop = value) %>%
  select(-GEOID)
```

```
## Getting data from the 2020 decennial Census

## Using the PL 94-171 Redistricting Data Summary File

## Note: 2020 decennial Census data use differential privacy, a technique that
## introduces errors into data to preserve respondent confidentiality.
## i Small counts should be interpreted with caution.
## i See https://www.census.gov/library/fact-sheets/2021/protecting-the-confidentiality-of-the-2020-cens
## This message is displayed once per session.
```

Expanded MPLS Carjacking (Crime Incidents) Data

```
cj_exp <- read_csv("Car Jacking/MPDdata_2017to2022.csv") %>%
   mutate(date=mdy_hm(reporteddate),
         year=isoyear(date),
         week=isoweek(date)) %>%
  select(casenumber, year, week, latitude, longitude) %>%
  distinct(casenumber, .keep_all = TRUE) %>% #collapsing to incident-level
  drop_na(latitude, longitude) %>%
  st_as_sf(coords = c("longitude", "latitude"), crs = "NAD83", remove=F) %>%
  st_join(mpls_tract) %>% #spatial join neighborhoods
  st_drop_geometry() %>%
  drop_na(GEOID) %>%
  group_by(year, week, GEOID, .drop=F) %>%
  tally(name = "car_jack") %>%
  ungroup() %>%
  complete(year, week, GEOID=mpls_tract$GEOID, fill = list(car_jack = 0)) %>%
  filter(!(year==2021 & week==53)) %>%
  arrange(GEOID, year, week) %>%
  left_join(mpls_tract, by = "GEOID") %>%
  left_join(dc2020, by = c("NAME")) %>%
  mutate(car_jack_rate = car_jack/total_pop*1000) %>%
  st_as_sf()
```

```
## Rows: 3779 Columns: 28
## -- Column specification --------------------------------------------------------
```

```
## Delimiter: ","
## chr (23): casenumber, dataset, closurecode, closurecode_mpd, reporteddate, c...
## dbl  (4): precinct, latitude, longitude, age
## lgl  (1): dateofbirth
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

MPLS Carjackings by Week - MPD Extended Data

```r
#aggregate to week over tracts
cj_exp_week <- cj_exp %>%
  group_by(year, week) %>%
  summarize(car_jack = sum(car_jack, na.rm = T),
            total_pop = sum(total_pop, na.rm = T)) %>%
  mutate(begin_date = ISOweek2date(paste(year, paste0("W", sprintf("%02d", week)), 1,sep = "-")),
         end_date = begin_date+weeks(1)-days(1),
         car_jack_rate = car_jack/total_pop*1000,
         pre_post_floyd = ifelse(end_date <= as.Date("2020-05-25"), 0, 1)) %>%
  filter(end_date <= as.Date("2022-12-31")) %>%
  ungroup() %>%
  mutate(csma = forecast::ma(car_jack_rate, order=5, centre=TRUE),
         tsma = TTR::SMA(car_jack_rate, n=5))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
## Registered S3 method overwritten by 'quantmod': method from as.zoo.data.frame
## zoo
```

```r
pre_mean <- mean(cj_exp_week$car_jack_rate[cj_exp_week$pre_post_floyd==0], na.rm = T)
post_mean <- mean(cj_exp_week$car_jack_rate[cj_exp_week$pre_post_floyd==1], na.rm = T)

c(pre_mean, post_mean)
```

```
## [1] 0.004835884 0.030500004
```

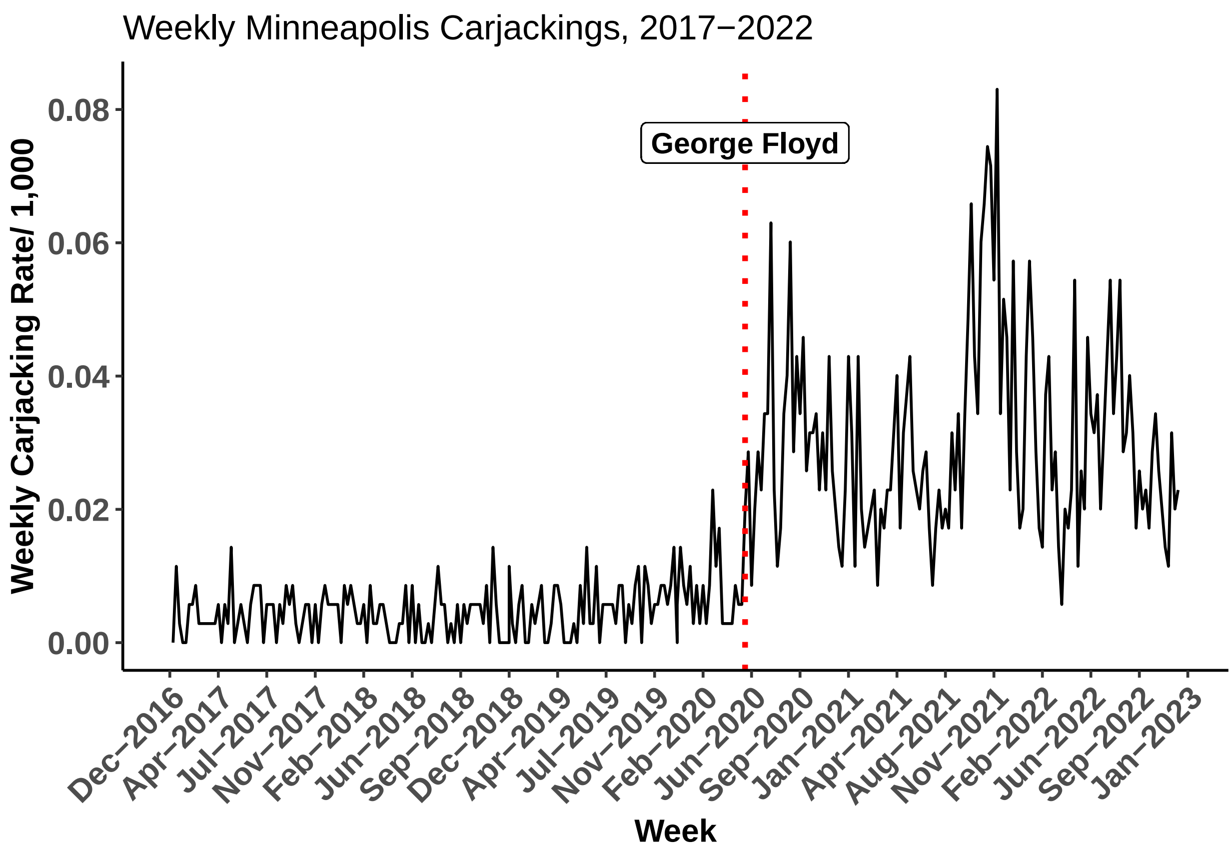
```r
post_mean/pre_mean
```

```
## [1] 6.307018
```

```r
ggplot(cj_exp_week)+
  geom_line(aes(x=begin_date, y=car_jack_rate))+
  scale_x_date(date_labels = "%b-%Y", date_breaks = "15 weeks",
               limits = c(min(cj_exp_week$begin_date), max(cj_exp_week$begin_date)))+
  geom_vline(xintercept=cj_exp_week$begin_date[cj_exp_week$year==2020 &
                                                 cj_exp_week$week==isoweek(date("2020-05-25"))],
              linetype="dotted", color="red", size=1)+
   geom_label(aes(x=cj_exp_week$begin_date[cj_exp_week$year==2020 &
                                             cj_exp_week$week==isoweek(date("2020-05-25"))],
                y=0.075),
             label = "George Floyd", show.legend = FALSE,
             fontface = "bold")+
  labs(title = "Weekly Minneapolis Carjackings, 2017-2022",
       x = "Week",
       y = "Weekly Carjacking Rate/ 1,000",
```

```r
      color = NULL)+
  #geom_line(aes(x=begin_date, y=csma, color = "CSMA(5)"))+
  theme_classic()+
  theme(axis.text = element_text(face = "bold", size = 12),
        axis.title.x = element_text(face = "bold", size = 12),
        axis.title.y = element_text(face = "bold", size = 12),,
        legend.key.size = unit(0.8, "cm"),
        legend.position = "bottom",
        strip.text.x = element_text(face = "bold", size = 12),
        axis.text.x=element_text(angle=45, hjust=1))
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## Warning: Use of `cj_exp_week$begin_date` is discouraged.
## i Use `begin_date` instead.

## Warning: Use of `cj_exp_week$year` is discouraged.
## i Use `year` instead.

## Warning: Use of `cj_exp_week$week` is discouraged.
## i Use `week` instead.
```




```r
ggsave(filename = "Car Jacking/Figures for publication/fig1.png", bg="white", width = 10, height = 8)
```

```
## Warning: Use of `cj_exp_week$begin_date` is discouraged.
## i Use `begin_date` instead.

## Warning: Use of `cj_exp_week$year` is discouraged.
## i Use `year` instead.

## Warning: Use of `cj_exp_week$week` is discouraged.
## i Use `week` instead.
```

Structural Change in Carjacking

```r
library(strucchange)
```

```
## Loading required package: sandwich

##
## Attaching package: 'strucchange'

## The following object is masked from 'package:stringr':
##
##     boundary
```

```r
cj_exp_week <- cj_exp_week %>%
  ungroup() %>%
  mutate(t = row_number())

sctest(cj_exp_week$car_jack_rate~cj_exp_week$t,
       type = "Chow",
       point = 180)
```

```
##
##  Chow test
##
## data:  cj_exp_week$car_jack_rate ~ cj_exp_week$t
## F = 51.174, p-value < 2.2e-16
```

```r
breakpoints(car_jack_rate~t,
            data = cj_exp_week,
            breaks = 1)
```

```
##
##   Optimal 2-segment partition:
##
## Call:
## breakpoints.formula(formula = car_jack_rate ~ t, breaks = 1,
##     data = cj_exp_week)
##
## Breakpoints at observation number:
## 180
##
## Corresponding to breakdates:
## 0.5714286
```

MPLS ZCTA Carjackings Map - MPD Extended Data

```r
#aggregate to neighborhood-year level
cj_exp_tract_year <- cj_exp %>%
  group_by(GEOID, year) %>%
  summarize(car_jack = sum(car_jack, na.rm = T),
            total_pop = sum(B01001_001E, na.rm = T),
            car_jack_rate = car_jack/total_pop*1000) %>%
  mutate(GEOID = as.character(GEOID))
```

```
## `summarise()` has grouped output by 'GEOID'. You can override using the
## `.groups` argument.
```

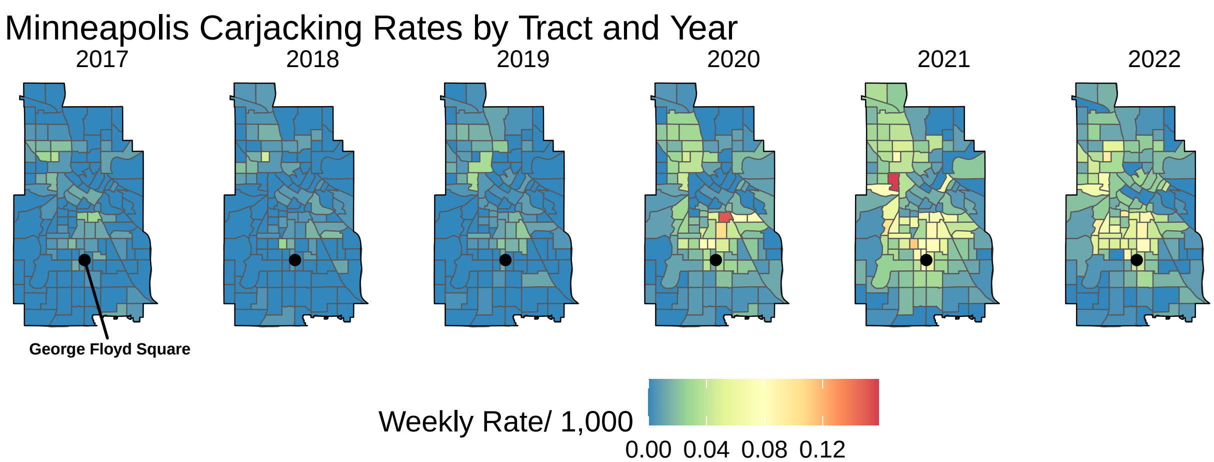
```r
#george floyd square
gfs_label <- geocode("George Floyd Square, Minneapolis", output = "latlon") %>%
  st_as_sf(coords = c("lon", "lat"), crs = "NAD83", remove=F) %>%
  mutate(name = "George Floyd Square") %>%
  mutate(year = 2017)
```

```
## i <https://maps.googleapis.com/maps/api/geocode/json?address=George+Floyd+Square,+Minneapolis&key=xx
```

```r
gfs <- geocode("George Floyd Square, Minneapolis", output = "latlon") %>%
  st_as_sf(coords = c("lon", "lat"), crs = "NAD83", remove=F) %>%
  mutate(name = "George Floyd Square")
```

```
## i <https://maps.googleapis.com/maps/api/geocode/json?address=George+Floyd+Square,+Minneapolis&key=xx
```

```r
ggplot() +
  geom_sf(data = cj_exp_tract_year, aes(geometry = geometry, fill = car_jack_rate)) +
  geom_sf(data = mpls, aes(geometry = geometry), color = "black", alpha = 0)+
  geom_sf(data = gfs, aes(geometry = geometry), color = "black")+
  geom_text_repel(data = gfs_label, aes(x=lon, y=lat, label = name),
                  size = 2,
                  fontface = "bold",
                  nudge_x = .1, nudge_y = -.06)+
  facet_grid(~year)+
  scale_fill_distiller(palette = "Spectral")+
  labs(title = "Minneapolis Carjacking Rates by Tract and Year",
       fill = "Weekly Rate/ 1,000")+
  theme_void()+
  theme(legend.position="bottom")
```




```r
ggsave(filename = "Car Jacking/Figures for publication/fig2.png", bg="white", width = 10, height = 8)
```

MPLS Murder (Crime Incidents) Data

```r
#pre-pims
mpd_2016 <- read_csv("Data/Police_Incidents_2016.csv")
```

```
## Rows: 20155 Columns: 20
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (12): PublicAddress, CCN, Precinct, ReportedDate, BeginDate, Offense, D...
## dbl   (7): FID, ControlNbr, GBSID, Lat, Long, X, Y
## time  (1): Time
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
mpd_2017 <- read_csv("Data/Police_Incidents_2017.csv")
```

```
## Rows: 22085 Columns: 20
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (12): PublicAddress, CCN, Precinct, ReportedDate, BeginDate, Offense, D...
## dbl   (7): FID, ControlNbr, GBSID, Lat, Long, X, Y
## time  (1): Time
##
```

```
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
mpd_2018a <- read_csv("Data/Police_Incidents_2018.csv")
```

```
## Rows: 7350 Columns: 20
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (12): PublicAddress, CCN, Precinct, ReportedDate, BeginDate, Offense, D...
## dbl   (7): FID, ControlNbr, GBSID, Lat, Long, X, Y
## time  (1): Time
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
#pims
mpd_2018b <- read_csv("Data/Police_Incidents_2018_PIMS.csv")
```

```
## Rows: 11603 Columns: 23
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (13): publicaddress, caseNumber, precinct, reportedDate, beginDate, repo...
## dbl (10): X, Y, reportedTime, beginTime, centergbsid, centerLong, centerLat,...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
mpd_2019 <- read_csv("Data/Police_Incidents_2019.csv")
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)

## Rows: 22934 Columns: 23
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (12): publicaddress, caseNumber, reportedDate, beginDate, reportedDateTi...
## dbl (11): X, Y, precinct, reportedTime, beginTime, centergbsid, centerLong, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
mpd_2020 <- read_csv("Data/Police_Incidents_2020.csv")
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)

## Rows: 24136 Columns: 23
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (12): publicaddress, caseNumber, reportedDate, beginDate, reportedDateTi...
## dbl (11): X, Y, precinct, reportedTime, beginTime, centergbsid, centerLong, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
mpd_2021 <- read_csv("Data/Police_Incidents_2021.csv")
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)

## Rows: 24755 Columns: 23
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (12): publicaddress, caseNumber, reportedDate, beginDate, reportedDateTi...
## dbl (11): X, Y, precinct, reportedTime, beginTime, centergbsid, centerLong, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
mpd_2022 <- read_csv("Data/Police_Incidents_2022.csv")
```

```
## Rows: 26526 Columns: 23
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (13): publicaddress, caseNumber, precinct, reportedDate, beginDate, repo...
## dbl  (9): X, Y, OBJECTID, reportedTime, beginTime, centerLong, centerLat, ce...
## lgl  (1): centergbsid
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
pre_pims_base <- mpd_2016 %>%
  rbind(mpd_2017) %>%
  rbind(mpd_2018a) %>%
  rename(reportedDate = ReportedDate,
         centerLong = Long,
         centerLat = Lat) %>%
  select(FID, centerLong, centerLat, Offense, reportedDate) %>%
  rename(OBJECTID = FID,
         X = centerLong,
         Y = centerLat,
         offense = Offense)

post_pims_base <- mpd_2018b %>%
  rbind(mpd_2019) %>%
  rbind(mpd_2020) %>%
  rbind(mpd_2021) %>%
  rbind(mpd_2022) %>%
  select(OBJECTID, X, Y, offense, reportedDate)

mpd <- pre_pims_base %>%
  rbind(post_pims_base)

#aggregate homicides to tract-week
homicide <- mpd %>%
  mutate(date=ymd_hms(reportedDate),
         year=isoyear(date),
         week=isoweek(date)) %>%
```

```r
  filter(offense=="MURDR" & year!=2016 & year!=2015) %>% #filter homicides
  select(OBJECTID, year, week, Y, X) %>%
  st_as_sf(coords = c("X", "Y"), crs = "NAD83", remove=F) %>%
  st_join(mpls_tract) %>% #spatial join neighborhoods
  st_drop_geometry() %>%
  filter(!is.na(GEOID)) %>%
  group_by(year, week, GEOID, .drop=F) %>%
  tally(name = "homicide") %>%
  ungroup() %>%
  complete(year, week, GEOID=mpls_tract$GEOID, fill = list(homicide = 0)) %>%
  filter(!(year==2021 & week==53)) %>%
  arrange(GEOID, year, week) %>%
  left_join(mpls_tract, by = "GEOID") %>%
  left_join(dc2020, by = c("NAME")) %>%
  mutate(homicide_rate = homicide/total_pop*1000) %>%
  st_as_sf()
```

MPLS Murder by Week

```r
#aggregate to week over tracts
homicide_week <- homicide %>%
  group_by(year, week) %>%
  summarize(homicide = sum(homicide, na.rm = T),
            total_pop = sum(total_pop, na.rm = T)) %>%
  mutate(begin_date = ISOweek2date(paste(year, paste0("W", sprintf("%02d", week)), 1,sep = "-")),
         end_date = begin_date+weeks(1)-days(1),
         homicide_rate = homicide/total_pop*1000,
         pre_post_floyd = ifelse(end_date <= as.Date("2020-05-25"), 0, 1)) %>%
  filter(end_date <= as.Date("2022-12-31")) %>%
  ungroup() %>%
  mutate(csma = forecast::ma(homicide_rate, order=5,centre=TRUE),
         tsma = TTR::SMA(homicide_rate, n=5))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

```r
pre_mean <- mean(homicide_week$homicide_rate[homicide_week$pre_post_floyd==0], na.rm = T)
post_mean <- mean(homicide_week$homicide_rate[homicide_week$pre_post_floyd==1], na.rm = T)

c(pre_mean, post_mean)
```

```
## [1] 0.001749826 0.004984354
```

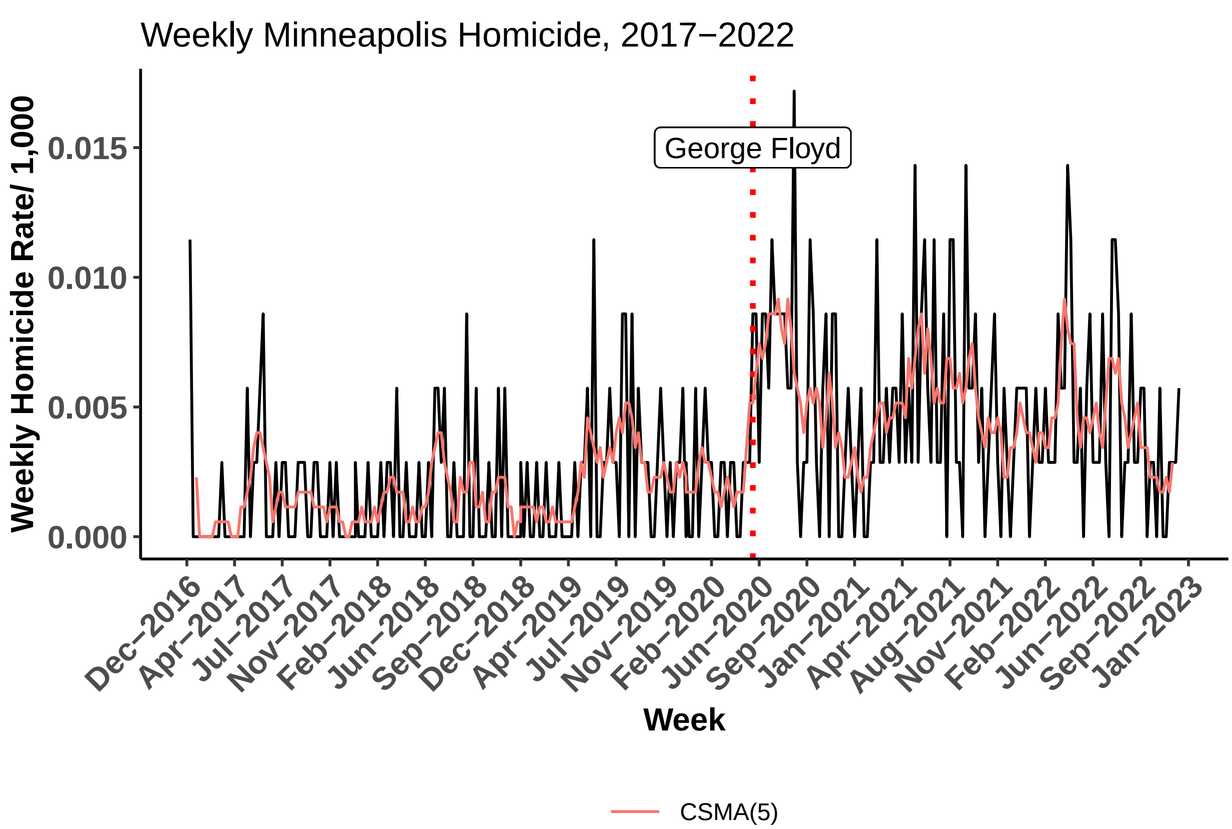
```r
post_mean/pre_mean
```

```
## [1] 2.848485
```

```r
ggplot(homicide_week)+
  geom_line(aes(x=begin_date, y=homicide_rate))+
 scale_x_date(date_labels = "%b-%Y", date_breaks = "15 weeks",
              limits = c(min(homicide_week$begin_date), max(homicide_week$begin_date)))+
    labs(title = "Weekly Minneapolis Homicide, 2017-2022",
      x = "Week",
      y = "Weekly Homicide Rate/ 1,000",
```

```
      color = NULL)+
  geom_vline(xintercept=homicide_week$begin_date[homicide_week$year==2020 &
                                                   homicide_week$week==isoweek(date("2020-05-25"))],
             linetype="dotted", color="red", size=1)+
   geom_label(aes(x=homicide_week$begin_date[homicide_week$year==2020 &
                                               homicide_week$week==isoweek(date("2020-05-25"))],
                 y=0.015),
             label = "George Floyd", show.legend = FALSE)+
  geom_line(aes(x=begin_date, y=csma, color = "CSMA(5)"))+     theme_classic()+
  theme(axis.text = element_text(face = "bold", size = 12),
        axis.title.x = element_text(face = "bold", size = 12),
        axis.title.y = element_text(face = "bold", size = 12),
        legend.key.size = unit(0.8, "cm"),
        legend.position = "bottom",
        strip.text.x = element_text(face = "bold", size = 12),
        axis.text.x=element_text(angle=45, hjust=1))
```

```
## Warning: Use of `homicide_week$begin_date` is discouraged.
## i Use `begin_date` instead.

## Warning: Use of `homicide_week$year` is discouraged.
## i Use `year` instead.

## Warning: Use of `homicide_week$week` is discouraged.
## i Use `week` instead.

## Warning: Removed 4 rows containing missing values (`geom_line()`).
```



```r
ggsave(filename = "Car Jacking/Figures for publication/fig3.png", bg="white", width = 10, height = 8)
```

```
## Warning: Use of `homicide_week$begin_date` is discouraged.
## i Use `begin_date` instead.

## Warning: Use of `homicide_week$year` is discouraged.
## i Use `year` instead.

## Warning: Use of `homicide_week$week` is discouraged.
## i Use `week` instead.

## Warning: Removed 4 rows containing missing values (`geom_line()`).
```

Structural Change in Carjacking

```r
homicide_week <- homicide_week %>%
  ungroup() %>%
  mutate(t = row_number())

sctest(homicide_week$homicide_rate~homicide_week$t,
       type = "Chow",
       point = 180)
```

```
##
##  Chow test
##
## data:  homicide_week$homicide_rate ~ homicide_week$t
## F = 16.772, p-value = 1.21e-07
```

```r
breakpoints(homicide_rate~t,
            data = homicide_week,
            breaks = 1)
```

```
##
##   Optimal 2-segment partition:
##
## Call:
## breakpoints.formula(formula = homicide_rate ~ t, breaks = 1,
##     data = homicide_week)
##
## Breakpoints at observation number:
## 180
##
## Corresponding to breakdates:
## 0.5714286
```

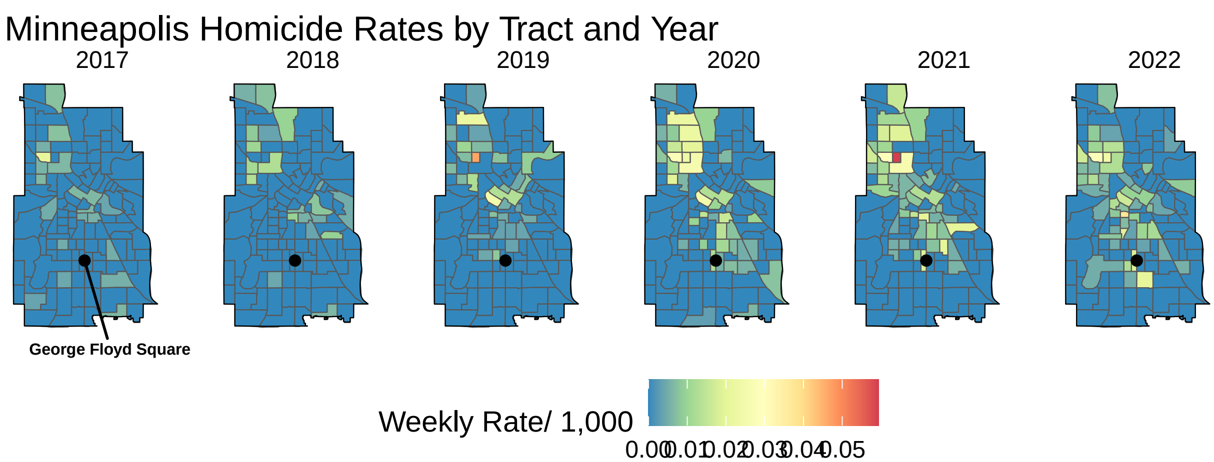
MPLS ZCTA Murder Map - MPD Extended Data

```r
#aggregate to neighborhood-year level
homicide_tract_year <- homicide %>%
  group_by(GEOID, year) %>%
  summarize(homicide = sum(homicide, na.rm = T),
            total_pop = sum(B01001_001E, na.rm = T),
```

```r
            homicide_rate = homicide/total_pop*1000) %>%
  mutate(GEOID = as.character(GEOID))
```

```
## `summarise()` has grouped output by 'GEOID'. You can override using the
## `.groups` argument.
```

```r
ggplot() +
  geom_sf(data = homicide_tract_year, aes(geometry = geometry, fill = homicide_rate)) +
  geom_sf(data = mpls, aes(geometry = geometry), color = "black", alpha = 0)+
  geom_sf(data = gfs, aes(geometry = geometry), color = "black")+
  geom_text_repel(data = gfs_label, aes(x=lon, y=lat, label = name),
                  size = 2,
                  fontface = "bold",
                  nudge_x = .1, nudge_y = -.06)+
  facet_grid(~year)+
  scale_fill_distiller(palette = "Spectral")+
  labs(title = "Minneapolis Homicide Rates by Tract and Year",
       fill = "Weekly Rate/ 1,000")+
  theme_void()+
  theme(legend.position="bottom")
```




```r
ggsave(filename = "Car Jacking/Figures for publication/fig4.png", bg="white", width = 10, height = 8)
```

Dispersion of Change from 2017-2019 to 2020-2021

Car Jacking

```
crimedispersion <- function
(data1, unitID, time1, time2, method = "match") {

  # define variables to limit build warnings
  adjusted <- Ut1 <- Ut2 <- Rt1 <- Rt2 <- chg <- pct <- NULL

  # ERROR CHECKING. Has user passed a data frame?
  if (!is.data.frame(data1)) {
    stop("The input data specified is not a data.frame object. Please fix.")
  }

  # Build a local data.frame and populate with passed arguments
  source_rows <- nrow(data1)
  df1 <- data.frame(matrix(ncol = 3, nrow = source_rows))
  colnames(df1) <- c("unit", "time1", "time2")
  df1$unit <- data1[, unitID]
  df1$time1 <- data1[, time1]
  df1$time2 <- data1[, time2]
  if (method == "remove") {
    analysisMethod <- "remove"
  } else {
    analysisMethod <- "match"
  }


  # ERROR CHECKING. Did user pass numeric columns where needed?
  try (df1$time1 <- as.numeric(df1$time1), silent = TRUE)
  try (df1$time2 <- as.numeric(df1$time2), silent = TRUE)

  if (!class(df1$time1)[1] == "numeric") {
    stop("The time1 field is not a numeric object. Please fix.")
  }
  if (!class(df1$time2)[1] == "numeric") {
    stop("The time2 field is not a numeric object. Please fix.")
  }

  # MORE ERROR CHECKING:
  # What if the user has NA or missing data?
  # What if the crime problem is decreasing?
  # Fun tasks for later...


  # Set up parameters ------------------------------------------------------

  # Set up initial parameters
  count_Rt1 <- sum(df1$time1)
  count_Rt2 <- sum(df1$time2)
  chg_Rt1_Rt2 <- count_Rt2 - count_Rt1
  pct_Rt1_Rt2 <- (chg_Rt1_Rt2 / count_Rt1) *100
```

```r
# Add the field that has the volume of change, and order by it
df1 <- df1 %>%
  mutate (diff = time2 - time1) %>%
  mutate (diffPct = 100*(diff/time1)) %>%
  arrange(desc(diff))

# Grab some basic statistics here
numPositive <- length(which(df1$diff > 0))
numNeutral <- length(which(df1$diff == 0))
numNegative <- length(which(df1$diff < 0))


# Create the new data frame to hold the result
df2 <- data.frame(matrix(ncol =8, nrow = 0))
colnames(df2) <- c("unit", "adjusted", "Ut1", "Ut2", "Rt1", "Rt2", "chg", "pct")
df2 <- df2 %>%
  mutate(unit = as.character(unit)) %>%
  mutate(adjusted = as.numeric(adjusted)) %>%
  mutate(Ut1 = as.numeric(Ut1)) %>%
  mutate(Ut2 = as.numeric(Ut2)) %>%
  mutate(Rt1 = as.numeric(Rt1)) %>%
  mutate(Rt2 = as.numeric(Rt2)) %>%
  mutate(chg = as.numeric(chg)) %>%
  mutate(pct = as.numeric(pct))

# set up the initial row in the result data frame
df2 <- df2 %>% add_row(unit = "[ ALL AREAS ]", adjusted = 0,
                       Ut1 = 0, Ut2 = 0,
                       Rt1 = count_Rt1, Rt2 = count_Rt2,
                       chg = chg_Rt1_Rt2, pct = pct_Rt1_Rt2)

gain_from_row_removal <- row_to_remove <- NULL


# Loop through each row of the data
for (master_loop in 1:(source_rows)){

  df1 <- df1 %>%  # order the data frame
    arrange(desc(diff))

  if (analysisMethod == "match"){
    #### 'Zero change the row' approach
    count_Rt1_temp <- count_Rt1
    count_Rt2_temp <- count_Rt2 - df1$diff[master_loop]
    pct_Rt1_Rt2 <- ((count_Rt1_temp - count_Rt2_temp) / count_Rt1) *100
  }
  else { #analysisMethod == "remove"
    #### 'Remove entire row' approach, including remove t1 value
    count_Rt1_temp <- count_Rt1 - df1$time1[master_loop]
    count_Rt2_temp <- count_Rt2 - df1$time2[master_loop]
    pct_Rt1_Rt2 <- ((count_Rt1_temp - count_Rt2_temp) / count_Rt1) *100
  }
```

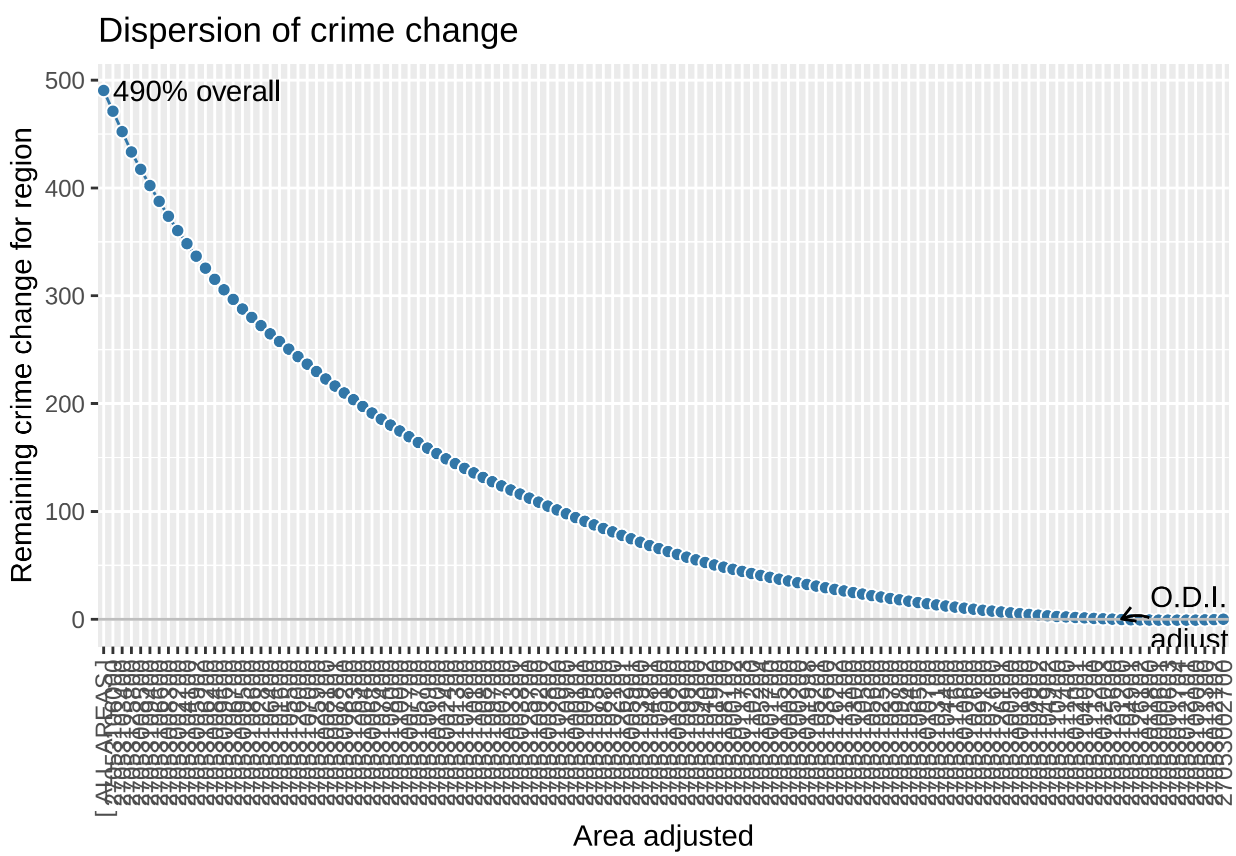
```
  row_to_remove <- 1  # Always row 1, but this is a legacy from
  # when I used a different approach...
  # Here, the row we are removing is
  # stored in row_to_remove

  if (analysisMethod == "remove"){
    #### Remove entire row approach
    #    This approach removes the impact of the area by subtracting
    #    both Rt1 and Rt2
    count_Rt1 <- count_Rt1 - df1$time1[row_to_remove]
    count_Rt2 <- count_Rt2 - df1$time2[row_to_remove]
    chg_Rt1_Rt2 <- count_Rt2 - count_Rt1
    pct_Rt1_Rt2 <- (chg_Rt1_Rt2 / count_Rt1) *100
    named_areas <- df1$unit[row_to_remove]
  }

  if (analysisMethod == "match"){
    #### Zero change the row approach, as if Rt2 == Rt1 in the row
    #    The best row to remove is has been exhaustively calculated
    #    Here, the row we are removing is stored in row_to_remove
    count_Rt1 <- count_Rt1
    count_Rt2 <- count_Rt2 - df1$diff[row_to_remove]
    chg_Rt1_Rt2 <- count_Rt2 - count_Rt1
    pct_Rt1_Rt2 <- (chg_Rt1_Rt2 / count_Rt1) *100
    named_areas <- df1$unit[row_to_remove]
  }

  # Add result to the output data frame
  df2 <- df2 %>% add_row(unit = named_areas, adjusted = master_loop,
                         Ut1 = df1$time1[row_to_remove], Ut2 = df1$time2[row_to_remove],
                         Rt1 = count_Rt1, Rt2 = count_Rt2,
                         chg = chg_Rt1_Rt2, pct = pct_Rt1_Rt2)

  # Adjust the row we just used in one of two ways:
  # 1. remove the actual row entirely
  if (analysisMethod == "remove"){
    df1 <-df1[-c(row_to_remove), ]
  }
  #2. adjust the Rt2 to match Rt1 resulting in a zero diff
  #   but show that diff as < lowest diff in the data set so that
  #   the program does not stall with too many zeros
  if (analysisMethod == "match"){
    df1$time2[row_to_remove] <- df1$time1[row_to_remove]
    df1$diff[row_to_remove] <- -999 # this should be changed to always less than
    # the lowest diff score in the data set
  }

} # end master_loop


# Calculate ODI and NCDI indices ----------------------------------------
NumContributed <- length(which(df2$chg > 0))
ODI <- NumContributed / source_rows
```

```r
  NCDI <- (numPositive - NumContributed) / source_rows
  ODI.text <- paste("O.D.I. = ", format(ODI, digits = 3), "after \nadjusting",
                    NumContributed, "of the", source_rows, "units")

  # Tidy up names for data frame ------------------------------------------

  df2 <- df2 %>%
    rename(unit_t1 = Ut1, unit_t2 = Ut2, region_t1 = Rt1, region_t2 = Rt2)

  # Plot ------------------------------------------------------------------

  df3 <- df2
  plot.adjustment <- ""
  if (nrow(df3) > 151)  {
    df3 <- df3[1:151, ]
    plot.adjustment <- "Plot only shows first\n100 areas adjusted"
  }


  p <- ggplot(df3, aes(x=reorder(unit, adjusted), y=pct, group = 1)) +
    geom_line(color="#3277a8") +
    geom_point(shape=21, color="white", fill="#3277a8", size=2) +
    geom_hline(color="grey", yintercept=0) +
    labs(title="Dispersion of crime change",
         x ="Area adjusted", y = "Remaining crime change for region") +
    annotate(
      geom = "curve", x = NumContributed+4, y = 1.5,
      xend = NumContributed+1, yend = 0.2,
      curvature = .2, arrow = arrow(length = unit(2, "mm"))
    ) +
    annotate(geom = "text", x = NumContributed+4.1, y = 1.5,
             label = ODI.text, hjust = "left") +
    theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +

    annotate(geom = "text", x = 2, y = df2$pct[1],
             label = paste0(format(df2$pct[1], digits = 3),"% overall"), hjust = "left") +
    theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))

  if (plot.adjustment != "") {
    p <- p +
      annotate(geom = "text", x = 100, y = df3$pct[1]-1, label = plot.adjustment, hjust = "right")
  }

  p


  # Create return list ----------------------------------------------------

  output <- list(df2, p, NumContributed, ODI, NCDI)
  return(output)
}
```

```r
prepost_cj <- cj_exp %>%
   mutate(begin_date = ISOweek2date(paste(year, paste0("W", sprintf("%02d", week)), 1,sep = "-")),
         end_date = begin_date+weeks(1)-days(1),
         pre_post_floyd = ifelse(end_date <= as.Date("2020-05-25"), 0, 1)) %>%
  filter(end_date <= as.Date("2022-12-31")) %>%
  group_by(GEOID, pre_post_floyd) %>%
  summarize(car_jack = sum(car_jack, na.rm = T),
            total_pop = sum(total_pop, na.rm = T)) %>%
  mutate(car_jack_rate = car_jack/total_pop*1000) %>%
  select(GEOID, pre_post_floyd, car_jack, car_jack_rate) %>%
  st_drop_geometry() %>%
  pivot_wider(names_from = pre_post_floyd, values_from = c(car_jack, car_jack_rate)) %>%
  mutate(GEOID = as.character(GEOID))
```

```
## `summarise()` has grouped output by 'GEOID'. You can override using the
## `.groups` argument.
```

```r
output <- crimedispersion(as.data.frame(prepost_cj), 'GEOID', 'car_jack_rate_0', 'car_jack_rate_1')

ouput_data <- output[[1]]
n_remove <- output[[3]]
odi <- output[[4]] #ratio of n removed to n overall
ncdi <- output[[5]] #ratio of areas not contributing to overall increase but still increase to overall

output[[2]]
```



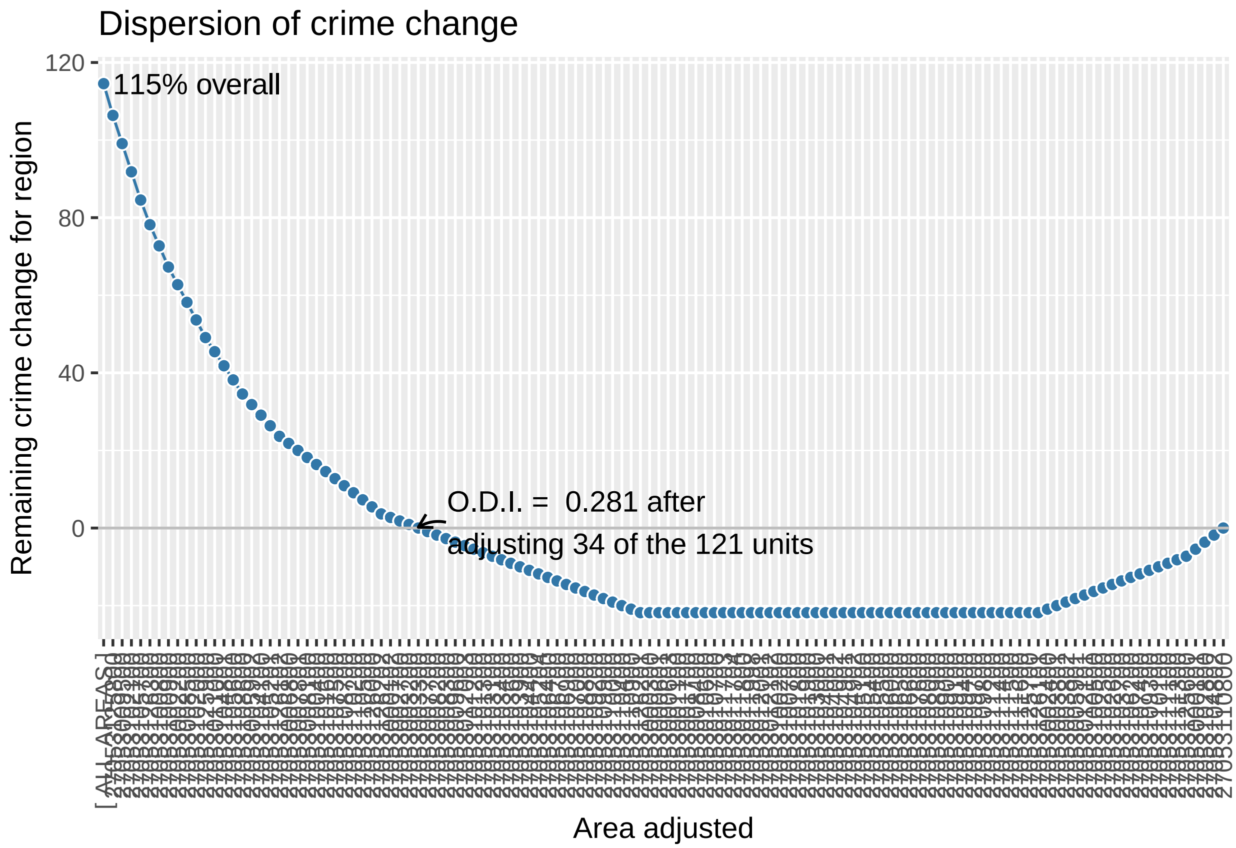
```
ggsave(filename = "Car Jacking/Figures for publication/fig5.png", bg="white", width = 10, height = 8)
```

Homicide

```
prepost_hom <- homicide %>%
 mutate(begin_date = ISOweek2date(paste(year, paste0("W", sprintf("%02d", week)), 1,sep = "-")),
        end_date = begin_date+weeks(1)-days(1),
        pre_post_floyd = ifelse(end_date <= as.Date("2020-05-25"), 0, 1)) %>%
  group_by(GEOID, pre_post_floyd) %>%
  summarize(homicide = sum(homicide, na.rm = T),
            total_pop = sum(total_pop, na.rm = T)) %>%
  mutate(homicide_rate = homicide/total_pop*1000) %>%
  st_drop_geometry() %>%
  select(GEOID, pre_post_floyd, homicide, homicide_rate) %>%
  pivot_wider(names_from = pre_post_floyd, values_from = c(homicide, homicide_rate)) %>%
  mutate(GEOID = as.character(GEOID))
```

```
## `summarise()` has grouped output by 'GEOID'. You can override using the
## `.groups` argument.
```

```
output_homicide <- crimedispersion(as.data.frame(prepost_hom), 'GEOID', 'homicide_0', 'homicide_1')

ouput_data <- output_homicide[[1]]
n_remove <- output_homicide[[3]]
odi <- output_homicide[[4]] #ratio of n removed to n overall
ncdi <- output_homicide[[5]] #ratio of areas not contributing to overall increase but still increase to

output_homicide[[2]]
```




```
ggsave(filename = "Car Jacking/Figures for publication/fig6.png", bg="white", width = 10, height = 8)
```

Spatial Correlation *Change* in Carjackings and Homicide

Carjacking

```
set.seed(7188)
library(sfdep)
library(spdep)
```

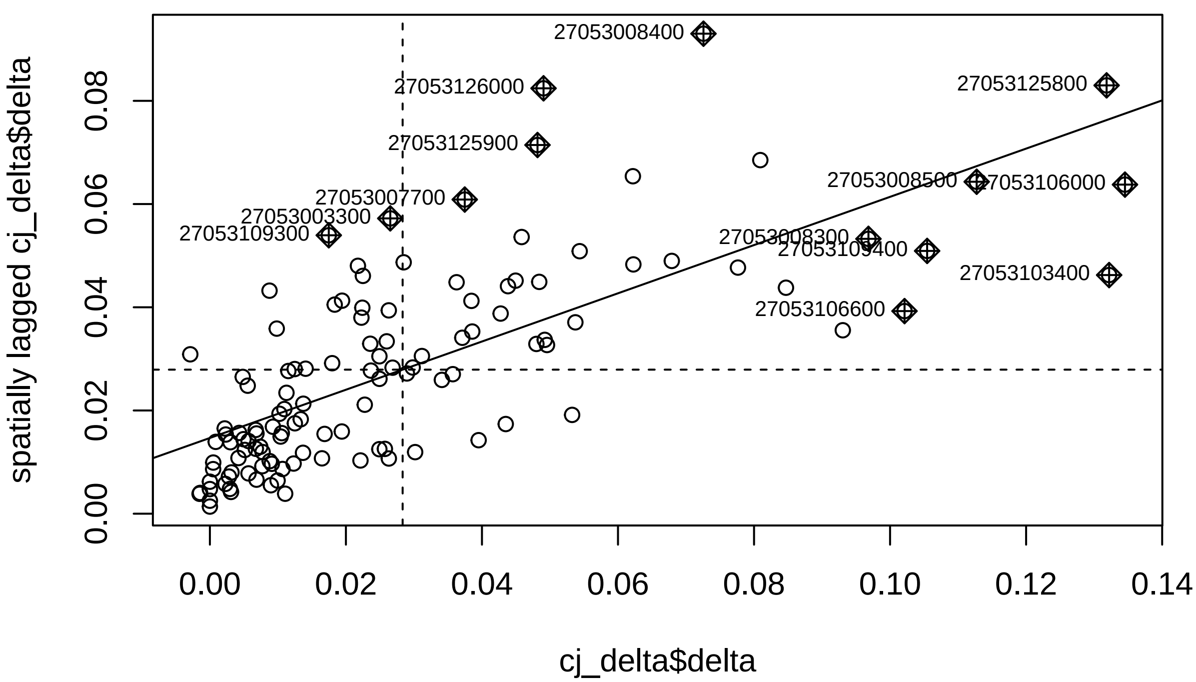
```
## Loading required package: spData

## To access larger datasets in this package, install the spDataLarge
## package with: `install.packages('spDataLarge',
## repos='https://nowosad.github.io/drat/', type='source')`
```

```
cj_delta <- prepost_cj %>%
  mutate(delta = car_jack_rate_1-car_jack_rate_0,
         GEOID = as.numeric(GEOID)) %>%
  left_join(mpls_tract, by = "GEOID") %>%
  st_as_sf()

nb <- st_contiguity(cj_delta, queen=TRUE)
wt <-  st_weights(nb, style = "W")

global_moran_test(
```

```r
  cj_delta$delta,
  nb,
  wt,
  alternative = "greater",
  randomization = TRUE)
```

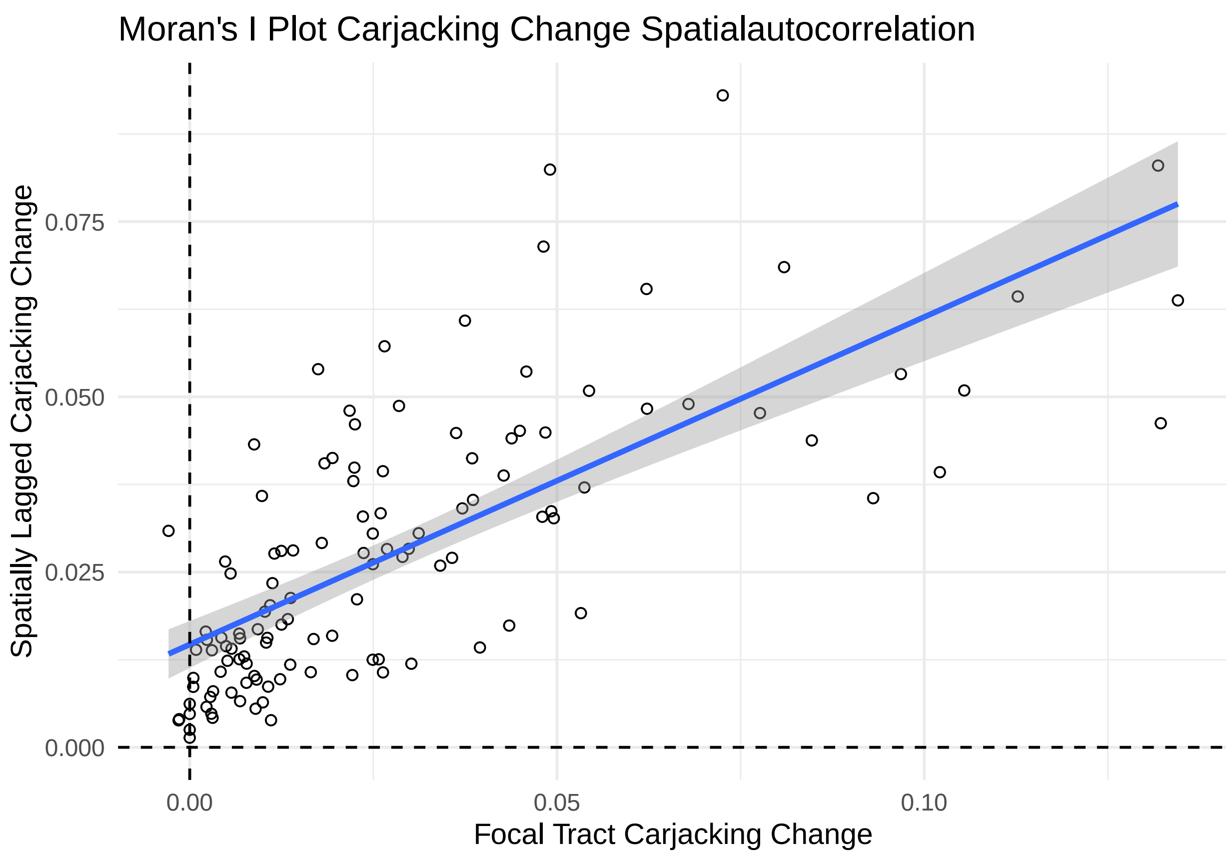
```
##
##  Moran I test under randomisation
##
## data:  x
## weights: listw
##
## Moran I statistic standard deviate = 8.9895, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.467232307      -0.008333333       0.002798653
```

```r
mp <- moran.plot(cj_delta$delta, nb2listw(nb),
                 labels = as.character(cj_delta$GEOID))
```




```r
ggplot(mp, aes(x=x, y=wx)) +
  geom_point(shape=1) +
    geom_smooth(formula=y ~ x, method="lm") +
    #geom_hline(yintercept=mean(mp$wx), lty=2) +
    geom_hline(yintercept=0, lty=2) +
    #geom_vline(xintercept=mean(mp$x), lty=2) +
```

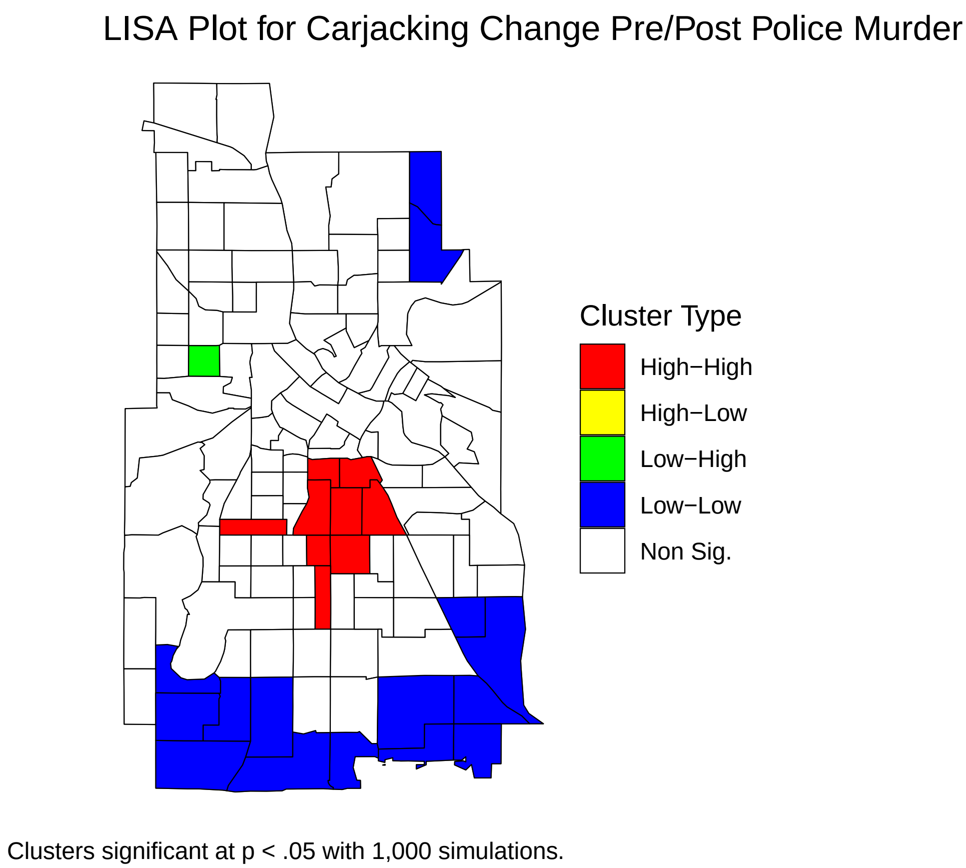
```r
    geom_vline(xintercept=0, lty=2) +
    theme_minimal() +
    labs(title = "Moran's I Plot Carjacking Change Spatialautocorrelation",
         x = "Focal Tract Carjacking Change",
         y = "Spatially Lagged Carjacking Change")
```




```r
ggsave(filename = "Car Jacking/Figures for publication/MI_CJ.png", bg="white", width = 10, height = 8)

global_moran_test(
  cj_delta$car_jack_rate_1,
  nb,
  wt,
  alternative = "greater",
  randomization = TRUE)
```

```
##
##  Moran I test under randomisation
##
## data:  x
## weights: listw
##
## Moran I statistic standard deviate = 9.7351, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.506295621      -0.008333333       0.002794527
```

```r
#LISA
cj_lisa <-  local_moran(cj_delta$delta,
                       nb = nb,
                       wt = wt,
                       nsim = 1000,
                       iseed = set.seed(7188)) %>%
  mutate(mean_p = ifelse(p_ii_sim <= 0.05, as.character(pysal), "Non Sig."),
         mean_p = factor(mean_p, levels = c("High-High", "High-Low", "Low-High",
                                            "Low-Low", "Non Sig.")))

cj_lisa_rate <-  local_moran(cj_delta$car_jack_rate_1,
                       nb = nb,
                       wt = wt,
                       nsim = 1000,
                       iseed = set.seed(7188)) %>%
  mutate(mean_p = ifelse(p_ii_sim <= 0.05, as.character(pysal), "Non Sig."),
         mean_p = factor(mean_p, levels = c("High-High", "High-Low", "Low-High",
                                            "Low-Low", "Non Sig.")))

cj_delta %>%
  cbind(cj_lisa) %>%
  ggplot(aes(fill = mean_p)) +
  geom_sf() +
  geom_sf(lwd = 0.2, color = "black") +
  theme_void() +
  scale_fill_manual(values = c("red", "yellow", "green", "blue","white"), drop = FALSE)+
  labs(title = "LISA Plot for Carjacking Change Pre/Post Police Murder",
       fill = "Cluster Type",
       caption = "Clusters significant at p < .05 with 1,000 simulations.")
```




```
ggsave(filename = "Car Jacking/Figures for publication/fig7.png", bg="white", width = 10, height = 8)
```

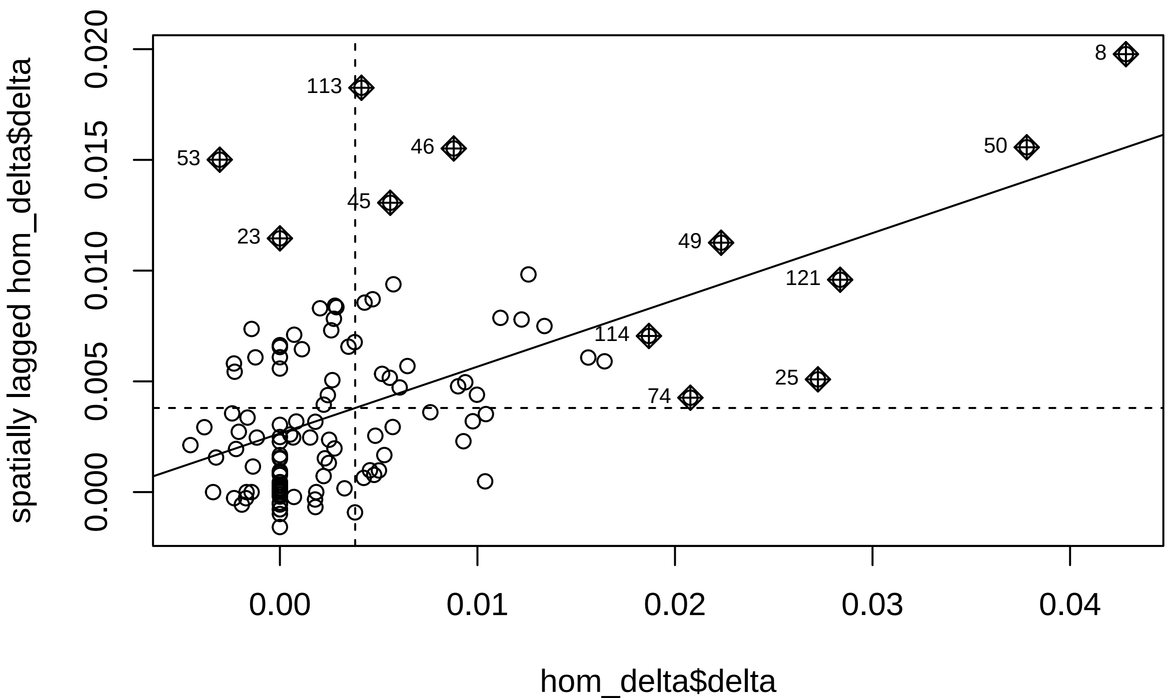
Homicide

```
hom_delta <- prepost_hom %>%
  mutate(delta = homicide_rate_1-homicide_rate_0,
         GEOID = as.numeric(GEOID)) %>%
  left_join(mpls_tract, by = "GEOID") %>%
  st_sf()


nb <- st_contiguity(hom_delta, queen=TRUE)
wt <-  st_weights(nb, style = "W")

global_moran_test(
  hom_delta$delta,
  nb,
  wt,
  alternative = "greater",
  randomization = TRUE)
```

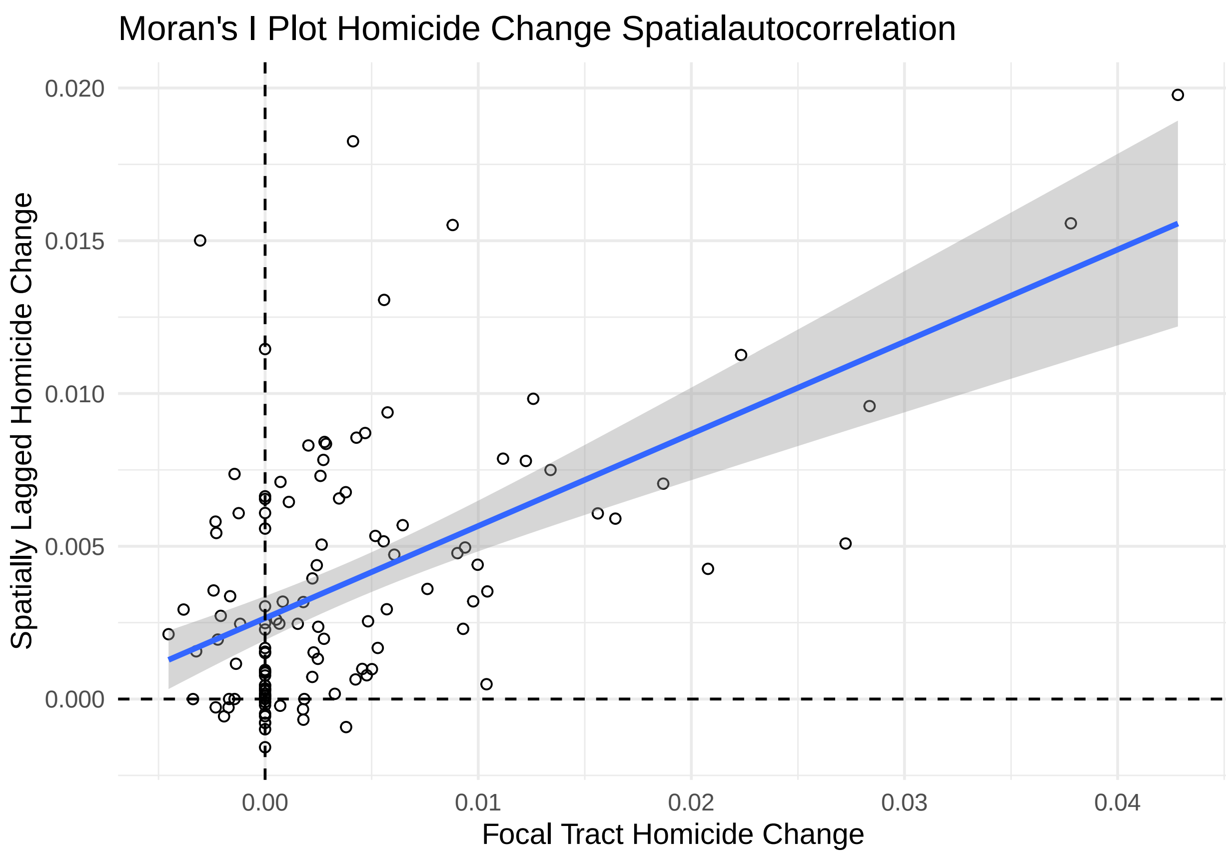
```
##
##  Moran I test under randomisation
##
## data:  x
## weights: listw
```

```
##
## Moran I statistic standard deviate = 6.0187, p-value = 8.791e-10
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.301559446      -0.008333333       0.002651046
```

```
mp <- moran.plot(hom_delta$delta, nb2listw(nb))
```




```
ggplot(mp, aes(x=x, y=wx)) +
  geom_point(shape=1) +
    geom_smooth(formula=y ~ x, method="lm") +
    #geom_hline(yintercept=mean(mp$wx), lty=2) +
    geom_hline(yintercept=0, lty=2) +
    #geom_vline(xintercept=mean(mp$x), lty=2) +
    geom_vline(xintercept=0, lty=2) +
    theme_minimal() +
    labs(title = "Moran's I Plot Homicide Change Spatialautocorrelation",
         x = "Focal Tract Homicide Change",
         y = "Spatially Lagged Homicide Change")
```



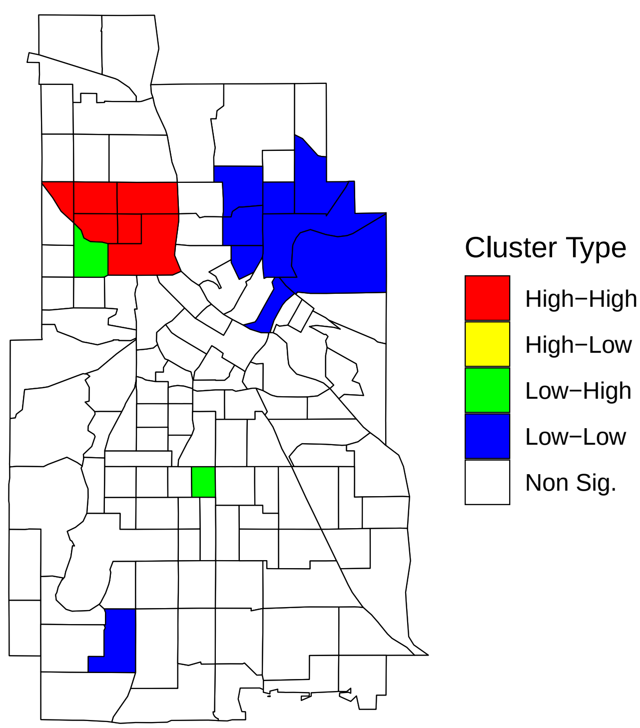

```r
ggsave(filename = "Car Jacking/Figures for publication/MI_HOM.png", bg="white", width = 10, height = 8)

global_moran_test(
  hom_delta$homicide_1,
  nb,
  wt,
  alternative = "greater",
  randomization = TRUE)
```

```
##
##  Moran I test under randomisation
##
## data:  x
## weights: listw
##
## Moran I statistic standard deviate = 6.725, p-value = 8.777e-12
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.347800619      -0.008333333       0.002804380
```

```r
#LISA
hom_lisa <-  local_moran(hom_delta$delta,
                       nb = nb,
                       wt = wt,
                       nsim = 1000,
                       iseed = set.seed(7188)) %>%
```

```r
  mutate(mean_p = ifelse(p_ii_sim <= 0.05, as.character(pysal), "Non Sig."),
         mean_p = factor(mean_p, levels = c("High-High", "High-Low", "Low-High",
                                             "Low-Low", "Non Sig.")))

hom_lisa_rate <-  local_moran(hom_delta$homicide_1,
                    nb = nb,
                    wt = wt,
                    nsim = 1000,
                    iseed = set.seed(7188)) %>%
  mutate(mean_p = ifelse(p_ii_sim <= 0.05, as.character(pysal), "Non Sig."),
         mean_p = factor(mean_p, levels = c("High-High", "High-Low", "Low-High",
                                             "Low-Low", "Non Sig.")))

hom_delta %>%
  cbind(hom_lisa) %>%
  ggplot(aes(fill = mean_p)) +
  geom_sf() +
  geom_sf(lwd = 0.2, color = "black") +
  theme_void() +
  scale_fill_manual(values = c("red", "yellow", "green","blue", "white"), drop = F)+
  labs(title = "LISA Plot for Homicide Change Pre/Post Police Murder",
       fill = "Cluster Type",
       caption = "Clusters significant at p < .05 with 1,000 simulations.")
```

LISA Plot for Homicide Change Pre/Post Police Murder



Clusters significant at p < .05 with 1,000 simulations.

```r
ggsave(filename = "Car Jacking/Figures for publication/fig8.png", bg="white", width = 10, height = 8)
```

RE CJ Models

```r
cj_exp_prepost <- cj_exp %>%
  group_by(GEOID) %>%
  mutate(begin_date = ISOweek2date(paste(year, paste0("W", sprintf("%02d", week)), 1,sep = "-")),
         end_date = begin_date+weeks(1)-days(1),
         post_floyd = as.numeric(begin_date >= as.Date("2020-05-25")),
         post_floyd_3 = as.numeric(begin_date >= as.Date("2020-05-25")+months(3)),
         weeks_post = as.numeric(begin_date-as.Date("2020-05-25"))/7,
         t_post_floyd = ifelse(weeks_post >=0,
                               weeks_post,
                               0),
        # stay_at_home = as.numeric(begin_date >= as.Date("2020-03-28") &
         #state_of_emerg = as.numeric(begin_date >= as.Date("2020-03-13")),
         period = factor(case_when(
           post_floyd==0 & post_floyd_3==0 ~ "Pre-Killing",
           post_floyd>=1 & post_floyd_3==0 ~ "0-3 Months Post-Killing",
           post_floyd>=1 & post_floyd_3>=1 ~ "3+ Months Post-Killing"),
           levels = c("Pre-Killing", "0-3 Months Post-Killing", "3+ Months Post-Killing")),
        GEOID = as.character(GEOID),
        anyjack = ifelse(car_jack==0, 0, 1),
        t = 1:length(car_jack_rate)) %>%
  left_join(acs_2020, by = c("GEOID")) %>%
  select(-med_hh_inc) %>%
  drop_na()



library(lme4)
library(lmerTest)


library(lavaan)

cd_model_1  <- ' cd =~ unemp_rate + pov_rate +  female_hh_rate + no_hs_dip_rate + black_perc
                  black_perc ~~ unemp_rate
                  unemp_rate ~~ pov_rate
                  no_hs_dip_rate    ~~  black_perc'

cfa_cd <- cfa(cd_model_1, data = cj_exp_prepost, std.lv = T)
```

```
## Warning in lav_object_post_check(object): lavaan WARNING: some estimated ov
## variances are negative
```

```r
modificationindices(cfa_cd)
```

```
## Warning in sqrt(var.lhs.value * var.rhs.value): NaNs produced

## Warning in lav_start_check_cov(lavpartable = lavpartable, start = start.values): lavaan WARNING: sta
##                    variables involved are: unemp_rate black_perc

## Warning in lav_start_check_cov(lavpartable = lavpartable, start = start.values): lavaan WARNING: sta
```

```
##                    variables involved are:  no_hs_dip_rate   black_perc
```

```
##    lhs op rhs      mi    epc sepc.lv sepc.all sepc.nox
## 14  cd ~~  cd 107.741 -0.505      -1       -1       -1
```

```r
summary(cfa_cd, fit.measures=TRUE, standardized = T)
```

```
## lavaan 0.6.17 ended normally after 56 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        13
##
##   Number of observations                         38357
##
## Model Test User Model:
##
##   Test statistic                                94.974
##   Degrees of freedom                                 2
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                             96458.724
##   Degrees of freedom                                10
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.999
##   Tucker-Lewis Index (TLI)                       0.995
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)            -605800.708
##   Loglikelihood unrestricted model (H1)    -605753.221
##
##   Akaike (AIC)                             1211627.416
##   Bayesian (BIC)                           1211738.627
##   Sample-size adjusted Bayesian (SABIC)    1211697.313
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.035
##   90 Percent confidence interval - lower         0.029
##   90 Percent confidence interval - upper         0.041
##   P-value H_0: RMSEA <= 0.050                    1.000
##   P-value H_0: RMSEA >= 0.080                    0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.006
##
## Parameter Estimates:
##
```

```
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   cd =~
##     unemp_rate        3.211    0.024  134.911    0.000    3.211    0.681
##     pov_rate          7.634    0.069  110.159    0.000    7.634    0.543
##     female_hh_rate    2.081    0.014  149.687    0.000    2.081    0.714
##     no_hs_dip_rate    5.031    0.029  170.706    0.000    5.031    0.844
##     black_perc       18.544    0.086  214.961    0.000   18.544    1.014
##
## Covariances:
##                     Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##  .unemp_rate ~~
##    .black_perc        -8.629    0.268  -32.162    0.000   -8.629   -0.808
##    .pov_rate           8.635    0.258   33.422    0.000    8.635    0.212
##  .no_hs_dip_rate ~~
##    .black_perc       -11.908    0.528  -22.541    0.000  -11.908   -1.204
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .unemp_rate       11.918    0.112  106.481    0.000   11.918    0.536
##    .pov_rate        139.458    1.053  132.426    0.000  139.458    0.705
##    .female_hh_rate    4.156    0.037  111.329    0.000    4.156    0.490
##    .no_hs_dip_rate   10.217    0.182   56.241    0.000   10.217    0.288
##    .black_perc       -9.567    2.076   -4.608    0.000   -9.567   -0.029
##     cd                1.000                               1.000    1.000
```

```
cd_predict <- as.vector(lavPredict(cfa_cd, newdata = as.data.frame(cj_exp_prepost)))
cj_exp_prepost$conc_dis <- cd_predict

cd_model_2  <- ' cd =~ unemp_rate + pov_rate +  female_hh_rate + no_hs_dip_rate
                  unemp_rate ~~ pov_rate'

cfa_cd_2 <- cfa(cd_model_2, data = cj_exp_prepost, std.lv = T)
modificationindices(cfa_cd_2)
```

```
##           lhs op            rhs     mi    epc sepc.lv sepc.all sepc.nox
## 11 unemp_rate ~~ female_hh_rate 20.785  0.302   0.302    0.042    0.042
## 12 unemp_rate ~~ no_hs_dip_rate 20.785 -0.744  -0.744   -0.068   -0.068
## 13   pov_rate ~~ female_hh_rate 20.785 -0.736  -0.736   -0.030   -0.030
## 14   pov_rate ~~ no_hs_dip_rate 20.785  1.814   1.814    0.049    0.049
```

```
summary(cfa_cd_2, fit.measures=TRUE, standardized = T)
```

```
## lavaan 0.6.17 ended normally after 32 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                         9
##
##   Number of observations                         38357
##
```

```
## Model Test User Model:
##
##   Test statistic                                20.791
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                             48310.688
##   Degrees of freedom                                 6
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    1.000
##   Tucker-Lewis Index (TLI)                       0.998
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)            -463947.614
##   Loglikelihood unrestricted model (H1)    -463937.219
##
##   Akaike (AIC)                              927913.229
##   Bayesian (BIC)                            927990.221
##   Sample-size adjusted Bayesian (SABIC)     927961.619
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.023
##   90 Percent confidence interval - lower         0.015
##   90 Percent confidence interval - upper         0.032
##   P-value H_0: RMSEA <= 0.050                    1.000
##   P-value H_0: RMSEA >= 0.080                    0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.004
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   cd =~
##     unemp_rate        3.233    0.024  134.525    0.000    3.233    0.685
##     pov_rate          7.888    0.075  105.151    0.000    7.888    0.561
##     female_hh_rate    2.044    0.015  140.274    0.000    2.044    0.702
##     no_hs_dip_rate    5.039    0.030  170.663    0.000    5.039    0.845
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
```

```
##   .unemp_rate ~~
##     .pov_rate          7.916    0.271   29.162    0.000    7.916    0.198
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     .unemp_rate      11.809    0.113  104.177    0.000   11.809    0.530
##     .pov_rate       135.515    1.130  119.917    0.000  135.515    0.685
##     .female_hh_rate   4.309    0.042  103.654    0.000    4.309    0.508
##     .no_hs_dip_rate  10.136    0.183   55.430    0.000   10.136    0.285
##      cd               1.000                               1.000    1.000
```

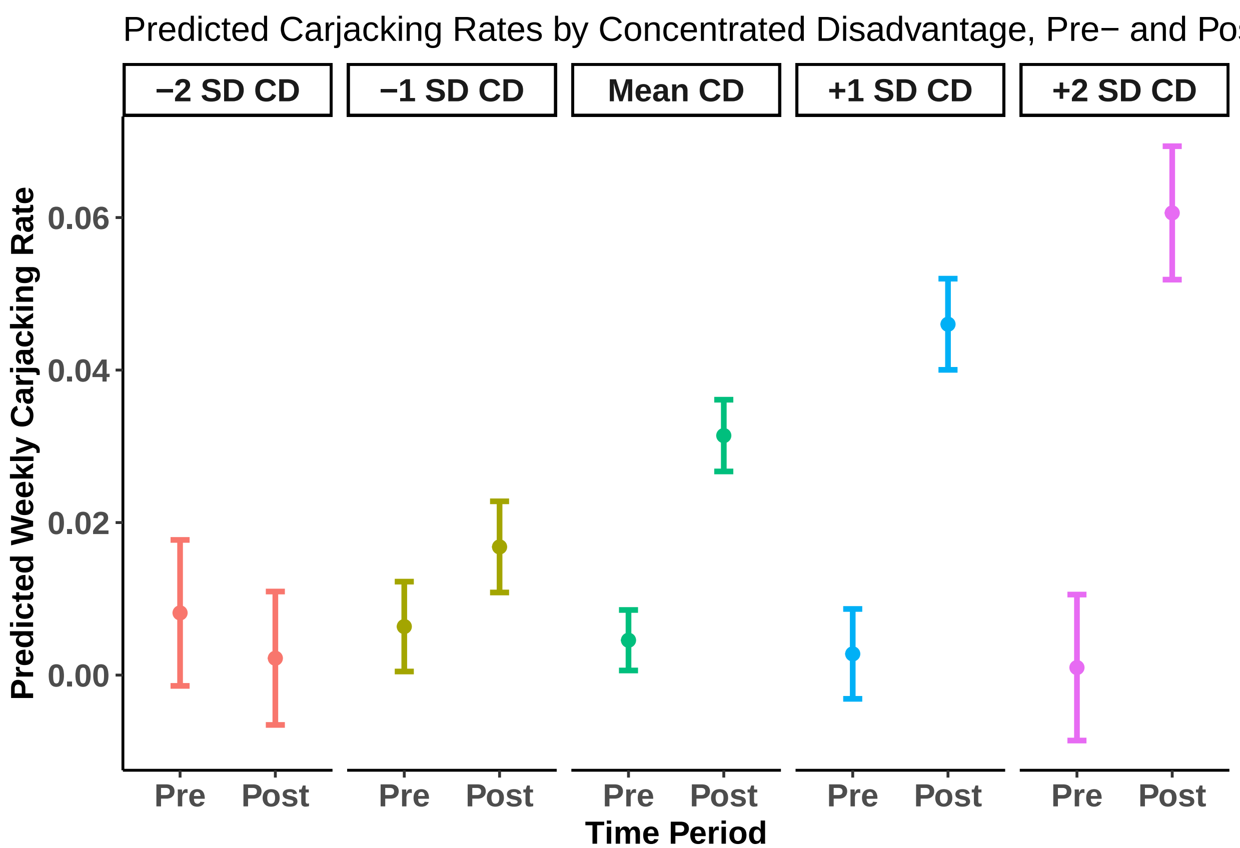
```
cd_predict_2 <- as.vector(lavPredict(cfa_cd_2, newdata = as.data.frame(cj_exp_prepost)))
cj_exp_prepost$conc_dis_no_black <- cd_predict_2


write_csv(cj_exp_prepost, file = "cj_exp_prepost.csv")


#predicted probability/rate plots
  #over time
  #stratified by CD (mean, +1SD, -1SD, +2SD, -2SD)


re <- lmer(car_jack_rate~t+post_floyd+t_post_floyd+
             conc_dis+
             age_19_29_perc+age_30_49_perc+age_50_69_perc+
             age_70_plus_perc+ post_floyd:conc_dis+t_post_floyd:conc_dis+
             (1|GEOID),
          data = cj_exp_prepost)
summary(re)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## car_jack_rate ~ t + post_floyd + t_post_floyd + conc_dis + age_19_29_perc +
##     age_30_49_perc + age_50_69_perc + age_70_plus_perc + post_floyd:conc_dis +
##     t_post_floyd:conc_dis + (1 | GEOID)
##    Data: cj_exp_prepost
##
## REML criterion at convergence: -71430.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.2639 -0.2541 -0.1167  0.0169 31.5976
##
## Random effects:
##  Groups   Name        Variance  Std.Dev.
##  GEOID    (Intercept) 0.0001661 0.01289
##  Residual             0.0090021 0.09488
## Number of obs: 38357, groups:  GEOID, 121
##
## Fixed effects:
##                         Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)            5.007e-02  1.557e-02  1.164e+02   3.215  0.00169 **
## t                      1.490e-05  1.237e-05  3.823e+04   1.204  0.22847
```

```
## post_floyd             2.684e-02  1.953e-03  3.823e+04  13.739  < 2e-16 ***
## t_post_floyd          -1.775e-05  2.237e-05  3.823e+04  -0.794  0.42735
## conc_dis               6.052e-04  1.848e-03  1.273e+02   0.327  0.74389
## age_19_29_perc        -4.414e-04  1.637e-04  1.150e+02  -2.697  0.00806 **
## age_30_49_perc        -3.714e-04  2.562e-04  1.150e+02  -1.450  0.14989
## age_50_69_perc        -9.986e-04  3.027e-04  1.150e+02  -3.300  0.00129 **
## age_70_plus_perc      -6.633e-04  3.699e-04  1.150e+02  -1.793  0.07559 .
## post_floyd:conc_dis    1.639e-02  1.557e-03  3.823e+04  10.524  < 2e-16 ***
## t_post_floyd:conc_dis -3.525e-05  1.813e-05  3.823e+04  -1.944  0.05187 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) t      pst_fl t_pst_ cnc_ds a_19_2 a_30_4 a_50_6 a_70__
## t            -0.072
## post_floyd    0.028 -0.573
## t_post_flyd   0.040 -0.553 -0.223
## conc_dis     -0.692  0.000  0.000  0.000
## ag_19_29_pr -0.932  0.000  0.000  0.000  0.678
## ag_30_49_pr -0.870  0.000  0.000  0.000  0.546  0.746
## ag_50_69_pr -0.736  0.000  0.000  0.000  0.600  0.714  0.396
## ag_70_pls_p -0.264  0.000  0.000  0.000  0.137  0.243  0.245 -0.170
## pst_flyd:c_  0.000  0.000  0.000  0.000 -0.136  0.000  0.000  0.000  0.000
## t_pst_fly:_  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000
##              pst_:_
## t
## post_floyd
## t_post_flyd
## conc_dis
## ag_19_29_pr
## ag_30_49_pr
## ag_50_69_pr
## ag_70_pls_p
## pst_flyd:c_
## t_pst_fly:_ -0.792
```

```r
library(emmeans)

pred_raw <- emmeans(re, c("conc_dis", "post_floyd"),
                    at = list(conc_dis = c(-2, -1, 0, 1, 2),
                              post_floyd = c(0,1),
                              t_post_floyd = mean(cj_exp_prepost$t_post_floyd[cj_exp_prepost$post_floyd
                              t = mean(cj_exp_prepost$t[cj_exp_prepost$post_floyd==0]))) %>%
  as.data.frame() %>%
  mutate(conc_dis = factor(case_when(
    conc_dis==-2~"-2 SD CD",
    conc_dis==-1~"-1 SD CD",
    conc_dis==0~"Mean CD",
    conc_dis==1~"+1 SD CD",
    conc_dis==2~"+2 SD CD"),
    levels = c("-2 SD CD",
               "-1 SD CD",
               "Mean CD",
               "+1 SD CD",
```

```
                "+2 SD CD")),
    post_floyd = factor(case_when(
    post_floyd==1~"Post",
    post_floyd==0~"Pre"),
    levels = c("Pre", "Post")))

ggplot(pred_raw, aes(x = post_floyd, y = emmean, color = conc_dis))+
  geom_point(size = 2)+
  geom_errorbar(aes(ymin=asymp.LCL, ymax=asymp.UCL), width = .2, size = 1)+
  labs(title = "Predicted Carjacking Rates by Concentrated Disadvantage, Pre- and Post-Killing",
       y = "Predicted Weekly Carjacking Rate",
       x = "Time Period")+
  facet_grid(~conc_dis)+
  guides(color = "none")+
  theme_classic()+
  theme(axis.text = element_text(face = "bold", size = 12),
        axis.title.x = element_text(face = "bold", size = 12),
        axis.title.y = element_text(face = "bold", size = 12),
        strip.text.x = element_text(face = "bold", size = 12))
```




```
ggsave(filename = "Car Jacking/Figures for publication/fig9.png", bg="white", width = 8, height = 8)
```

CD Map

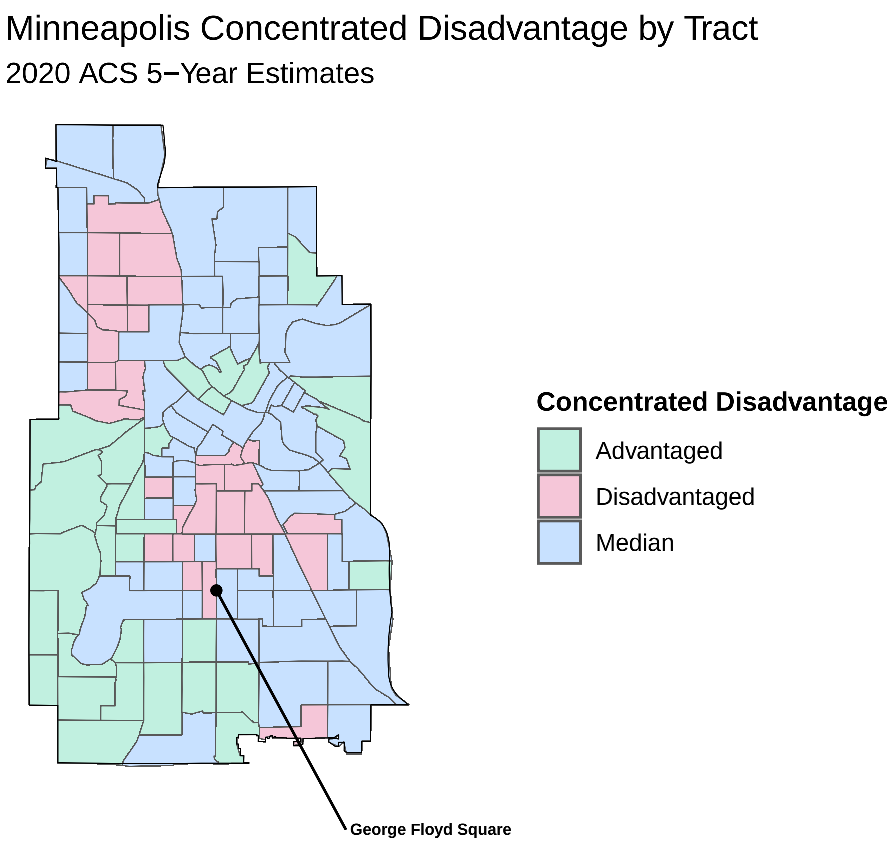
```
cd_cat <- read_csv("Data/cd_noblack_forRyan_031324.csv")
```

```
## Rows: 121 Columns: 3
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): Concentrated Disadvantage
## dbl (2): GEOID, conc_dis_no_black
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
cj_exp_tract_year_cd <- cj_exp_prepost %>%
  filter(year==2020)

q_25 <- quantile(cj_exp_tract_year_cd$conc_dis_no_black, .25)
q_75 <- quantile(cj_exp_tract_year_cd$conc_dis_no_black, .75)


cj_exp_tract_year_cd <- cj_exp_prepost %>%
  mutate(`Concentrated Disadvantage` = case_when(
    conc_dis_no_black < q_25 ~ "Advantaged",
    conc_dis_no_black >= q_25 & conc_dis_no_black <= q_75  ~ "Median",
    conc_dis_no_black > q_75 ~ "Disadvantaged"
  ))

cj_exp_tract_year_cd <-  cj_exp_tract_year_cd %>%
  mutate(GEOID = as.numeric(GEOID)) %>%
  left_join(cd_cat, by = "GEOID")

table(cj_exp_tract_year_cd$`Concentrated Disadvantage.x`==cj_exp_tract_year_cd$`Concentrated Disadvanta
```

```
##
##  TRUE
## 38357
```

```
MyColour <- c("#c1f0e0", "#c8e1ff", "#f5c6d8")
names(MyColour) <- c("Advantaged", "Median", "Disadvantaged")

ggplot() +
  geom_sf(data = cj_exp_tract_year_cd,
          aes(geometry = geometry, fill = `Concentrated Disadvantage.x`)) +
  geom_sf(data = mpls, aes(geometry = geometry), color = "black", alpha = 0)+
  geom_sf(data = gfs, aes(geometry = geometry), color = "black")+
  geom_text_repel(data = gfs_label, aes(x=lon, y=lat, label = name),
                  size = 2,
                  fontface = "bold",
                  nudge_x = .1, nudge_y = -.06)+
  scale_fill_manual(values = MyColour)+
  labs(title = "Minneapolis Concentrated Disadvantage by Tract",
       subtitle = "2020 ACS 5-Year Estimates",
       fill = "Concentrated Disadvantage")+
  theme_void()+
  theme(strip.text.x = element_text(face = "bold", size = 12),
        legend.title = element_text(face = "bold", size = 10))
```




```
ggsave(filename = "Car Jacking/Figures for publication/A_CD.png", bg="white", width = 8, height = 8)
```

```
results_table<-standardizedSolution(cfa_cd) %>%
  filter(row_number() %in% c(1:8)) %>%
  dplyr::select(LHS=lhs, Specification=op, RHS=rhs, 'Std(Beta)'=est.std, SE=se,
              'P-Value'=pvalue) %>%
  mutate(LHS = case_when(
    LHS=="cd"~"Conc. Dis.",
    LHS=="unemp_rate"~"Unemp. Rate",
    LHS=="no_hs_dip_rate"~"No HS Diploma Rate"),
         RHS = case_when(
           RHS=="unemp_rate"~"Unemp. Rate",
           RHS=="pov_rate"~"Poverty Rate",
           RHS=="female_hh_rate"~"Female-HH Rate",
           RHS=="no_hs_dip_rate"~"No HS Diploma Rate",
           RHS=="black_perc"~"Percent Black"
         ),
    Specification = case_when(
      Specification=="=~"~"FL",
      Specification=="~~"~"Cov."))



stargazer(results_table, summary = FALSE, header = F,
           type="latex", style="aer", align = T,
           title="CFA Measurement Model of Concentrated Disadvantage",
```

```
          notes="$LR\\chi^2$ vs. saturated (1) = 20.79***, RMSEA = .023 (PCLOSE = 1.0), CFI = 1.0, SRMR
```

Table 1: CFA Measurement Model of Concentrated Disadvantage

	LHS	Specification	RHS	Std(Beta)	SE	P-Value
1	Conc. Dis.	FL	Unemp. Rate	0.681	0.004	0
2	Conc. Dis.	FL	Poverty Rate	0.543	0.004	0
3	Conc. Dis.	FL	Female-HH Rate	0.714	0.003	0
4	Conc. Dis.	FL	No HS Diploma Rate	0.844	0.003	0
5	Conc. Dis.	FL	Percent Black	1.014	0.003	0
6	Unemp. Rate	Cov.	Percent Black	−0.808	0.079	0
7	Unemp. Rate	Cov.	Poverty Rate	0.212	0.006	0
8	No HS Diploma Rate	Cov.	Percent Black	−1.204	0.088	0

$LR\chi^2$ vs. saturated (1) = 20.79***, RMSEA = .023 (PCLOSE = 1.0), CFI = 1.0, SRMR = .004

```
results_table<-standardizedSolution(cfa_cd_2) %>%
  filter(row_number() %in% c(1:5)) %>%
  dplyr::select(LHS=lhs, Specification=op, RHS=rhs, 'Std(Beta)'=est.std, SE=se,
              'P-Value'=pvalue) %>%
  mutate(LHS = case_when(
    LHS=="cd"~"Conc. Dis.",
    LHS=="unemp_rate"~"Unemp. Rate",
    LHS=="no_hs_dip_rate"~"No HS Diploma Rate"),
         RHS = case_when(
           RHS=="unemp_rate"~"Unemp. Rate",
           RHS=="pov_rate"~"Poverty Rate",
           RHS=="female_hh_rate"~"Female-HH Rate",
           RHS=="no_hs_dip_rate"~"No HS Diploma Rate"
         ),
    Specification = case_when(
      Specification=="=~"~"FL",
      Specification=="~~"~"Cov."))



stargazer(results_table, summary = FALSE, header = F,
          type="latex", style="aer", align = T,
          title="CFA Measurement Model of Concentrated Disadvantage",
          notes="$LR\\chi^2$ vs. saturated (2) = 94.97, RMSEA = .035 (PCLOSE = 1.0), CFI = .999, SRMR =
```

Table 2: CFA Measurement Model of Concentrated Disadvantage

	LHS	Specification	RHS	Std(Beta)	SE	P-Value
1	Conc. Dis.	FL	Unemp. Rate	0.685	0.004	0
2	Conc. Dis.	FL	Poverty Rate	0.561	0.004	0
3	Conc. Dis.	FL	Female-HH Rate	0.702	0.003	0
4	Conc. Dis.	FL	No HS Diploma Rate	0.845	0.003	0
5	Unemp. Rate	Cov.	Poverty Rate	0.198	0.006	0

$LR\chi^2$ vs. saturated (2) = 94.97, RMSEA = .035 (PCLOSE = 1.0), CFI = .999, SRMR = .006

```r
re_logit_cd <- glmer(anyjack ~ t+post_floyd+t_post_floyd+conc_dis+
             age_19_29_perc+age_30_49_perc+age_50_69_perc+
             age_70_plus_perc+ post_floyd:conc_dis+t_post_floyd:conc_dis+
             (1|GEOID),
                  data = cj_exp_prepost, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.013781 (tol = 0.002, component 1)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unider
##  - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```r
summary(re_logit_cd)
```

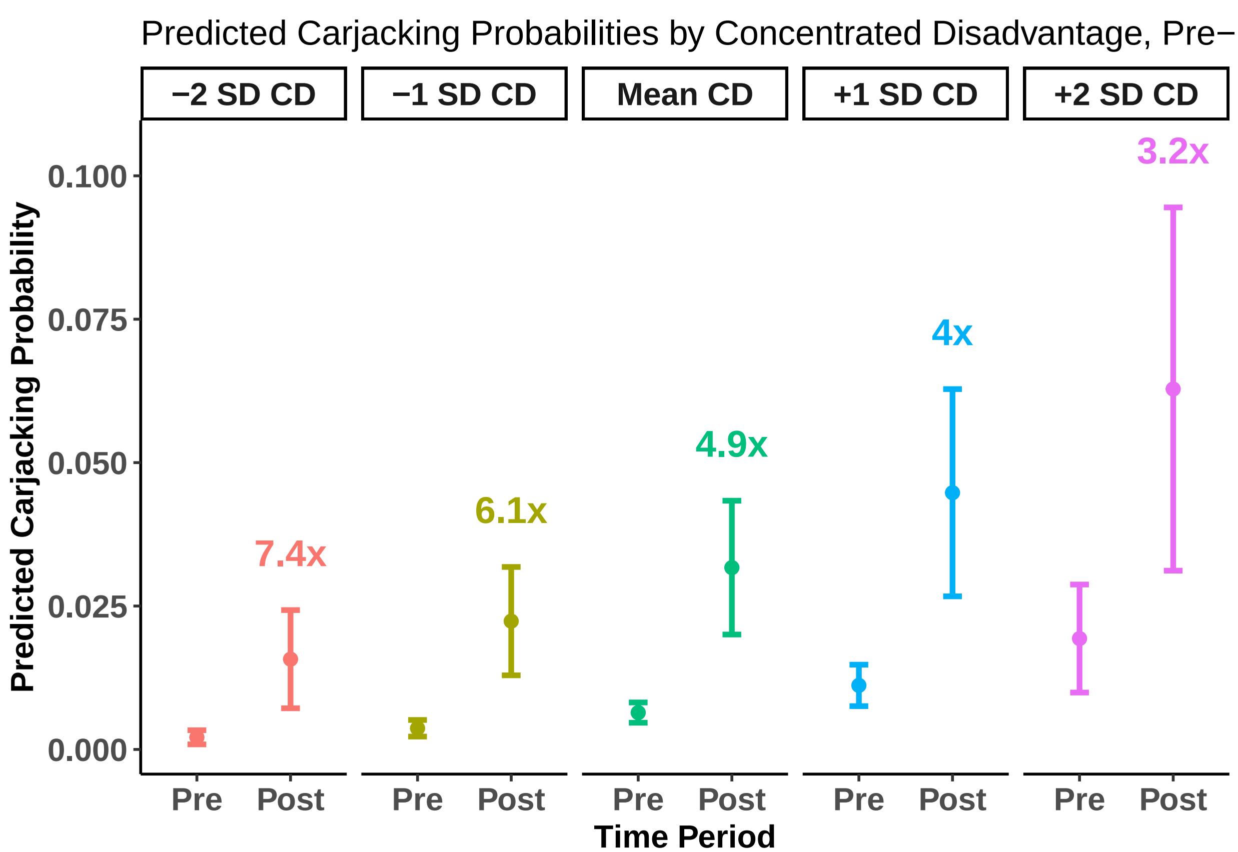
```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: anyjack ~ t + post_floyd + t_post_floyd + conc_dis + age_19_29_perc +
##     age_30_49_perc + age_50_69_perc + age_70_plus_perc + post_floyd:conc_dis +
##     t_post_floyd:conc_dis + (1 | GEOID)
##    Data: cj_exp_prepost
##
##      AIC      BIC   logLik deviance df.resid
##  11180.3  11282.9  -5578.1  11156.3    38345
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -0.7210 -0.2190 -0.1384 -0.0723 21.3407
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  GEOID  (Intercept) 0.4458   0.6677
## Number of obs: 38357, groups:  GEOID, 121
##
## Fixed effects:
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -4.5893181  0.8639692  -5.312 1.08e-07 ***
## t                      0.0038429  0.0011368   3.380 0.000724 ***
## post_floyd             1.6219023  0.1316410  12.321  < 2e-16 ***
## t_post_floyd          -0.0031135  0.0013874  -2.244 0.024828 *
## conc_dis               0.6387035  0.1092871   5.844 5.09e-09 ***
## age_19_29_perc        -0.0008097  0.0090043  -0.090 0.928349
## age_30_49_perc         0.0147866  0.0140641   1.051 0.293088
## age_50_69_perc        -0.0458157  0.0168100  -2.725 0.006420 **
## age_70_plus_perc      -0.0175089  0.0202488  -0.865 0.387209
## post_floyd:conc_dis   -0.1990058  0.0752563  -2.644 0.008184 **
## t_post_floyd:conc_dis -0.0011956  0.0007052  -1.695 0.089998 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) t      pst_fl t_pst_ cnc_ds a_19_2 a_30_4 a_50_6 a_70__
## t           -0.134
## post_floyd   0.050 -0.697
```

```
## t_post_flyd  0.109 -0.819  0.331
## conc_dis    -0.682  0.003  0.136 -0.001
## ag_19_29_pr -0.924  0.002 -0.001 -0.001  0.653
## ag_30_49_pr -0.864  0.002  0.000 -0.001  0.535  0.745
## ag_50_69_pr -0.726  0.001 -0.002 -0.001  0.558  0.710  0.394
## ag_70_pls_p -0.243  0.000 -0.004 -0.001  0.118  0.226  0.233 -0.189
## pst_flyd:c_  0.034 -0.003 -0.316  0.123 -0.328 -0.002 -0.004 -0.003  0.008
## t_pst_fly:_  0.003  0.000  0.134 -0.180 -0.007 -0.001 -0.003 -0.002  0.003
##             pst_:_
## t
## post_floyd
## t_post_flyd
## conc_dis
## ag_19_29_pr
## ag_30_49_pr
## ag_50_69_pr
## ag_70_pls_p
## pst_flyd:c_
## t_pst_fly:_ -0.625
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.013781 (tol = 0.002, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
## Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```
pred_prob_raw <- emmeans(re_logit_cd, c("conc_dis", "post_floyd"),
                   at = list(conc_dis = c(-2, -1, 0, 1, 2),
                             post_floyd = c(0,1),
                             t_post_floyd = mean(cj_exp_prepost$t_post_floyd[cj_exp_prepost$post_floyd=
                             t = mean(cj_exp_prepost$t[cj_exp_prepost$post_floyd==0])),
                   trans = "response") %>%
  as.data.frame() %>%
  mutate(conc_dis = factor(case_when(
    conc_dis==-2~"-2 SD CD",
    conc_dis==-1~"-1 SD CD",
    conc_dis==0~"Mean CD",
    conc_dis==1~"+1 SD CD",
    conc_dis==2~"+2 SD CD"),
    levels = c("-2 SD CD",
               "-1 SD CD",
               "Mean CD",
               "+1 SD CD",
               "+2 SD CD")),
    post_floyd = factor(case_when(
    post_floyd==1~"Post",
    post_floyd==0~"Pre"),
    levels = c("Pre", "Post")))

ratios <- pred_prob_raw %>%
  select(conc_dis, post_floyd, prob) %>%
  pivot_wider(names_from = post_floyd, values_from = prob) %>%
  mutate(ratio = Post/Pre)
```

```r
pred_prob_raw <- pred_prob_raw %>%
  left_join(ratios, by = "conc_dis") %>%
  mutate(ratio = ifelse(post_floyd=="Pre", NA_integer_, round(ratio,1)),
         ratio = ifelse(is.na(ratio), ratio, paste0(ratio, "x")))

ggplot(pred_prob_raw, aes(x = post_floyd, y = prob, color = conc_dis))+
  geom_point(size = 2)+
  geom_errorbar(aes(ymin=asymp.LCL, ymax=asymp.UCL), width = .2, size = 1)+
  labs(title = "Predicted Carjacking Probabilities by Concentrated Disadvantage, Pre- and Post-Killing"
       y = "Predicted Carjacking Probability",
       x = "Time Period")+
  theme_classic()+
  facet_grid(~conc_dis)+
  guides(color = "none")+
  geom_text(aes(x = post_floyd, y = asymp.UCL+.01, label = ratio),
            fontface = "bold", size = 5)+
  theme(axis.text = element_text(face = "bold", size = 12),
        axis.title.x = element_text(face = "bold", size = 12),
        axis.title.y = element_text(face = "bold", size = 12),
        strip.text.x = element_text(face = "bold", size = 12))
```

```
## Warning: Removed 5 rows containing missing values (`geom_text()`).
```




```r
ggsave(filename = "Car Jacking/Figures for publication/fig10.png", bg="white", width = 8, height = 8)
```

```
## Warning: Removed 5 rows containing missing values (`geom_text()`).
```

```r
homicide <- homicide %>% st_drop_geometry()

cj_exp_prepost <- cj_exp_prepost %>%
  mutate(GEOID = as.numeric(GEOID)) %>%
  left_join(homicide, by = c("GEOID", "year","week")) %>%
  mutate(anyhom = ifelse(homicide==0, 0, 1))

re_homicide <- lmer(homicide_rate~t+post_floyd+t_post_floyd+conc_dis+
             age_19_29_perc+age_30_49_perc+age_50_69_perc+
             age_70_plus_perc+ post_floyd:conc_dis+t_post_floyd:conc_dis+
             (1|GEOID),
          data = cj_exp_prepost)
summary(re_homicide)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## homicide_rate ~ t + post_floyd + t_post_floyd + conc_dis + age_19_29_perc +
##     age_30_49_perc + age_50_69_perc + age_70_plus_perc + post_floyd:conc_dis +
##     t_post_floyd:conc_dis + (1 | GEOID)
##    Data: cj_exp_prepost
##
## REML criterion at convergence: -131576.1
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -1.027 -0.107 -0.033 -0.007 44.543
##
## Random effects:
##  Groups   Name        Variance  Std.Dev.
##  GEOID    (Intercept) 1.582e-05 0.003977
##  Residual             1.879e-03 0.043349
## Number of obs: 38357, groups:  GEOID, 121
##
## Fixed effects:
##                         Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)            1.054e-02  5.221e-03  1.176e+02   2.019 0.045761 *
## t                      1.165e-05  5.653e-06  3.823e+04   2.061 0.039293 *
## post_floyd             4.206e-03  8.925e-04  3.823e+04   4.713 2.45e-06 ***
## t_post_floyd          -3.297e-05  1.022e-05  3.823e+04  -3.226 0.001257 **
## conc_dis               1.297e-03  6.312e-04  1.384e+02   2.055 0.041796 *
## age_19_29_perc        -7.918e-05  5.473e-05  1.150e+02  -1.447 0.150714
## age_30_49_perc        -1.845e-04  8.567e-05  1.150e+02  -2.154 0.033352 *
## age_50_69_perc        -4.694e-06  1.012e-04  1.150e+02  -0.046 0.963084
## age_70_plus_perc      -3.238e-04  1.237e-04  1.150e+02  -2.617 0.010059 *
## post_floyd:conc_dis    5.893e-03  7.116e-04  3.823e+04   8.281  < 2e-16 ***
## t_post_floyd:conc_dis -3.112e-05  8.284e-06  3.823e+04  -3.756 0.000173 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) t      pst_fl t_pst_ cnc_ds a_19_2 a_30_4 a_50_6 a_70__
## t            -0.098
## post_floyd    0.038 -0.573
```

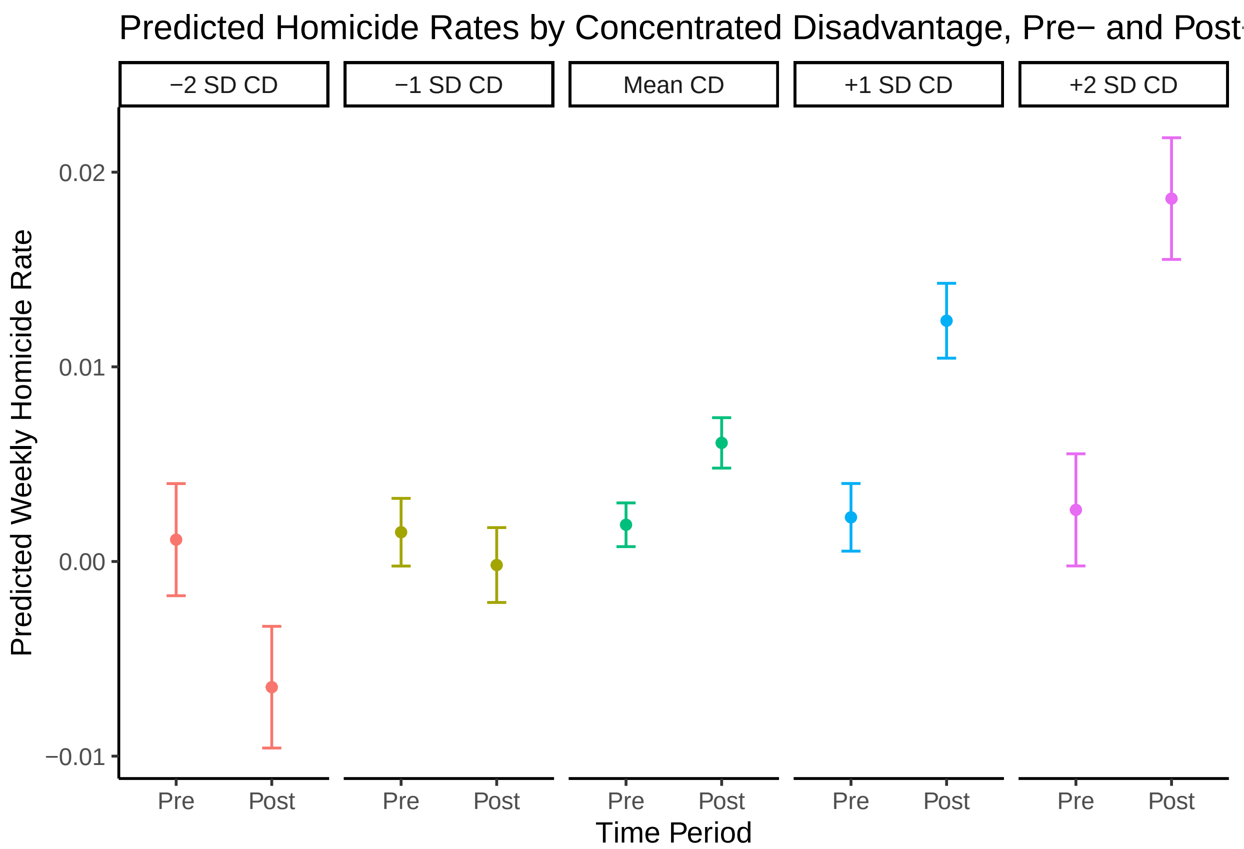
```
## t_post_flyd  0.054 -0.553 -0.223
## conc_dis    -0.676  0.000  0.000  0.000
## ag_19_29_pr -0.930  0.000  0.000  0.000  0.664
## ag_30_49_pr -0.868  0.000  0.000  0.000  0.535  0.746
## ag_50_69_pr -0.734  0.000  0.000  0.000  0.587  0.714  0.396
## ag_70_pls_p -0.264  0.000  0.000  0.000  0.134  0.243  0.245 -0.170
## pst_flyd:c_  0.000  0.000  0.000  0.000 -0.182  0.000  0.000  0.000  0.000
## t_pst_fly:_  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000
##             pst_:_
## t
## post_floyd
## t_post_flyd
## conc_dis
## ag_19_29_pr
## ag_30_49_pr
## ag_50_69_pr
## ag_70_pls_p
## pst_flyd:c_
## t_pst_fly:_ -0.792
```

```r
pred_raw_homicide <- emmeans(re_homicide, c("conc_dis", "post_floyd"),
                    at = list(conc_dis = c(-2, -1, 0, 1, 2),
                              post_floyd = c(0,1))) %>%
  as.data.frame() %>%
  mutate(conc_dis = factor(case_when(
    conc_dis==-2~"-2 SD CD",
    conc_dis==-1~"-1 SD CD",
    conc_dis==0~"Mean CD",
    conc_dis==1~"+1 SD CD",
    conc_dis==2~"+2 SD CD"),
    levels = c("-2 SD CD",
               "-1 SD CD",
               "Mean CD",
               "+1 SD CD",
               "+2 SD CD")),
    post_floyd = factor(case_when(
    post_floyd==1~"Post",
    post_floyd==0~"Pre"),
    levels = c("Pre", "Post")))
```

```
## Note: D.f. calculations have been disabled because the number of observations exceeds 3000.
## To enable adjustments, add the argument 'pbkrtest.limit = 38357' (or larger)
## [or, globally, 'set emm_options(pbkrtest.limit = 38357)' or larger];
## but be warned that this may result in large computation time and memory use.

## Note: D.f. calculations have been disabled because the number of observations exceeds 3000.
## To enable adjustments, add the argument 'lmerTest.limit = 38357' (or larger)
## [or, globally, 'set emm_options(lmerTest.limit = 38357)' or larger];
## but be warned that this may result in large computation time and memory use.
```

```r
ggplot(pred_raw_homicide, aes(x = post_floyd, y = emmean, color = conc_dis))+
  geom_point()+
  geom_errorbar(aes(ymin=asymp.LCL, ymax=asymp.UCL), width=0.2)+
  labs(title = "Predicted Homicide Rates by Concentrated Disadvantage, Pre- and Post-Killing",
       y = "Predicted Weekly Homicide Rate",
       x = "Time Period")+
```

```r
  theme_classic()+
  facet_grid(~conc_dis)+
  guides(color = "none")
```




```r
re_logit_cd_homicide <- glmer(anyhom ~ t+post_floyd+t_post_floyd+conc_dis+
             age_19_29_perc+age_30_49_perc+age_50_69_perc+
             age_70_plus_perc+ post_floyd:conc_dis+t_post_floyd:conc_dis+
             (1|GEOID),
                  data = cj_exp_prepost, family = binomial)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.0606537 (tol = 0.002, component 1)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly uniden
##  - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```r
summary(re_logit_cd_homicide)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: anyhom ~ t + post_floyd + t_post_floyd + conc_dis + age_19_29_perc +
##     age_30_49_perc + age_50_69_perc + age_70_plus_perc + post_floyd:conc_dis +
##     t_post_floyd:conc_dis + (1 | GEOID)
##    Data: cj_exp_prepost
##
```

```
##      AIC      BIC   logLik deviance df.resid
##   3457.2   3559.8  -1716.6   3433.2    38345
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
## -0.334 -0.092 -0.060 -0.042 33.923
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  GEOID  (Intercept) 0.6529   0.808
## Number of obs: 38357, groups:  GEOID, 121
##
## Fixed effects:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -7.473166   1.231468  -6.068 1.29e-09 ***
## t                     0.004926   0.001892   2.604  0.00921 **
## post_floyd            0.757974   0.246947   3.069  0.00215 **
## t_post_floyd         -0.007602   0.002839  -2.678  0.00741 **
## conc_dis              0.867038   0.163990   5.287 1.24e-07 ***
## age_19_29_perc        0.016861   0.013042   1.293  0.19608
## age_30_49_perc        0.017916   0.020305   0.882  0.37760
## age_50_69_perc        0.022501   0.025188   0.893  0.37169
## age_70_plus_perc     -0.055565   0.031831  -1.746  0.08088 .
## post_floyd:conc_dis   0.178816   0.139286   1.284  0.19921
## t_post_floyd:conc_dis -0.002551   0.001549  -1.647  0.09963 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) t      pst_fl t_pst_ cnc_ds a_19_2 a_30_4 a_50_6 a_70__
## t           -0.158
## post_floyd   0.049 -0.595
## t_post_flyd  0.103 -0.666 -0.014
## conc_dis    -0.700  0.001  0.140  0.007
## ag_19_29_pr -0.905  0.000  0.001  0.000  0.660
## ag_30_49_pr -0.829  0.000  0.000  0.002  0.517  0.689
## ag_50_69_pr -0.670  0.000 -0.001  0.002  0.530  0.646  0.276
## ag_70_pls_p -0.226  0.000  0.004 -0.005  0.089  0.208  0.230 -0.246
## pst_flyd:c_  0.032 -0.001 -0.480  0.305 -0.314 -0.001  0.001  0.002 -0.009
## t_pst_fly:_  0.005  0.000  0.314 -0.423 -0.018 -0.001 -0.004 -0.004  0.012
##             pst_:_
## t
## post_floyd
## t_post_flyd
## conc_dis
## ag_19_29_pr
## ag_30_49_pr
## ag_50_69_pr
## ag_70_pls_p
## pst_flyd:c_
## t_pst_fly:_ -0.655
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.0606537 (tol = 0.002, component 1)
## Model is nearly unidentifiable: very large eigenvalue
```

```
##   - Rescale variables?
## Model is nearly unidentifiable: large eigenvalue ratio
##   - Rescale variables?
```

```
class(re) <- "lmerMod"
class(re_logit_cd) <- "lmerMod"

stargazer(re, re_logit_cd,
          title = "Interrupted Time Series Models of Carjackings, MPLS 2017-2022",
          covariate.labels = c("T", "Post-Killing", "T Post-Killing",
                              "Conc. Dis.", "Age 19-29", "Age 30-49",
                              "Age 50-69", "Age 70+",
                              "Post-Killing X Conc. Dis.",
                              "T Post-Killing X Conc. Dis."),
          header = F,
          dep.var.caption = "Carjacking",
          dep.var.labels = c("Rate per 1,000", "Any Carjacking"),
          model.names = FALSE,
          column.labels = c("RE HLM","RE Logit"),
          report = "vcs",
          ci=TRUE,
          ci.level=0.95,
          ci.separator = "|",
          notes = "95\\% Confidence Intervals in parentheses",
          single.row = F,
          omit.stat = c("adj.rsq", "aic", "bic"),
          #star.cutoffs = c(.05, .01, .001), star.char = c("*","**","***"),
          add.lines = list(c("SD(Tract)", .013, .674),
                          c("SD(Residual)",  .094, "")),
          notes.append = F)
```

Spatial Panel Models

```
library(splm)
library(plm)
```

```
##
## Attaching package: 'plm'

## The following object is masked from 'package:lavaan':
##
##     nobs

## The following objects are masked from 'package:dplyr':
##
##     between, lag, lead
```

```
nb <- st_contiguity(mpls_tract, queen=TRUE)
wt <-  st_weights(nb, style = "W", allow_zero = T)
w <- recreate_listw(nb, wt)

cj_exp_prepost <- cj_exp_prepost %>%
  group_by(GEOID) %>%
  arrange(year, week) %>%
  mutate(WEEKID = row_number())
```

Table 3: Interrupted Time Series Models of Carjackings, MPLS 2017-2022

	Carjacking	
	Rate per 1,000 RE HLM	Any Carjacking RE Logit
	(1)	(2)
T	0.00001	0.004
	(−0.00001|0.00004)	(0.002|0.006)
Post-Killing	0.027	1.622
	(0.023|0.031)	(1.364|1.880)
T Post-Killing	−0.00002	−0.003
	(−0.0001|0.00003)	(−0.006|−0.0004)
Conc. Dis.	0.001	0.639
	(−0.003|0.004)	(0.425|0.853)
Age 19-29	−0.0004	−0.001
	(−0.001|−0.0001)	(−0.018|0.017)
Age 30-49	−0.0004	0.015
	(−0.001|0.0001)	(−0.013|0.042)
Age 50-69	−0.001	−0.046
	(−0.002|−0.0004)	(−0.079|−0.013)
Age 70+	−0.001	−0.018
	(−0.001|0.0001)	(−0.057|0.022)
Post-Killing X Conc. Dis.	0.016	−0.199
	(0.013|0.019)	(−0.347|−0.052)
T Post-Killing X Conc. Dis.	−0.00004	−0.001
	(−0.0001|0.00000)	(−0.003|0.0002)
Constant	0.050	−4.589
	(0.020|0.081)	(−6.283|−2.896)
SD(Tract)	0.013	0.674
SD(Residual)	0.094	
Observations	38,357	38,357
Log Likelihood	35,715.190	−5,578.143
Note:	95% Confidence Intervals in parentheses	

```r
#write_csv(cj_exp_prepost, "Car Jacking/cj_exp_prepost.csv")

cj_exp_prepost_panel <- pdata.frame(cj_exp_prepost, index = c("GEOID", "WEEKID"), drop.index = F)

slmtest(car_jack_rate~t+post_floyd+t_post_floyd+
             conc_dis+
             age_19_29_perc+age_30_49_perc+age_50_69_perc+
             age_70_plus_perc+ post_floyd:conc_dis+t_post_floyd:conc_dis,
        data=cj_exp_prepost_panel,
        listw =w,
        model = "random",
        test = "lml")
```

```
##
##  LM test for spatial lag dependence
##
## data:  formula (random transformation)
## LM = 2.4208, df = 1, p-value = 0.1197
## alternative hypothesis: spatial lag dependence
```

```r
#SAR Linear CJ
sar <- spml(car_jack_rate~t+post_floyd+t_post_floyd+
             conc_dis+
             age_19_29_perc+age_30_49_perc+age_50_69_perc+
             age_70_plus_perc+ post_floyd:conc_dis+t_post_floyd:conc_dis,
                      data = cj_exp_prepost,
             index = c("GEOID"),
            effect="individual",
             model="random",
             listw = w,
            lag=T,
            spatial.error="none")
summary(sar)
```

```
## Warning in sqrt(diag(object$vcov.arcoef)): NaNs produced

## ML panel with spatial lag, random effects
##
## Call:
## spreml(formula = formula, data = data, index = index, w = listw2mat(listw),
##     w2 = listw2mat(listw2), lag = lag, errors = errors, cl = cl)
##
## Residuals:
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.12754 -0.02268 -0.01054  0.00023 -0.00324  3.03511
##
## Error variance parameters:
##      Estimate Std. Error t-value Pr(>|t|)
## phi 0.0174063  0.0023195  7.5042 6.18e-14 ***
##
## Spatial autoregressive coefficient:
##        Estimate Std. Error t-value Pr(>|t|)
## lambda 0.012598        NaN     NaN      NaN
##
## Coefficients:
```

```
##                           Estimate  Std. Error t-value  Pr(>|t|)
## (Intercept)             5.0130e-02  1.5194e-02  3.2992 0.0009695 ***
## t                       1.4765e-05  1.2371e-05  1.1934 0.2326970
## post_floyd              2.6473e-02  1.9532e-03 13.5537 < 2.2e-16 ***
## t_post_floyd           -1.7419e-05  2.2365e-05 -0.7788 0.4360735
## conc_dis                6.0435e-04  1.8055e-03  0.3347 0.7378228
## age_19_29_perc         -4.4143e-04  1.5967e-04 -2.7647 0.0056983 **
## age_30_49_perc         -3.7533e-04  2.4992e-04 -1.5018 0.1331503
## age_50_69_perc         -9.9901e-04  2.9524e-04 -3.3837 0.0007151 ***
## age_70_plus_perc       -6.6219e-04  3.6087e-04 -1.8349 0.0665137 .
## post_floyd:conc_dis     1.6419e-02  1.5572e-03 10.5436 < 2.2e-16 ***
## t_post_floyd:conc_dis -3.5298e-05  1.8129e-05 -1.9470 0.0515311 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#SAR LPM CJ
sar_lpm <- spml(anyjack~t+post_floyd+t_post_floyd+
             conc_dis+
             age_19_29_perc+age_30_49_perc+age_50_69_perc+
             age_70_plus_perc+ post_floyd:conc_dis+t_post_floyd:conc_dis,
                      data = cj_exp_prepost,
             index = c("GEOID"),
            effect="individual",
             model="random",
             listw = w,
            lag=T,
            spatial.error="none")
summary(sar_lpm)
```

```
## ML panel with spatial lag, random effects
##
## Call:
## spreml(formula = formula, data = data, index = index, w = listw2mat(listw),
##     w2 = listw2mat(listw2), lag = lag, errors = errors, cl = cl)
##
## Residuals:
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.22452 -0.05778 -0.02641  0.00128 -0.00784  1.01973
##
## Error variance parameters:
##      Estimate Std. Error t-value  Pr(>|t|)
## phi 0.0202918  0.0029469  6.8858 5.748e-12 ***
##
## Spatial autoregressive coefficient:
##         Estimate Std. Error t-value  Pr(>|t|)
## lambda 0.0313151  0.0063409  4.9386 7.869e-07 ***
##
## Coefficients:
##                           Estimate  Std. Error t-value  Pr(>|t|)
## (Intercept)             6.9039e-02  3.2736e-02  2.1090  0.034948 *
## t                       4.9001e-05  2.4969e-05  1.9625  0.049708 *
## post_floyd              5.4693e-02  3.9421e-03 13.8741 < 2.2e-16 ***
## t_post_floyd           -2.9460e-05  4.5139e-05 -0.6526  0.513982
## conc_dis                3.7377e-03  3.8788e-03  0.9636  0.335240
## age_19_29_perc         -4.8123e-04  3.4413e-04 -1.3984  0.161999
```

```
## age_30_49_perc        -1.2904e-04  5.3865e-04 -0.2396  0.810675
## age_50_69_perc        -1.9176e-03  6.3633e-04 -3.0135  0.002583 **
## age_70_plus_perc      -1.1173e-03  7.7779e-04 -1.4366  0.150840
## post_floyd:conc_dis    2.4504e-02  3.1430e-03  7.7964  6.37e-15 ***
## t_post_floyd:conc_dis -7.9872e-05  3.6590e-05 -2.1829  0.029045 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```