Supplemental Material for **Exposure to community homicide during pregnancy and adverse birth outcomes: A within-community matched design**

Covariate categories

We categorized maternal age as younger than 20, 20-24, 25-29, 30-34, and 35 or older. Educational attainment was categorized as less than high school, high school graduate, some college or associate's degree, bachelor's degree, graduate degree, or missing. Insurance types were no prenatal care, Medi-Cal without Comprehensive Perinatal Services Program (CPSP) services, other government insurance, private insurance, self-pay, Medi-Cal with CPSP services, and other. Parity was categorized as 1, 2, or 3 or more.

Data access

Interested readers who wish to access the data used in this study may do so by requesting it from California's Department of Public Health (see application form at https://www.cdph.ca.gov/Programs/CHSI/Pages/Data-Applications.aspx).
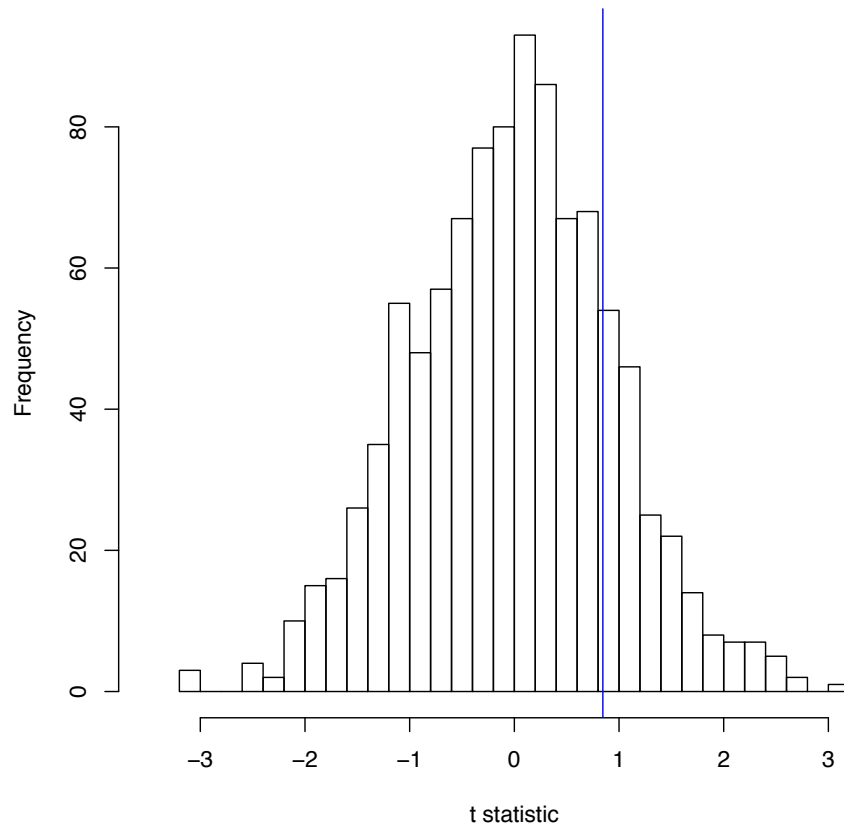
Randomization inference

Randomization inference was used as an additional check to assess the likelihood of our results occurring due to random error. Randomization inference is a type of permutation test in which the exposure values are randomly shuffled across individuals, allowing us to compare the observed association with the distribution of associations that occur due to chance. We performed the randomization inference by permuting the homicide exposure values within Census tracts, rerunning the weighting and regression analysis, and recording the t-statistic. This was done 1,000 times to generate a distribution of the t-statistic under the null hypothesis of no relationship between homicide exposure and adverse birth outcomes. We then compared the observed t-statistic to the null distribution to assess the likelihood that our results could be attributed to random error.

Supplemental Figure 1. Randomization inference for homicide exposure during the pre-conception period on risk of preterm birth.
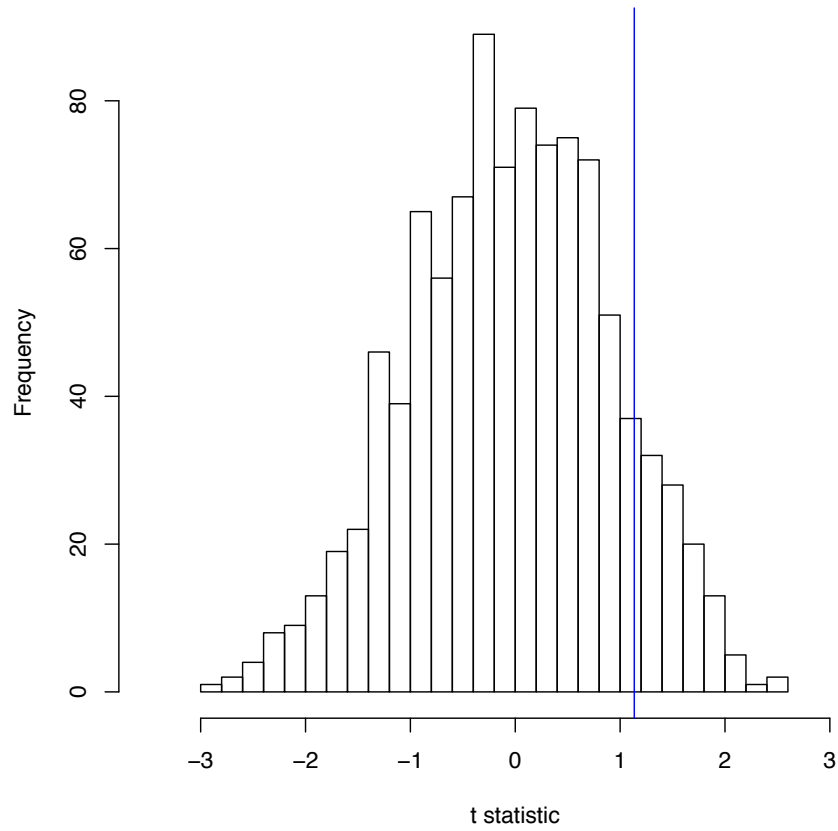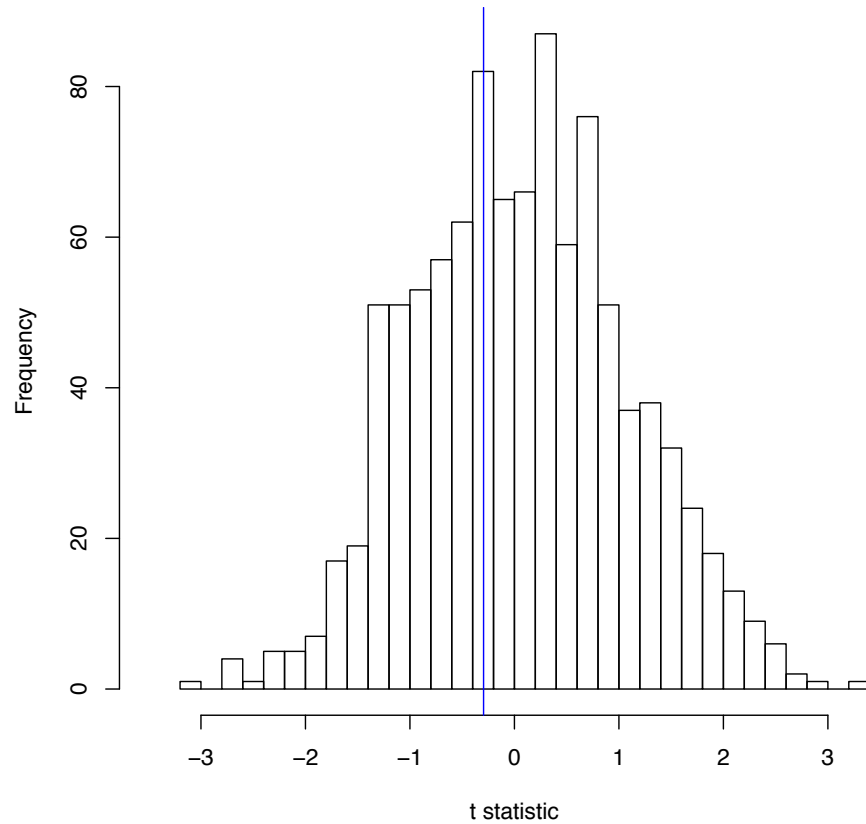

Note: The blue vertical line is the t statistic for the observed association of pre-conception homicide exposure and risk of preterm birth.

Supplemental Figure 2. Randomization inference for homicide exposure during the first trimester on risk of preterm birth.

Note: The blue vertical line is the t statistic for the observed association of first trimester homicide exposure and risk of preterm birth.

Supplemental Figure 3. Randomization inference for homicide exposure during the second trimester on risk of preterm birth.
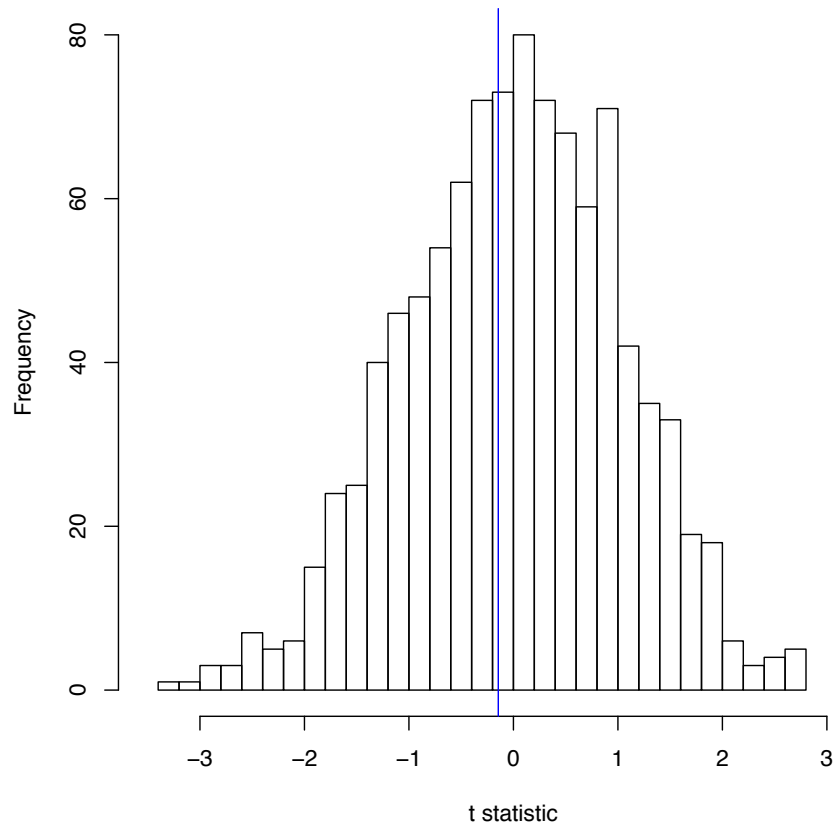

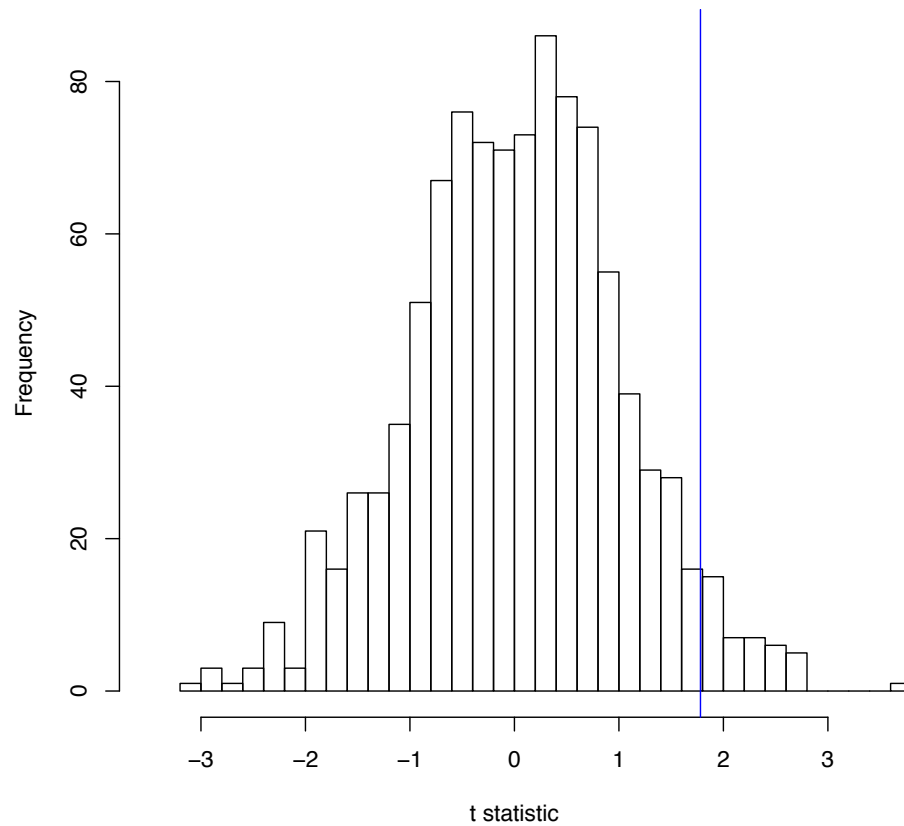Note: The blue vertical line is the t statistic for the observed association of second trimester homicide exposure and risk of preterm birth.

Supplemental Figure 4. Randomization inference for homicide exposure during the pre-conception period on risk of SGA.
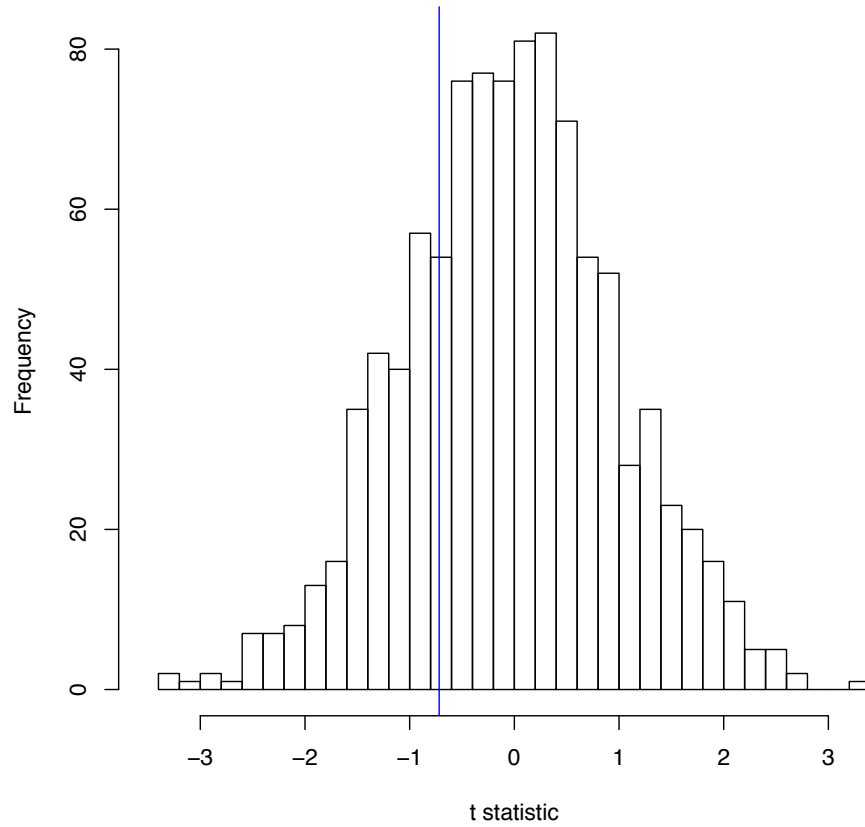

Note: The blue vertical line is the t statistic for the observed association of pre-conception homicide exposure and risk of SGA.

Supplemental Figure 5. Randomization inference for homicide exposure during the first trimester on risk of SGA.

Note: The blue vertical line is the t statistic for the observed association of first trimester homicide exposure and risk of SGA.

Supplemental Figure 6. Randomization inference for homicide exposure during the second trimester on risk of SGA.

Note: The blue vertical line is the t statistic for the observed association of second trimester homicide exposure and risk of SGA.

Supplemental Figure 7. Randomization inference for three or more homicides during the pre-conception period, first trimester, or second trimester on risk of preterm birth.

Note: The blue vertical line is the t statistic for the observed association of exposure to three or more homicides and risk of preterm birth.

Supplemental Figure 8. Randomization inference for three or more homicides during the pre-conception period, first trimester, or second trimester on risk of SGA.
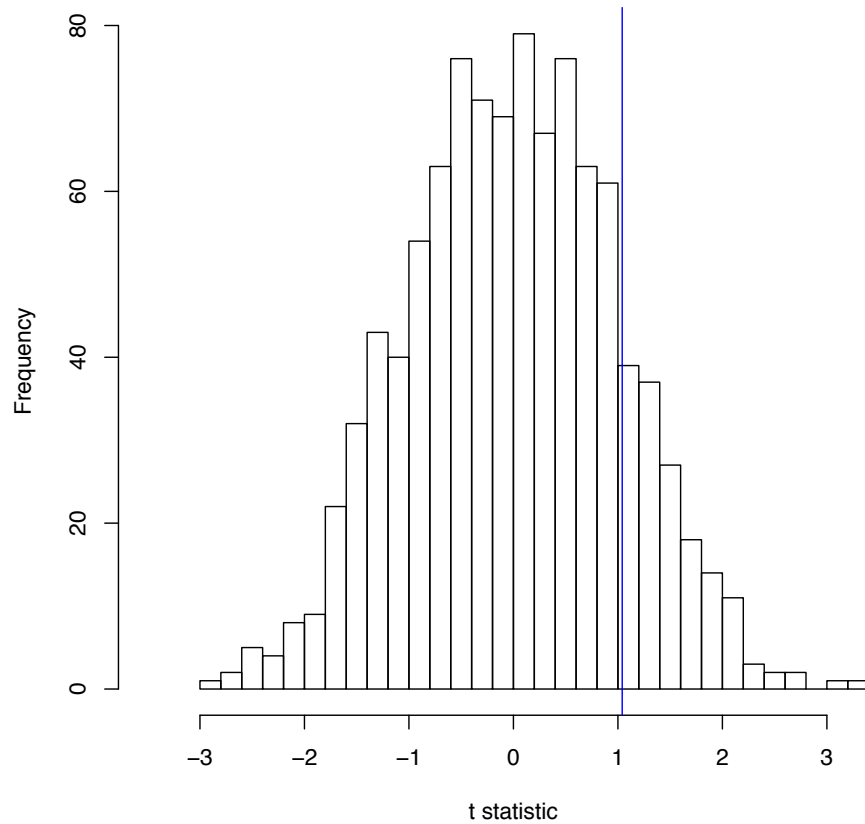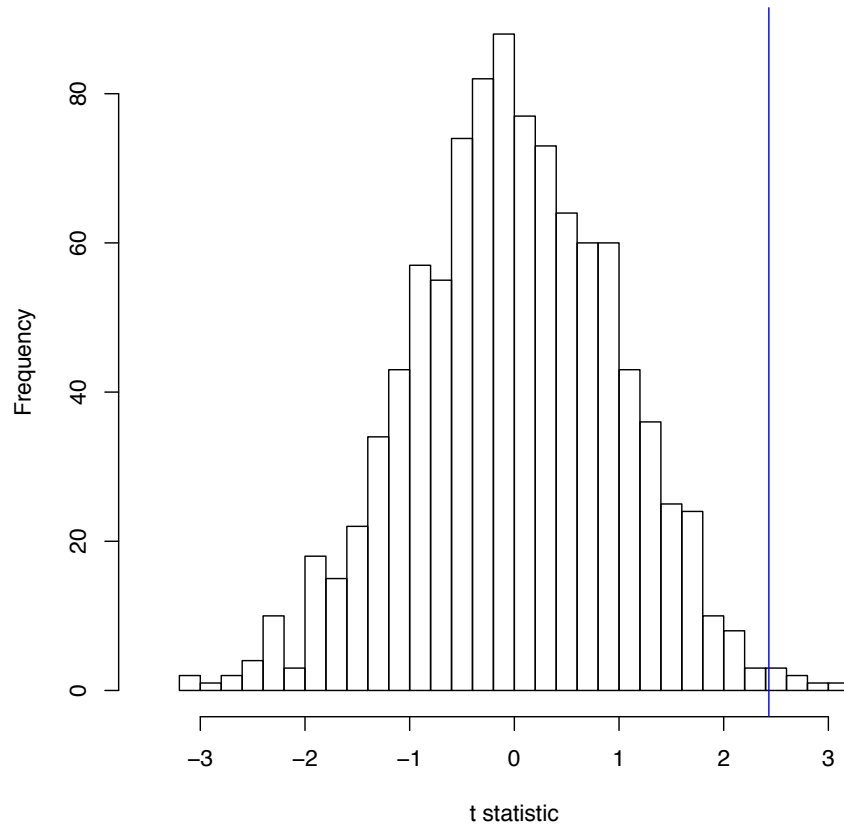

Note: The blue vertical line is the t statistic for the observed association of exposure to three or more homicides and risk of SGA.

Supplemental Figure 9. Randomization inference for exposure during the pre-conception period, first trimester, and the second trimester on risk of SGA.

Note: The blue vertical line is the t statistic for the observed association of exposure to homicides during all three exposure windows and risk of SGA.

Supplemental Figure 10. Randomization inference for three or more homicides during the pre-conception period, first trimester, or second trimester on risk of preterm birth.
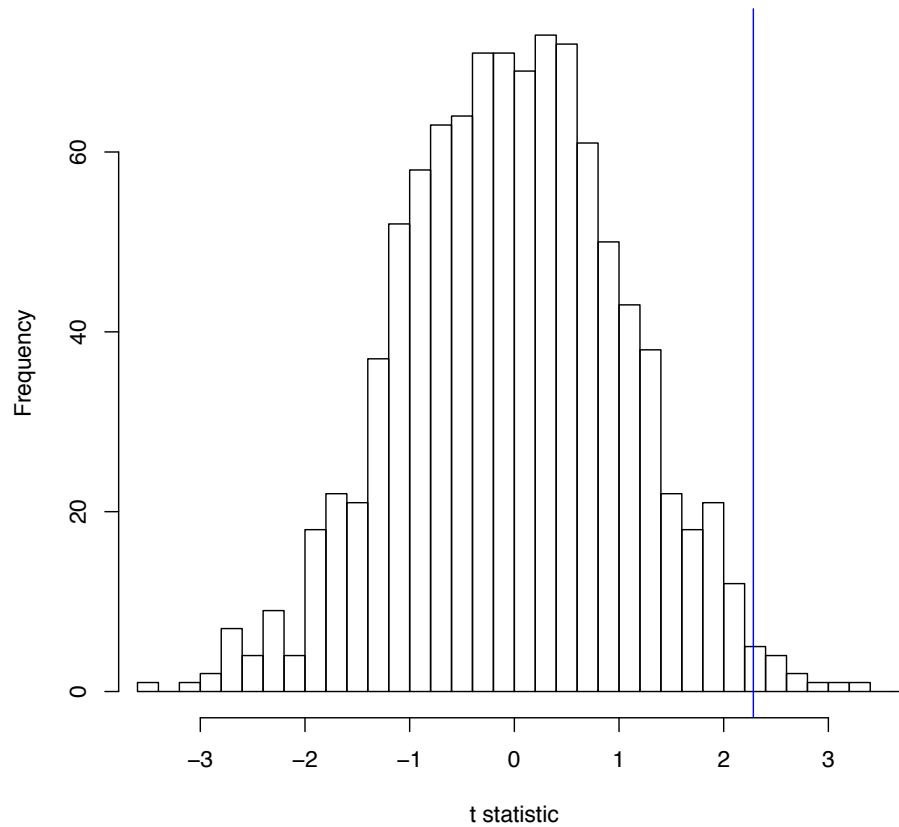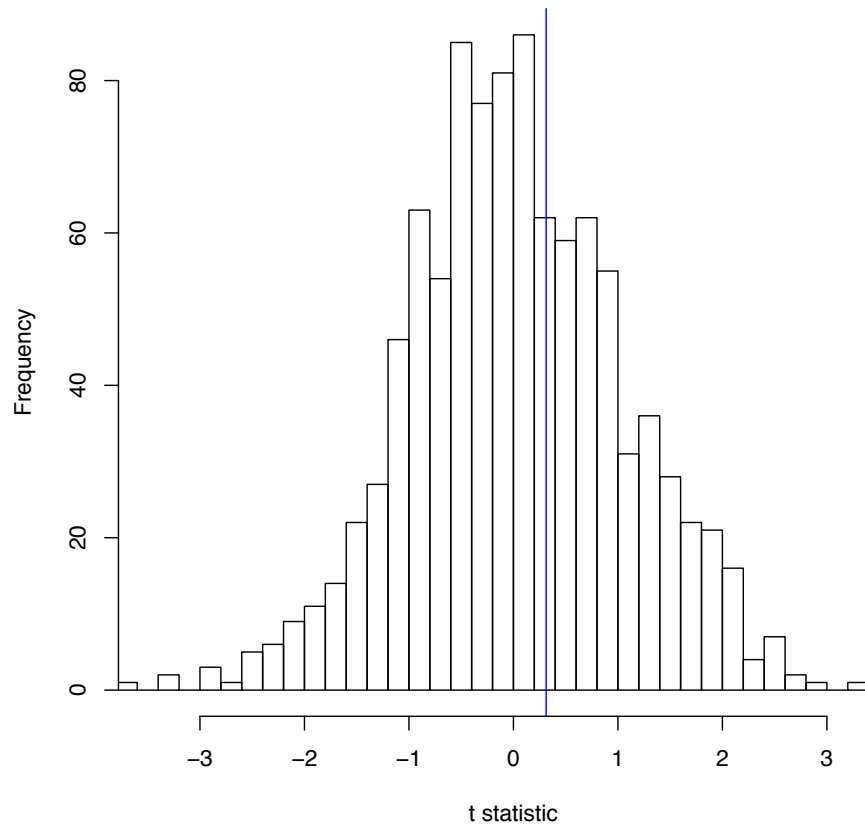

Note: The blue vertical line is the t statistic for the observed association of exposure to homicides during all three exposure windows and risk of preterm birth.

Supplemental Table 1. Randomization inference p-values.

Exposure period	Outcome	P-value from randomization inference
Pre-conception period	Preterm birth	0.109
	SGA	0.569
First trimester	Preterm birth	0.182
	SGA	0.043
Second trimester	Preterm birth	0.630
	SGA	0.747
Exposure to three or more homicides during the pre-conception period, first trimester, and second trimester	Preterm birth	0.148
	SGA	0.006
Exposure to homicide in the pre-conception period, the first trimester, and the second trimester	Preterm birth	0.366
	SGA	0.015

Sensitivity analyses

Due to the high level of missingness of education attainment in the birth record, we excluded it as a covariate from the main analysis. However, we reran our results while controlling for educational attainment, with a missingness indicator. These results are below.

Supplemental Figure 11. Estimated average treatment effect on the treated (ATT) of homicide exposure (adjusting for educational attainment with a missing category) on risk of preterm birth during the pre-conception period (1-12 weeks before conception), first trimester (weeks 0-12), second trimester (weeks 13-22), and the negative controls (0-12 weeks and 13-24 weeks after the pseudo-birth week) in California, 2007-2011.
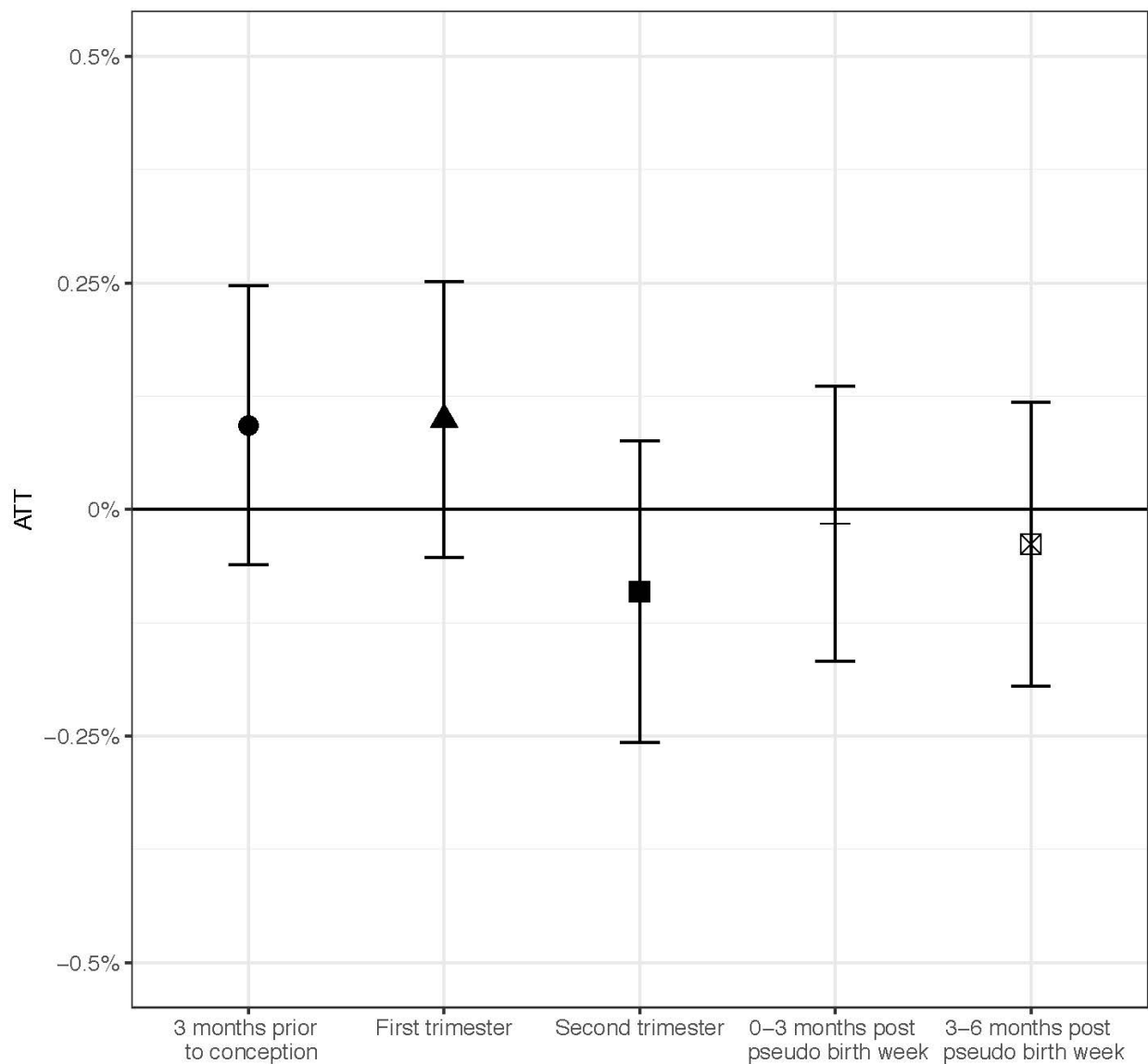
Supplemental Figure 12. Estimated average treatment effect on the treated (ATT) of homicide exposure (adjusting for educational attainment with a missing category) on risk of SGA during the pre-conception period (1-12 weeks before conception), first trimester (weeks 0-12), second trimester (weeks 13-22), and the negative controls (0-12 weeks and 13-24 weeks after the pseudo-birth week) in California, 2007-2011.
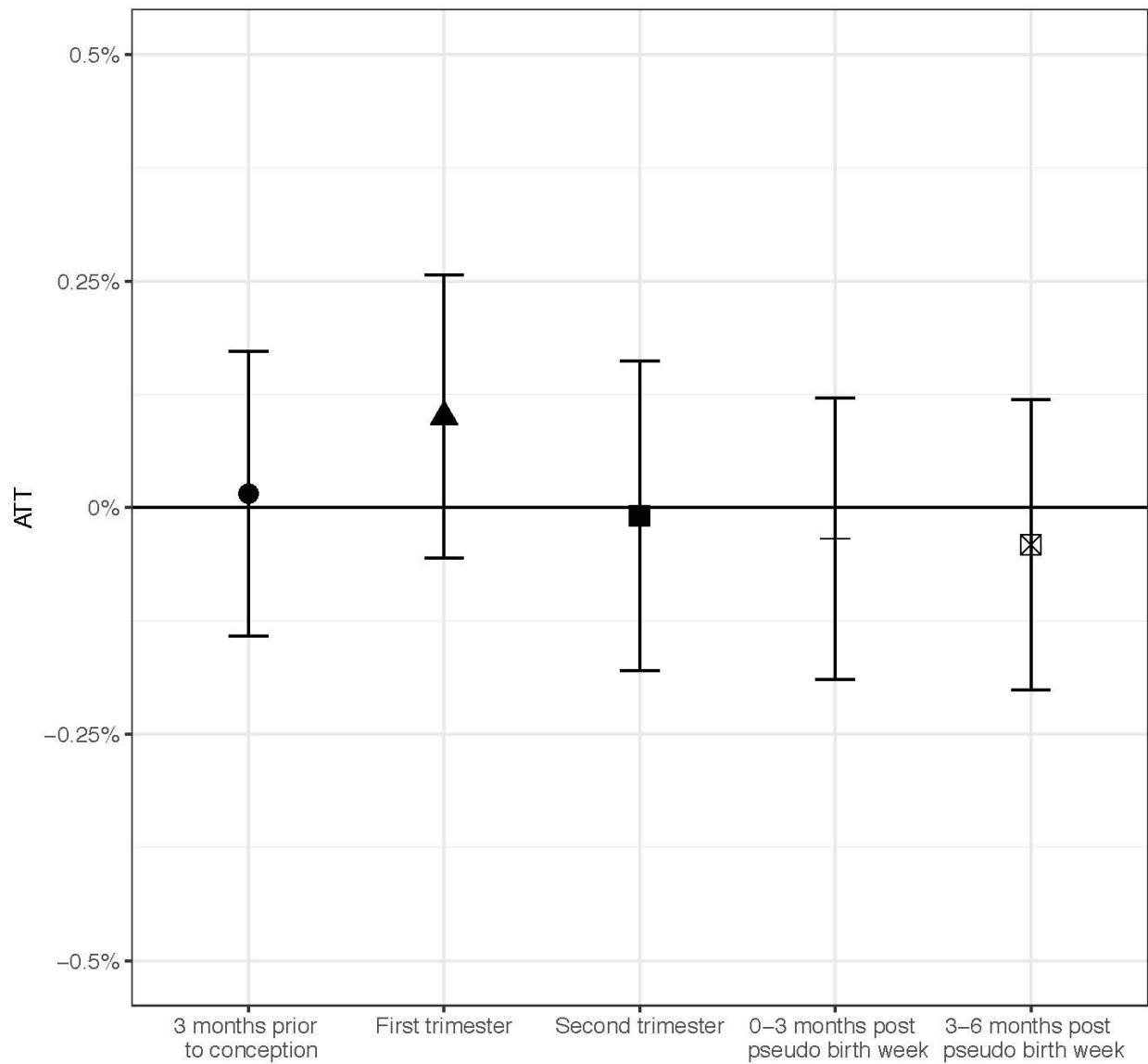
Supplemental Figure 13. Estimated average treatment effect on the treated (ATT) of homicide exposure during all three exposure windows and exposed to at least three homicides (adjusting for educational attainment with a missing category) on risk of preterm birth and SGA in California, 2007-2011.
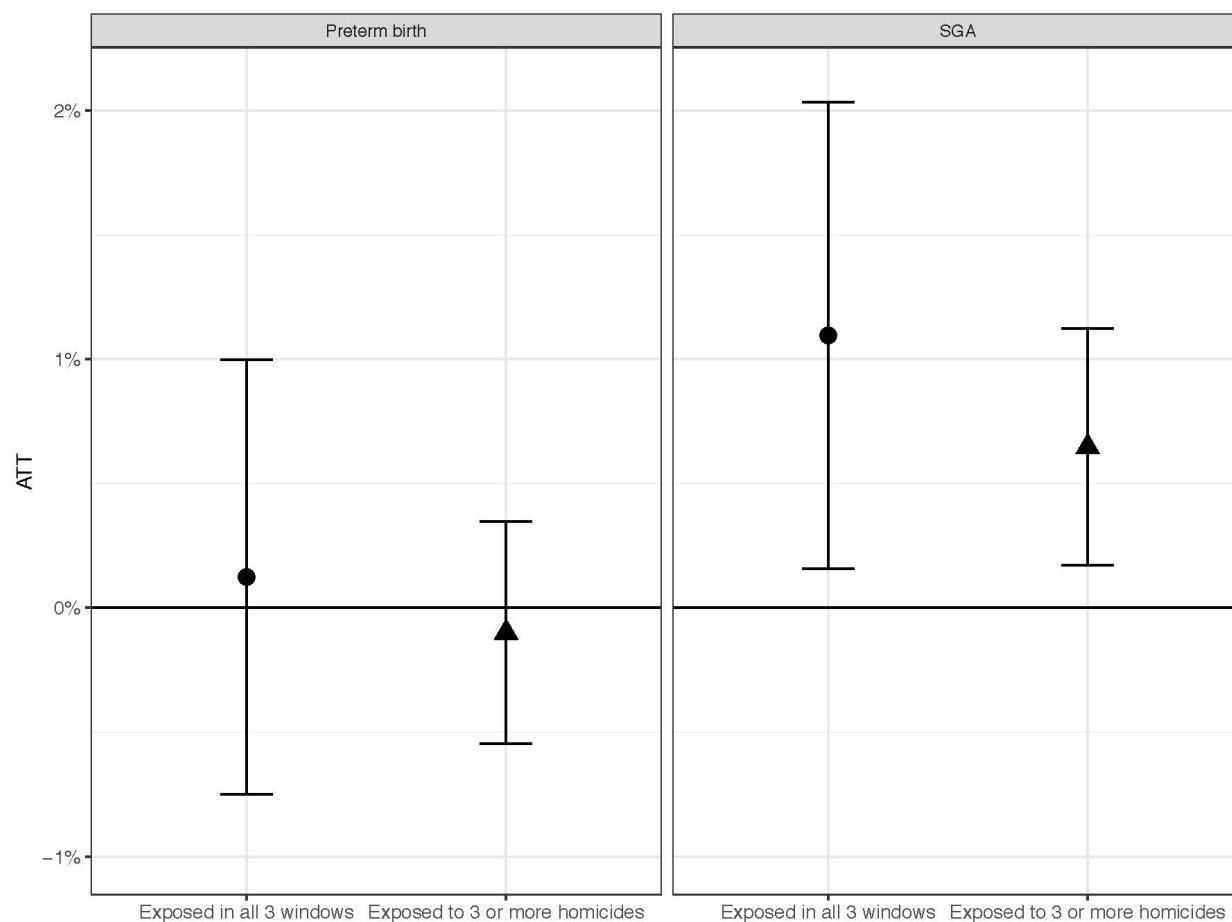
R code

```
rm(list=ls())
library(foreign)
library(ggplot2)
library(doBy)
library(dplyr)
library(mice)

set.seed(244534)

#------------------------------------------------------------------------
------
#      read in data
#------------------------------------------------------------------------
------
memory.limit(1e+13)
files <- list()
for (y in 2007:2011) {
  print(y)
    files[[y - 2007 +1 ]] <- read.csv(paste0("…/birth_hom_",y,".csv"),
colClasses="character")
}


vars <- data.frame(varnames = c(names(files[[1]]),names(files[[2]]),
names(files[[3]]), names(files[[4]]), names(files[[5]])))
kvars <- dplyr::count(vars, vars=varnames)
kvars <- kvars[kvars$n==5,]
kvars <- as.character(kvars$vars)

for (yr in 2007:2011) {
  print(yr)
  files[[yr - 2007 +1]] <- files[[yr - 2007 +1]][,colnames(files[[yr -
2007+1]]) %in% kvars]
}

file <- do.call(rbind, files)

# assign correct class to variables
file$year <- as.numeric(file$year)
file$tract <- ifelse(file$tract==".",NA,file$tract)
file$gestwks <- as.numeric(file$pgstest)
file$ptb <- ifelse(file$gestwks<22,NA, ifelse(file$gestwks>44,NA,
ifelse(file$gestwks<37,1,0)))
file$pparity <- as.numeric(file$pparity)
file$bthdate <- as.Date(file$nbthdate, "%y-%m-%d")

# remove people whose conception dates are outside range
file$condate <- file$bthdate - (file$gestwks*7)
file$con_month <- format(as.Date(file$condate),"%Y-%m")
# adjust for fixed cohort effect by
# remove any with conception dates before it was possible for early delivery
to have been included
```

```
#      in cohort or after it was possible for term delivery to be included

dim(file)
file <- file %>% filter(condate> as.Date("01jan2007", format="%d %b %Y") &
condate < as.Date("01mar2011",  format="%d %b %Y"))
dim(file)




file$hom_post3m_n <- file$hom_post3m
file$hom_post3m <-
ifelse(file$hom_post3m_n>=1,1,ifelse(file$hom_post3m_n==0,0,NA))

file$hom_post6m_n <- file$hom_post6m
file$hom_post6m <-
ifelse(file$hom_post6m_n>=1,1,ifelse(file$hom_post6m_n==0,0,NA))

file$hom_pre3m_n <- file$hom_pre3m
file$hom_pre3m <-
ifelse(file$hom_pre3m_n>=1,1,ifelse(file$hom_pre3m_n==0,0,NA))

file$hom_tri1_n <- file$hom_tri1
file$hom_tri1 <- ifelse(file$hom_tri1_n>=1,1,ifelse(file$hom_tri1_n==0,0,NA))

file$hom_tri2_n <- file$hom_tri2
file$hom_tri2 <- ifelse(file$hom_tri2_n>=1,1,ifelse(file$hom_tri2_n==0,0,NA))


# calculate conception season

getSeason <- function(DATES) {
    WS <- as.Date("2012-12-15", format = "%Y-%m-%d") # Winter Solstice
    SE <- as.Date("2012-3-15",  format = "%Y-%m-%d") # Spring Equinox
    SS <- as.Date("2012-6-15",  format = "%Y-%m-%d") # Summer Solstice
    FE <- as.Date("2012-9-15",  format = "%Y-%m-%d") # Fall Equinox

    # Convert dates from any year to 2012 dates
    d <- as.Date(strftime(DATES, format="2012-%m-%d"))

    ifelse (d >= WS | d < SE, "Winter",
      ifelse (d >= SE & d < SS, "Spring",
        ifelse (d >= SS & d < FE, "Summer", "Fall")))
}

file$con_season <- getSeason(as.Date(file$condate, format="%Y-%m-%d"))

file$conwk <- as.numeric(file$conwk)
file$conwk <- ifelse(file$conwk<= -52, NA, file$conwk)


# define race/ethnicity

# white : mracem=1
# black : mracem=2
# alaskan native / american indian: mracem=3
# asian  : mracem=4
```

```
# hawaiian / pi: mracem=5
# other: mracem=6
# two or more: mracem=7

# not hispanic: msporig=1
# mexican/chicano: msporig=2
# puerto rican: msporig=3
# cuban: msporig=4
# central/south american: msporig=5
# other hispanic born outside us: msporig= 6
# code not used: msporig=7
# other hispanic born in us: msporig=8

file$whitenh <- as.numeric(file$mracem==1 & file$msporig==1)
file$blacknh <- as.numeric(file$mracem==2 & file$msporig==1)
file$akainh <- as.numeric(file$mracem==3 & file$msporig==1)
file$asiannh <- as.numeric(file$mracem==4 & file$msporig==1)
file$hipinh <- as.numeric(file$mracem==5 & file$msporig==1)
file$othertwonh <- as.numeric((file$mracem==6 | file$mracem==7) &
file$msporig==1 )
file$hispanic <- as.numeric(file$msporig>1 )
file$mexican <- as.numeric(file$msporig==2)
file$puertorican <- as.numeric(file$msporig==3 )
file$cuban <- as.numeric(file$msporig==4)
file$csamerican <- as.numeric(file$msporig==5 )
file$ohforeign <- as.numeric(file$msporig==6 )
file$ohus <- as.numeric(file$msporig==8)

file$race <- ifelse(file$whitenh==1,1,
                ifelse(file$blacknh==1,2,
                    ifelse(file$akainh==1,3,
                        ifelse(file$asiannh==1,4,
                            ifelse(file$hipinh==1,5,
                                ifelse(file$othertwonh==1,6,

ifelse(file$hispanic==1,7,NA)))))))

# create 5-yr age groups
file$mage <- as.numeric(file$mage)
file$agecat <- ifelse(file$mage<20,1,ifelse(file$mage>=20 &
file$mage<25,2,ifelse(file$mage>=25 & file$mage<30,3,ifelse(file$mage>=30 &
file$mage<35,4, ifelse(file$mage>=35,5,NA)))))
file$educ <-
ifelse(file$meduc<2,1,ifelse(file$meduc==3,2,ifelse(file$meduc==4 |
file$meduc==5,3,ifelse(file$meduc==6,4,ifelse(file$meduc>=7 &
file$meduc<9,5,NA)))))

# ppncpay  -- principal source of payment for prenatal care
# 0 = no prenatal care
# 2 = medical, without cpsp
# 5 = other government program
# 7 = private insurance
# 9 = self-pay
# 13 = medical, with cpsp
# 14 = other insurance
# 99 = unknown
```

```r
file$ppncpay <- ifelse(file$ppncpay==".m",NA,file$ppncpay)

# create categories of parity
file$parity <-
ifelse(file$pparity<=1,1,ifelse(file$pparity==2,2,ifelse(file$pparity>=3,3,NA
)))

# create SGA based on Talge et al 2014
file$birthwt <- as.numeric(file$nbthwt1)
file$sga <- ifelse(file$nsex=="1" & file$gestwks==22 & file$birthwt<=375,1,
                   ifelse(file$nsex=="1" & file$gestwks==23 &
file$birthwt<=436,1,
                   ifelse(file$nsex=="1" & file$gestwks==24 &
file$birthwt<=497,1,
                   ifelse(file$nsex=="1" & file$gestwks==25 &
file$birthwt<=561,1,
                   ifelse(file$nsex=="1" & file$gestwks==26 &
file$birthwt<=629,1,
                   ifelse(file$nsex=="1" & file$gestwks==27 &
file$birthwt<=706,1,
                   ifelse(file$nsex=="1" & file$gestwks==28 &
file$birthwt<=802,1,
                   ifelse(file$nsex=="1" & file$gestwks==29 &
file$birthwt<=924,1,
                   ifelse(file$nsex=="1" & file$gestwks==30 &
file$birthwt<=1068,1,
                   ifelse(file$nsex=="1" & file$gestwks==31 &
file$birthwt<=1231,1,
                   ifelse(file$nsex=="1" & file$gestwks==32 &
file$birthwt<=1415,1,
                   ifelse(file$nsex=="1" & file$gestwks==33 &
file$birthwt<=1627,1,
                   ifelse(file$nsex=="1" & file$gestwks==34 &
file$birthwt<=1859,1,
                   ifelse(file$nsex=="1" & file$gestwks==35 &
file$birthwt<=2105,1,
                   ifelse(file$nsex=="1" & file$gestwks==36 &
file$birthwt<=2355,1,
                   ifelse(file$nsex=="1" & file$gestwks==37 &
file$birthwt<=2588,1,
                   ifelse(file$nsex=="1" & file$gestwks==38 &
file$birthwt<=2782,1,
                   ifelse(file$nsex=="1" & file$gestwks==39 &
file$birthwt<=2926,1,
                   ifelse(file$nsex=="1" & file$gestwks==40 &
file$birthwt<=3017,1,
                   ifelse(file$nsex=="1" & file$gestwks==41 &
file$birthwt<=3065,1,
                   ifelse(file$nsex=="1" & file$gestwks==42 &
file$birthwt<=3082,1,
                   ifelse(file$nsex=="1" & file$gestwks==43 &
file$birthwt<=3067,1,
                   ifelse(file$nsex=="1" & file$gestwks==44 &
file$birthwt<=3027,1,
                   ifelse(file$nsex=="2" & file$gestwks==22 &
file$birthwt<=354 ,1,
```

```
                  ifelse(file$nsex=="2" & file$gestwks==23 &
file$birthwt<=416 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==24 &
file$birthwt<=473 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==25 &
file$birthwt<=529 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==26 &
file$birthwt<=597 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==27 &
file$birthwt<=677 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==28 &
file$birthwt<=770 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==29 &
file$birthwt<=882 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==30 &
file$birthwt<=1018 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==31 &
file$birthwt<=1166 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==32 &
file$birthwt<=1335 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==33 &
file$birthwt<=1538 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==34 &
file$birthwt<=1772 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==35 &
file$birthwt<=2021 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==36 &
file$birthwt<=2261 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==37 &
file$birthwt<=2477 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==38 &
file$birthwt<=2665 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==39 &
file$birthwt<=2810 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==40 &
file$birthwt<=2904 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==41 &
file$birthwt<=2958 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==42 &
file$birthwt<=2985 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==43 &
file$birthwt<=2981 ,1,
                  ifelse(file$nsex=="2" & file$gestwks==44 &
file$birthwt<=2952 ,1,0))))))))))))))))))))))))))))))))))))))))))))))))


df <- file %>% dplyr::select(year,county,tract,nbthdate,bthdate,
              agecat,educ,ppncpay,race,bcf_recnum,
              birthwk,conwk,conyr,con_month,con_season,hom_tri1,hom_tri2,

gestwks,ptb,sga,hom_post3m,hom_post6m,hom_pre3m,parity,hom_pre3m_n,
              hom_tri1_n, hom_tri2_n, hom_post3m_n, hom_post6m_n)

dim(df)
# 2184209
```

```
dim(df)
df <- df %>% filter(!is.na(tract))
dim(df)
# removes the 32 people with the incorrect county and therefore missing tract
save(df, file="…/fulldata")


df_e <- df[,!colnames(df) %in% "educ"]

dfcc_e <- df_e[complete.cases(df_e),]
#2133662
save(dfcc_e, file="…fulldata_cc_e")



dfcc <- df[complete.cases(df),]
dim(dfcc)
#1784585

save(dfcc, file="…/fulldata_cc")



#-----------------------------------------------------------------------------
------
#      create matched data set for hom_tri1 just on tract
#-----------------------------------------------------------------------------
------
rm(list=ls())
library(doBy)
library(dplyr)

load("…/fulldata_cc_e")


# match on tract and conception year
df.s1 <- summaryBy(hom_tri1 ~ county + tract, data=dfcc_e, FUN=mean)
df.s2 <- summaryBy(hom_tri1 ~  county + tract, data=dfcc_e, FUN=sum)
df.s3 <- summaryBy(hom_tri1 ~  county + tract, data=dfcc_e, FUN=length)

df.s <- left_join(df.s1,df.s2, by=c("county","tract"))
df.s <- left_join(df.s, df.s3, by=c("county","tract"))
df.s$cw <- df.s$hom_tri1.sum/(df.s$hom_tri1.length - df.s$hom_tri1.sum)


df.t <- left_join(dfcc_e,df.s, by=c("county","tract"))

df.sk <- df.s[df.s$hom_tri1.mean>0 & df.s$hom_tri1.mean<1,]


df.m.z <- left_join(dfcc_e,df.sk, by=c("county","tract"))
dim(df.m.z)
# remove anyone who doesn't have a value for hom_tri1.mean
# this is where there's no variation in treatment
df_tri1_a_e <- df.m.z[complete.cases(df.m.z),]
dim(df_tri1_a_e)
# weight should be 1 for each treated unit
```

```
df_tri1_a_e$cw[df_tri1_a_e$hom_tri1==1] <- 1

save(df_tri1_a_e, file="…/hom_tri1_a_e")

#-----------------------------------------------------------------------
------
#      calculate ATT for  hom_tri1
#-----------------------------------------------------------------------
------

att_est <- function(outcome) {
mfit <- glm(get(outcome) ~ hom_tri1 + factor(conyr) + factor(con_season) +
factor(race) +
               factor(agecat) + factor(ppncpay) + factor(parity),
data=df_tri1_a_e, weights=cw)

var <- vcovHC(mfit, type="HC0")[2,2]

tri1_att <- data.frame(cbind(coef(summary(mfit))[2], var))
rownames(tri1_att) <- "tri1_att"

colnames(tri1_att) <- c("estimate","var")
tri1_att$lb <- tri1_att$estimate - 1.96*sqrt(tri1_att$var)
tri1_att$ub <- tri1_att$estimate + 1.96*sqrt(tri1_att$var)

return(tri1_att)
}
```