

ARE 213 Problem Set 1b

Nick Depsky, Will Gorman, Peter Worley

October 12, 2018

Import Data

```
#setwd("~/Dropbox/Berkeley_tings/Fall 2018/ARE213/Problem Sets/SharedFiles/are213/PS1b")
dat <- read.dta("ps1.dta")
#fix missing data
dat_drop <- dat %>% filter(herpes != 8 & tobacco != 9 & cigar != 99 & cigar6 != 6 &
                           alcohol != 9 & drink != 99 & drink5 != 5 & wgain != 99)
dat_drop$tobacco[dat_drop$tobacco == 2] <- 0
setwd("/Users/nicholasdepsky/Dropbox/Berkeley_tings/Fall 2018/ARE213/Problem Sets/SharedFiles/are213/PS1b")
load("white_df_ps1.Rdata")
load("vartab.Rdata")
dat.all <- cbind(dat_drop$dbrwt, white_df_ps1) %>% as.data.frame() %>% set_colnames(c("dbrwt",names(white_df_ps1)))
dat.mod <- dat_drop %>% select(dbrwt,stresfip,dimage,ormoth,mrace3,dmeduc,dmar,adequacy,
                                    dtotord,monpre,npervist,disllb,birmon,dgestat,csex,dplural,
                                    anemia,diabetes,herpes,chyper,pre4000,
                                    preterm,tobacco,cigar,alcohol,drink,wgain)

#dat.mod <- dat_drop %>% select(dbrwt,dimage,dmeduc,dmar,adequacy,
#                                    #dtotord,monpre,npervist,disllb,birmon,dgestat,csex,dplural,
#                                    #anemia,diabetes,herpes,chyper,pre4000,
#                                    #preterm,tobacco,cigar,alcohol,drink,wgain)

lm.mod <- lm(dbrwt ~ ., data = dat.mod)
summary(lm.mod)

##
## Call:
## lm(formula = dbrwt ~ ., data = dat.mod)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -2942.57  -289.14    -5.04   284.58  2741.78 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 6.420e+01 7.987e+01   0.804  0.42149  
## stresfip    9.177e-01 6.158e-01   1.490  0.13619  
## dimage      1.750e+00 3.134e-01   5.583 2.36e-08 *** 
## ormoth     -2.732e+01 2.606e+00 -10.485 < 2e-16 *** 
## mrace3     -6.862e+01 2.281e+00 -30.075 < 2e-16 *** 
## dmeduc      4.960e+00 7.031e-01   7.055 1.74e-12 *** 
## dmar       -4.816e+01 3.942e+00 -12.216 < 2e-16 *** 
## adequacy    1.011e+01 4.007e+00   2.524  0.01160 *  
## dtotord     8.995e+00 1.183e+00   7.606 2.84e-14 *** 
## monpre      1.564e+00 1.373e+00   1.139  0.25453  
## npervist    8.887e+00 5.195e-01  17.107 < 2e-16 *** 
## disllb     -1.951e-01 4.711e-03 -41.414 < 2e-16 ***
```

```

## birmon      -3.955e-01  3.928e-01  -1.007  0.31391
## dgestat     1.066e+02  5.862e-01  181.920  < 2e-16 ***
## csex        -1.329e+02  2.668e+00  -49.825  < 2e-16 ***
## dplural     -6.466e+02  7.965e+00  -81.177  < 2e-16 ***
## anemia      -2.601e+01  1.344e+01  -1.935  0.05296 .
## diabetes    -1.586e+02  8.325e+00  -19.047  < 2e-16 ***
## herpes       -1.500e+01  1.707e+01  -0.879  0.37945
## chyper       1.281e+02  1.531e+01   8.364  < 2e-16 ***
## pre4000     -3.768e+02  1.125e+01  -33.499  < 2e-16 ***
## preterm      2.528e+02  1.141e+01   22.166  < 2e-16 ***
## tobacco      -1.579e+02  6.569e+00  -24.033  < 2e-16 ***
## cigar         -3.805e+00  4.498e-01  -8.459  < 2e-16 ***
## alcohol      4.683e+01  1.585e+01   2.955  0.00313 **
## drink        -1.614e-01  2.494e+00  -0.065  0.94839
## wgain        8.257e+00  1.157e-01  71.353  < 2e-16 ***
##
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 450.9 on 114583 degrees of freedom
## Multiple R-squared:  0.4065, Adjusted R-squared:  0.4063
## F-statistic:  3018 on 26 and 114583 DF, p-value: < 2.2e-16

```

1a - Misspecification bias

There are a number of possible sources for misspecification bias in our linear model estimates from PS1a. For one, the control variables, including the treatment variable, are all assumed to have linear effects on the outcome variable, which may or may not be the correct functional form of the model, as some effects may in fact be non-linear, or there may exist important interaction effects between controls that were not considered. Secondly, it is possible that there still remain unobserved variables correlated with either our outcome or control variables that were omitted from the model, causing omitted variables bias. Perhaps most importantly, however, is the likelihood that there exists some issue of endogeneity between the outcome variable and some of the control. For example, smoking while pregnant may contribute to hypertension, cardiac issues, or other factors that were considered in our initial model as control variables independent of the outcome variable.

1b - Series Estimator, One Squared Term, One Interaction Term

```

dat.ser <- dat.mod
dat.ser$drinksq <- dat.ser$drink^2
dat.ser$cig_drink <- dat.ser$cigar*dat.ser$drink
poly.mod <- lm(dbrwt ~ ., data = dat.ser)
summary(poly.mod)

##
## Call:
## lm(formula = dbrwt ~ ., data = dat.ser)
##
## Residuals:
##      Min      1Q      Median      3Q      Max 
## -2942.37 -289.21    -4.99    284.62  2741.71 
## 
```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.530e+02 8.265e+01  1.851 0.06419 .
## stresfip    9.223e-01 6.158e-01  1.498 0.13422
## dimage     1.755e+00 3.134e-01  5.601 2.14e-08 ***
## ormooth   -2.733e+01 2.606e+00 -10.491 < 2e-16 ***
## mrace3     -6.853e+01 2.281e+00 -30.039 < 2e-16 ***
## dmeduc      4.937e+00 7.031e-01  7.022 2.20e-12 ***
## dmar      -4.801e+01 3.942e+00 -12.180 < 2e-16 ***
## adequacy   1.024e+01 4.006e+00  2.555 0.01062 *
## dtotord     9.005e+00 1.182e+00  7.616 2.65e-14 ***
## monpre      1.555e+00 1.373e+00  1.133 0.25741
## nprevist    8.882e+00 5.195e-01 17.099 < 2e-16 ***
## disllb     -1.951e-01 4.711e-03 -41.409 < 2e-16 ***
## birmon     -3.931e-01 3.927e-01 -1.001 0.31682
## dgestat    1.066e+02 5.861e-01 181.939 < 2e-16 ***
## csex       -1.329e+02 2.668e+00 -49.830 < 2e-16 ***
## dplural    -6.463e+02 7.964e+00 -81.155 < 2e-16 ***
## anemia     -2.605e+01 1.344e+01 -1.938 0.05261 .
## diabetes   -1.586e+02 8.324e+00 -19.051 < 2e-16 ***
## herpes      -1.487e+01 1.707e+01 -0.871 0.38356
## chyper      1.282e+02 1.531e+01  8.371 < 2e-16 ***
## pre4000    -3.767e+02 1.125e+01 -33.491 < 2e-16 ***
## preterm     2.527e+02 1.141e+01 22.153 < 2e-16 ***
## tobacco    -1.579e+02 6.579e+00 -24.006 < 2e-16 ***
## cigar       -3.788e+00 4.518e-01 -8.385 < 2e-16 ***
## alcohol     1.933e+00 1.916e+01  0.101 0.91965
## drink      -1.918e+01 5.944e+00 -3.228 0.00125 **
## wgain       8.256e+00 1.157e-01 71.347 < 2e-16 ***
## drinksq     3.241e-01 7.389e-02  4.387 1.15e-05 ***
## cig_drink   8.309e-02 2.229e-01  0.373 0.70927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 450.8 on 114581 degrees of freedom
## Multiple R-squared:  0.4066, Adjusted R-squared:  0.4064
## F-statistic:  2804 on 28 and 114581 DF, p-value: < 2.2e-16

```

1b - Series Estimator, Higher order Polynomial specifications

Performance was assessed using up to order 3, 4, 5 polynomial orders, and all were similar, with small R² gains from 3 to 4, but not from 4 to 5.

```
poly.mod <- lm(dbrwt ~ poly(stresfip, 4, raw=T)+poly(dimage, 4, raw=T)+poly(ormooth, 4, raw=T)+poly(mrace3, 4, raw=T))
summary(poly.mod)
```

```

##
## Call:
## lm(formula = dbrwt ~ poly(stresfip, 4, raw = T) + poly(dimage,
##     4, raw = T) + poly(ormooth, 4, raw = T) + poly(mrace3, 4,
##     raw = T) + poly(dmeduc, 4, raw = T) + poly(dmar, 4, raw = T) +
##     poly(adequacy, 4, raw = T) + poly(dtotord, 4, raw = T) +
##     poly(monpre, 4, raw = T) + poly(nprevist, 4, raw = T) + poly(disllb,
```

```

##      4, raw = T) + poly(birmon, 4, raw = T) + poly(dgestat, 4,
## raw = T) + poly(csex, 4, raw = T) + poly(dplural, 4, raw = T) +
## poly(anemia, 4, raw = T) + poly(diabetes, 4, raw = T) + poly(herpes,
## 4, raw = T) + poly(chyper, 4, raw = T) + poly(pre4000, 4,
## raw = T) + poly(preterm, 4, raw = T) + poly(tobacco, 4, raw = T) +
## poly(cigar, 4, raw = T) + poly(alcohol, 4, raw = T) + poly(drink,
## 4, raw = T) + poly(wgain, 4, raw = T), data = dat_drop)
##
## Residuals:
##      Min       1Q   Median      3Q      Max
## -2888.50  -286.96   -14.38  270.36  2983.43
##
## Coefficients: (35 not defined because of singularities)
##                                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 -1.796e+02  1.815e+03 -0.099  0.921202
## poly(stresfip, 4, raw = T)1  2.129e+01  1.734e+01  1.227  0.219671
## poly(stresfip, 4, raw = T)2 -1.537e+00  9.957e-01 -1.544  0.122612
## poly(stresfip, 4, raw = T)3  4.165e-02  2.296e-02  1.814  0.069645 .
## poly(stresfip, 4, raw = T)4 -3.757e-04  1.864e-04 -2.016  0.043822 *
## poly(dimage, 4, raw = T)1   -2.595e+01  4.604e+01 -0.564  0.572910
## poly(dimage, 4, raw = T)2   1.708e+00  2.520e+00  0.677  0.498107
## poly(dimage, 4, raw = T)3  -4.113e-02  5.987e-02 -0.687  0.492067
## poly(dimage, 4, raw = T)4   3.327e-04  5.213e-04  0.638  0.523378
## poly(ormoth, 4, raw = T)1   6.212e+01  8.031e+01  0.773  0.439252
## poly(ormoth, 4, raw = T)2  -1.428e+02  7.898e+01 -1.809  0.070527 .
## poly(ormoth, 4, raw = T)3   5.373e+01  2.403e+01  2.236  0.025330 *
## poly(ormoth, 4, raw = T)4  -5.648e+00  2.310e+00 -2.445  0.014494 *
## poly(mrace3, 4, raw = T)1  -5.064e+02  3.793e+01 -13.352 < 2e-16 ***
## poly(mrace3, 4, raw = T)2   1.110e+02  9.580e+00  11.586 < 2e-16 ***
## poly(mrace3, 4, raw = T)3     NA        NA        NA        NA
## poly(mrace3, 4, raw = T)4     NA        NA        NA        NA
## poly(dmeduc, 4, raw = T)1   7.656e+01  2.794e+01  2.740  0.006147 **
## poly(dmeduc, 4, raw = T)2  -1.847e+01  4.673e+00 -3.953  7.74e-05 ***
## poly(dmeduc, 4, raw = T)3   1.462e+00  3.227e-01  4.530  5.91e-06 ***
## poly(dmeduc, 4, raw = T)4  -3.740e-02  7.817e-03 -4.785  1.71e-06 ***
## poly(dmar, 4, raw = T)1   -3.209e+01  4.029e+00 -7.963  1.69e-15 ***
## poly(dmar, 4, raw = T)2     NA        NA        NA        NA
## poly(dmar, 4, raw = T)3     NA        NA        NA        NA
## poly(dmar, 4, raw = T)4     NA        NA        NA        NA
## poly(adequacy, 4, raw = T)1  5.798e+01  2.027e+01  2.861  0.004230 **
## poly(adequacy, 4, raw = T)2 -1.010e+01  6.149e+00 -1.642  0.100491
## poly(adequacy, 4, raw = T)3     NA        NA        NA
## poly(adequacy, 4, raw = T)4     NA        NA        NA
## poly(dtotord, 4, raw = T)1   4.964e+00  8.883e+00  0.559  0.576250
## poly(dtotord, 4, raw = T)2   1.099e+00  2.367e+00  0.464  0.642361
## poly(dtotord, 4, raw = T)3  -7.452e-02  2.312e-01 -0.322  0.747159
## poly(dtotord, 4, raw = T)4   4.490e-04  6.944e-03  0.065  0.948444
## poly(monpre, 4, raw = T)1   4.745e+01  1.897e+01  2.501  0.012400 *
## poly(monpre, 4, raw = T)2  -2.215e+01  8.614e+00 -2.571  0.010133 *
## poly(monpre, 4, raw = T)3   3.983e+00  1.500e+00  2.655  0.007935 **
## poly(monpre, 4, raw = T)4  -2.247e-01  8.730e-02 -2.574  0.010043 *
## poly(nprevist, 4, raw = T)1 -2.664e+00  5.419e+00 -0.492  0.623017
## poly(nprevist, 4, raw = T)2   2.143e+00  5.025e-01  4.265  2.00e-05 ***
## poly(nprevist, 4, raw = T)3 -1.116e-01  1.893e-02 -5.895  3.76e-09 ***

```

```

## poly(npervist, 4, raw = T)4  1.451e-03  2.309e-04  6.284 3.31e-10 ***
## poly(disllb, 4, raw = T)1   8.129e-01  3.273e-01  2.483 0.013013 *
## poly(disllb, 4, raw = T)2  -1.597e-02  4.298e-03 -3.715 0.000203 ***
## poly(disllb, 4, raw = T)3   5.806e-05  1.780e-05  3.261 0.001110 **
## poly(disllb, 4, raw = T)4  -5.036e-08  1.660e-08 -3.033 0.002419 **
## poly(birmon, 4, raw = T)1   2.291e+00  1.139e+01  0.201 0.840557
## poly(birmon, 4, raw = T)2  -9.345e-01  3.374e+00 -0.277 0.781808
## poly(birmon, 4, raw = T)3   1.136e-01  3.834e-01  0.296 0.766954
## poly(birmon, 4, raw = T)4  -4.606e-03  1.468e-02 -0.314 0.753711
## poly(dgestat, 4, raw = T)1   5.239e+02  2.128e+02  2.462 0.013825 *
## poly(dgestat, 4, raw = T)2  -3.310e+01  9.424e+00 -3.513 0.000444 ***
## poly(dgestat, 4, raw = T)3   1.023e+00  1.825e-01  5.607 2.06e-08 ***
## poly(dgestat, 4, raw = T)4  -1.056e-02  1.307e-03 -8.081 6.48e-16 ***
## poly(csex, 4, raw = T)1    -1.334e+02  2.589e+00 -51.517 < 2e-16 ***
## poly(csex, 4, raw = T)2      NA        NA        NA        NA
## poly(csex, 4, raw = T)3      NA        NA        NA        NA
## poly(csex, 4, raw = T)4      NA        NA        NA        NA
## poly(dplural, 4, raw = T)1  -2.626e+03  3.561e+02 -7.373 1.67e-13 ***
## poly(dplural, 4, raw = T)2   9.684e+02  1.839e+02  5.266 1.40e-07 ***
## poly(dplural, 4, raw = T)3  -1.267e+02  2.833e+01 -4.472 7.75e-06 ***
## poly(dplural, 4, raw = T)4      NA        NA        NA        NA
## poly(anemia, 4, raw = T)1   -2.470e+01  1.305e+01 -1.892 0.058433 .
## poly(anemia, 4, raw = T)2      NA        NA        NA        NA
## poly(anemia, 4, raw = T)3      NA        NA        NA        NA
## poly(anemia, 4, raw = T)4      NA        NA        NA        NA
## poly(diabetes, 4, raw = T)1  -1.588e+02  8.101e+00 -19.599 < 2e-16 ***
## poly(diabetes, 4, raw = T)2      NA        NA        NA        NA
## poly(diabetes, 4, raw = T)3      NA        NA        NA        NA
## poly(diabetes, 4, raw = T)4      NA        NA        NA        NA
## poly(herpes, 4, raw = T)1   -9.742e+00  1.656e+01 -0.588 0.556420
## poly(herpes, 4, raw = T)2      NA        NA        NA        NA
## poly(herpes, 4, raw = T)3      NA        NA        NA        NA
## poly(herpes, 4, raw = T)4      NA        NA        NA        NA
## poly(chyper, 4, raw = T)1   1.048e+02  1.487e+01  7.048 1.83e-12 ***
## poly(chyper, 4, raw = T)2      NA        NA        NA        NA
## poly(chyper, 4, raw = T)3      NA        NA        NA        NA
## poly(chyper, 4, raw = T)4      NA        NA        NA        NA
## poly(pre4000, 4, raw = T)1  -3.740e+02  1.092e+01 -34.250 < 2e-16 ***
## poly(pre4000, 4, raw = T)2      NA        NA        NA        NA
## poly(pre4000, 4, raw = T)3      NA        NA        NA        NA
## poly(pre4000, 4, raw = T)4      NA        NA        NA        NA
## poly(preterm, 4, raw = T)1   2.253e+02  1.108e+01  20.336 < 2e-16 ***
## poly(preterm, 4, raw = T)2      NA        NA        NA        NA
## poly(preterm, 4, raw = T)3      NA        NA        NA        NA
## poly(preterm, 4, raw = T)4      NA        NA        NA        NA
## poly(tobacco, 4, raw = T)1  -9.405e+01  1.372e+01 -6.855 7.14e-12 ***
## poly(tobacco, 4, raw = T)2      NA        NA        NA        NA
## poly(tobacco, 4, raw = T)3      NA        NA        NA        NA
## poly(tobacco, 4, raw = T)4      NA        NA        NA        NA
## poly(cigar, 4, raw = T)1   -1.670e+01  3.179e+00 -5.253 1.50e-07 ***
## poly(cigar, 4, raw = T)2   6.851e-01  2.177e-01  3.147 0.001650 **
## poly(cigar, 4, raw = T)3  -1.125e-02  5.188e-03 -2.169 0.030104 *
## poly(cigar, 4, raw = T)4   6.047e-05  3.619e-05  1.671 0.094773 .
## poly(alcohol, 4, raw = T)1  -4.229e+01  2.552e+01 -1.657 0.097503 .

```

```

## poly(alcohol, 4, raw = T)2      NA      NA      NA
## poly(alcohol, 4, raw = T)3      NA      NA      NA
## poly(alcohol, 4, raw = T)4      NA      NA      NA
## poly(drink, 4, raw = T)1     -4.602e+01 1.264e+01 -3.642 0.000271 ***
## poly(drink, 4, raw = T)2      3.098e+00 1.155e+00 2.682 0.007314 **
## poly(drink, 4, raw = T)3     -6.451e-02 2.950e-02 -2.187 0.028754 *
## poly(drink, 4, raw = T)4      4.118e-04 2.063e-04 1.996 0.045913 *
## poly(wgain, 4, raw = T)1     -9.557e+00 1.382e+00 -6.918 4.61e-12 ***
## poly(wgain, 4, raw = T)2      6.532e-01 6.375e-02 10.246 < 2e-16 ***
## poly(wgain, 4, raw = T)3     -8.873e-03 1.154e-03 -7.687 1.52e-14 ***
## poly(wgain, 4, raw = T)4      3.703e-05 6.958e-06 5.323 1.03e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 437.3 on 114540 degrees of freedom
## Multiple R-squared:  0.4418, Adjusted R-squared:  0.4415
## F-statistic:  1314 on 69 and 114540 DF,  p-value: < 2.2e-16

```

1b - Series Estimator, Squared Terms and Interaction Terms

```

all.mod <- lm(dbrwt ~ ., data = dat.all)
summary(all.mod)

##
## Call:
## lm(formula = dbrwt ~ ., data = dat.all)
##
## Residuals:
##    Min      1Q  Median      3Q     Max
## -3725.3 -275.6     3.6   283.2  2163.8
##
## Coefficients: (12 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -7.404e+03 3.929e+03 -1.885 0.059476 .
## stresfip     6.740e+01 4.870e+01  1.384 0.166390
## dimage      -1.622e+01 1.867e+01 -0.869 0.384955
## ormoth      2.719e+02 1.767e+02  1.538 0.123938
## mrace3      -2.794e+02 1.346e+02 -2.076 0.037865 *
## dmeduc      -6.434e+01 4.452e+01 -1.445 0.148349
## dmar        4.460e+02 2.376e+02  1.877 0.060484 .
## adequacy    1.558e+02 2.321e+02  0.672 0.501848
## dtotord     -1.439e+02 6.301e+01 -2.283 0.022424 *
## monpre       2.115e+02 7.620e+01  2.775 0.005521 **
## nprevist     2.186e+01 2.920e+01  0.749 0.454031
## disllb      -3.207e+00 1.157e+00 -2.773 0.005560 **
## birmon      -1.543e+01 2.441e+01 -0.632 0.527359
## dgestat      6.254e+02 3.262e+01 19.175 < 2e-16 ***
## csex         2.314e+02 1.625e+02  1.424 0.154592
## dplural      1.297e+03 5.314e+02  2.440 0.014688 *
## anemia      -1.490e+02 6.493e+02 -0.229 0.818543
## diabetes     -4.905e+02 4.414e+02 -1.111 0.266509
## herpes       1.118e+03 9.327e+02  1.199 0.230530

```

## chyper	-1.552e+03	1.251e+03	-1.241	0.214502
## pre4000	-7.087e+02	6.880e+02	-1.030	0.302993
## preterm	3.684e+02	5.863e+02	0.628	0.529713
## tobacco	-8.477e+02	3.857e+02	-2.198	0.027975 *
## cigar	-5.079e+01	2.600e+01	-1.953	0.050773 .
## alcohol	-2.333e+03	1.549e+03	-1.506	0.132044
## drink	-4.394e+02	8.422e+02	-0.522	0.601826
## wgain	5.033e+00	6.816e+00	0.738	0.460254
## stresfip_2	-3.064e-02	5.012e-02	-0.611	0.541069
## dimage_2	-3.926e-01	5.261e-02	-7.462	8.55e-14 ***
## ormoth_2	1.168e+01	1.875e+00	6.230	4.70e-10 ***
## mrace3_2	1.153e+02	1.002e+01	11.509	< 2e-16 ***
## dmeduc_2	6.047e-01	2.225e-01	2.718	0.006571 **
## dmar_2	NA	NA	NA	NA
## adequacy_2	-1.247e+01	1.086e+01	-1.148	0.250944
## dtotord_2	-1.992e+00	4.399e-01	-4.528	5.96e-06 ***
## monpre_2	2.973e+00	8.491e-01	3.502	0.000462 ***
## nprevist_2	-6.718e-01	6.459e-02	-10.400	< 2e-16 ***
## disllb_2	4.734e-04	8.918e-05	5.308	1.11e-07 ***
## birmon_2	-7.964e-02	1.237e-01	-0.644	0.519625
## dgestat_2	-5.512e+00	1.086e-01	-50.737	< 2e-16 ***
## csex_2	NA	NA	NA	NA
## dplural_2	9.307e+01	1.808e+01	5.147	2.65e-07 ***
## anemia_2	NA	NA	NA	NA
## diabetes_2	NA	NA	NA	NA
## herpes_2	NA	NA	NA	NA
## chyper_2	NA	NA	NA	NA
## pre4000_2	NA	NA	NA	NA
## preterm_2	NA	NA	NA	NA
## tobacco_2	NA	NA	NA	NA
## cigar_2	9.951e-02	3.007e-02	3.309	0.000937 ***
## alcohol_2	NA	NA	NA	NA
## drink_2	2.207e-01	1.823e-01	1.210	0.226185
## wgain_2	2.109e-02	5.517e-03	3.823	0.000132 ***
## stresfip_dimage	8.747e-02	1.404e-01	0.623	0.533173
## stresfip_ormoth	-9.259e-02	1.465e+00	-0.063	0.949603
## stresfip_mrace3	-1.078e-01	1.021e+00	-0.106	0.915878
## stresfip_dmeduc	2.155e-02	3.062e-01	0.070	0.943900
## stresfip_dmar	-1.489e+00	2.165e+00	-0.688	0.491456
## stresfip_adequacy	-3.243e+00	1.740e+00	-1.864	0.062278 .
## stresfip_dtotord	2.737e-01	5.585e-01	0.490	0.624125
## stresfip_monpre	4.119e-01	5.574e-01	0.739	0.459975
## stresfip_nprevist	-1.127e-01	2.388e-01	-0.472	0.636851
## stresfip_disllb	-1.901e-03	2.186e-03	-0.870	0.384462
## stresfip_birmon	1.040e-01	1.837e-01	0.566	0.571222
## stresfip_dgestat	-6.584e-01	2.651e-01	-2.483	0.013012 *
## stresfip_csex	7.279e-02	1.220e+00	0.060	0.952432
## stresfip_dplural	-4.799e+00	4.664e+00	-1.029	0.303478
## stresfip_anemia	-1.014e+00	6.007e+00	-0.169	0.865940
## stresfip_diabetes	3.209e+00	3.603e+00	0.890	0.373222
## stresfip_herpes	3.395e+00	4.947e+00	0.686	0.492504
## stresfip_chyper	-1.202e+01	6.291e+00	-1.910	0.056126 .
## stresfip_pre4000	-5.158e+00	4.765e+00	-1.082	0.279058
## stresfip_preterm	1.692e+00	5.187e+00	0.326	0.744276

## stresfip_tobacco	8.874e-01	3.836e+00	0.231	0.817078
## stresfip_cigar	2.633e-01	2.638e-01	0.998	0.318207
## stresfip_alcohol	-5.545e+00	1.977e+01	-0.280	0.779113
## stresfip_drink	-1.684e+01	1.387e+01	-1.214	0.224649
## stresfip_wgain	-8.034e-02	5.765e-02	-1.394	0.163435
## dimage_ormoth	3.996e-01	5.969e-01	0.669	0.503195
## dimage_mrace3	-2.218e-01	5.043e-01	-0.440	0.660035
## dimage_dmeduc	-1.607e-01	1.600e-01	-1.005	0.315097
## dimage_dmar	-2.386e+00	8.475e-01	-2.815	0.004876 **
## dimage_adequacy	-1.151e+00	9.094e-01	-1.265	0.205702
## dimage_dtotord	1.506e+00	2.674e-01	5.630	1.80e-08 ***
## dimage_monpre	-1.153e-01	3.149e-01	-0.366	0.714240
## dimage_nprevist	-1.683e-01	1.144e-01	-1.472	0.140992
## dimage_disllb	-2.942e-03	1.097e-03	-2.683	0.007306 **
## dimage_birmon	-2.782e-02	8.949e-02	-0.311	0.755934
## dimage_dgestat	7.873e-01	1.262e-01	6.238	4.45e-10 ***
## dimage_csex	-8.832e-01	6.094e-01	-1.449	0.147277
## dimage_dplural	2.011e+00	1.816e+00	1.107	0.268143
## dimage_anemia	9.208e+00	3.104e+00	2.966	0.003016 **
## dimage_diabetes	-1.556e-01	1.744e+00	-0.089	0.928902
## dimage_herpes	-6.091e+00	4.015e+00	-1.517	0.129306
## dimage_chyper	-1.539e-01	2.930e+00	-0.053	0.958109
## dimage_pre4000	3.983e+00	2.941e+00	1.354	0.175642
## dimage_preterm	-2.090e+00	2.710e+00	-0.771	0.440559
## dimage_tobacco	4.509e+00	1.446e+00	3.117	0.001827 **
## dimage_cigar	7.281e-02	9.574e-02	0.760	0.446979
## dimage_alcohol	-2.127e+00	3.904e+00	-0.545	0.585841
## dimage_drink	-1.245e+00	8.471e-01	-1.469	0.141790
## dimage_wgain	-5.492e-02	2.570e-02	-2.137	0.032599 *
## ormoth_mrace3	1.110e+01	4.972e+00	2.233	0.025581 *
## ormoth_dmeduc	-1.188e+00	1.266e+00	-0.938	0.348264
## ormoth_dmar	-2.404e+00	6.252e+00	-0.385	0.700551
## ormoth_adequacy	-1.425e+01	6.877e+00	-2.072	0.038253 *
## ormoth_dtotord	-1.252e+00	2.280e+00	-0.549	0.582855
## ormoth_monpre	3.041e+00	2.276e+00	1.336	0.181520
## ormoth_nprevist	-1.822e+00	8.220e-01	-2.217	0.026642 *
## ormoth_disllb	1.947e-03	9.261e-03	0.210	0.833523
## ormoth_birmon	-2.891e-01	7.494e-01	-0.386	0.699675
## ormoth_dgestat	-9.165e-01	9.852e-01	-0.930	0.352191
## ormoth_csex	7.468e+00	5.088e+00	1.468	0.142158
## ormoth_dplural	8.643e+00	1.720e+01	0.503	0.615231
## ormoth_anemia	8.068e+00	1.905e+01	0.423	0.671968
## ormoth_diabetes	-3.980e+01	1.620e+01	-2.456	0.014042 *
## ormoth_herpes	-1.984e+01	3.385e+01	-0.586	0.557851
## ormoth_chyper	-7.964e+01	4.031e+01	-1.976	0.048170 *
## ormoth_pre4000	-2.535e+00	2.358e+01	-0.107	0.914405
## ormoth_preterm	2.078e+01	1.896e+01	1.096	0.273131
## ormoth_tobacco	-1.034e+01	1.285e+01	-0.804	0.421212
## ormoth_cigar	-2.434e-01	9.852e-01	-0.247	0.804849
## ormoth_alcohol	-2.377e+01	4.055e+01	-0.586	0.557764
## ormoth_drink	-2.465e+00	1.062e+01	-0.232	0.816526
## ormoth_wgain	4.084e-02	2.038e-01	0.200	0.841128
## mrace3_dmeduc	-5.547e-01	1.288e+00	-0.431	0.666626
## mrace3_dmar	-5.761e+00	5.644e+00	-1.021	0.307406

```

## mrace3_adequacy    1.277e+00  6.075e+00  0.210  0.833502
## mrace3_dtotord   -1.820e+00  1.796e+00 -1.013  0.310922
## mrace3_monpre   -1.018e+00  1.953e+00 -0.521  0.602301
## mrace3_nprevist  1.592e+00  8.302e-01  1.917  0.055177 .
## mrace3_disllb    1.213e-02  7.989e-03  1.518  0.128939
## mrace3_birmon   -5.347e-01  6.486e-01 -0.824  0.409728
## mrace3_dgestat  -7.113e+00  8.403e-01 -8.465 < 2e-16 ***
## mrace3_csex      3.862e+00  4.440e+00  0.870  0.384405
## mrace3_dplural   9.287e+00  1.411e+01  0.658  0.510361
## mrace3_anemia    1.738e+01  1.735e+01  1.002  0.316480
## mrace3_diabetes  -3.353e+01  1.356e+01 -2.473  0.013401 *
## mrace3_herpes   -1.639e+01  2.524e+01 -0.649  0.516157
## mrace3_chyper   -1.606e+00  2.254e+01 -0.071  0.943213
## mrace3_pre4000  -1.771e+01  2.515e+01 -0.704  0.481398
## mrace3_preterm   3.871e+01  1.784e+01  2.170  0.030004 *
## mrace3_tobacco  -1.343e+01  1.001e+01 -1.342  0.179559
## mrace3_cigar     9.336e-01  8.187e-01  1.140  0.254131
## mrace3_alcohol   4.509e+01  2.603e+01  1.732  0.083271 .
## mrace3_drink     4.935e+00  5.066e+00  0.974  0.329953
## mrace3_wgain    -2.742e-01  1.770e-01 -1.550  0.121251
## dmeduc_dmar     6.023e+00  2.276e+00  2.647  0.008132 **
## dmeduc_adequacy 3.879e+00  2.085e+00  1.861  0.062794 .
## dmeduc_dtotord  -5.638e-01  5.626e-01 -1.002  0.316336
## dmeduc_monpre   -1.611e+00  7.128e-01 -2.260  0.023838 *
## dmeduc_nprevist 9.364e-02  2.756e-01  0.340  0.733986
## dmeduc_disllb   5.356e-04  2.429e-03  0.221  0.825468
## dmeduc_birmon   1.473e-01  2.007e-01  0.734  0.463180
## dmeduc_dgestat  1.008e+00  3.084e-01  3.269  0.001080 **
## dmeduc_csex      9.427e-01  1.368e+00  0.689  0.490697
## dmeduc_dplural   4.055e+00  4.024e+00  1.008  0.313561
## dmeduc_anemia   -1.189e+01  7.130e+00 -1.668  0.095332 .
## dmeduc_diabetes  8.008e+00  4.228e+00  1.894  0.058235 .
## dmeduc_herpes   7.251e+00  9.592e+00  0.756  0.449675
## dmeduc_chyper   3.566e+00  8.109e+00  0.440  0.660124
## dmeduc_pre4000 -1.990e-01  5.477e+00 -0.036  0.971024
## dmeduc_preterm  -2.909e+00  5.906e+00 -0.493  0.622303
## dmeduc_tobacco  -7.487e+00  4.281e+00 -1.749  0.080295 .
## dmeduc_cigar     3.389e-02  2.946e-01  0.115  0.908404
## dmeduc_alcohol   3.980e-01  1.007e+01  0.040  0.968464
## dmeduc_drink     1.401e+00  2.220e+00  0.631  0.528008
## dmeduc_wgain    2.699e-01  6.265e-02  4.309  1.64e-05 ***
## dmar_adequacy   -4.232e+00  1.096e+01 -0.386  0.699421
## dmar_dtotord    2.292e+00  3.362e+00  0.682  0.495369
## dmar_monpre     2.774e+00  3.701e+00  0.750  0.453479
## dmar_nprevist   1.154e+00  1.409e+00  0.819  0.412698
## dmar_disllb     1.977e-02  1.399e-02  1.414  0.157434
## dmar_birmon    -3.556e+00  1.123e+00 -3.167  0.001540 **
## dmar_dgestat   -1.192e+01  1.560e+00 -7.641  2.17e-14 ***
## dmar_csex       -4.860e+00  7.669e+00 -0.634  0.526303
## dmar_dplural   -3.154e+01  2.531e+01 -1.246  0.212816
## dmar_anemia    -1.819e+00  3.366e+01 -0.054  0.956893
## dmar_diabetes  -3.848e+00  2.422e+01 -0.159  0.873801
## dmar_herpes   -1.495e+01  4.623e+01 -0.323  0.746418
## dmar_chyper   5.565e+01  4.547e+01  1.224  0.221029

```

## dmar_pre4000	-7.243e+00	3.950e+01	-0.183	0.854515
## dmar_preterm	-1.961e+01	3.101e+01	-0.632	0.527193
## dmar_tobacco	-1.070e+00	1.519e+01	-0.070	0.943835
## dmar_cigar	2.023e+00	9.850e-01	2.054	0.039993 *
## dmar_alcohol	2.654e+01	5.090e+01	0.521	0.602094
## dmar_drink	-5.572e+00	1.056e+01	-0.527	0.597879
## dmar_wgain	1.310e-01	3.032e-01	0.432	0.665657
## adequacy_dtotord	4.248e-01	3.202e+00	0.133	0.894475
## adequacy_monpre	-1.300e+00	4.807e+00	-0.270	0.786813
## adequacy_nprevist	-5.676e+00	1.922e+00	-2.953	0.003144 **
## adequacy_disllb	-1.137e-03	1.370e-02	-0.083	0.933868
## adequacy_birmon	1.872e+00	1.150e+00	1.628	0.103630
## adequacy_dgestat	-6.750e-02	1.423e+00	-0.047	0.962162
## adequacy_csex	-9.206e+00	7.811e+00	-1.179	0.238537
## adequacy_dplural	-2.711e+01	2.381e+01	-1.139	0.254901
## adequacy_anemia	-6.830e+01	3.446e+01	-1.982	0.047464 *
## adequacy_diabetes	5.478e+01	2.483e+01	2.206	0.027361 *
## adequacy_herpes	-9.479e+00	5.484e+01	-0.173	0.862775
## adequacy_chyper	3.214e+01	4.485e+01	0.717	0.473621
## adequacy_pre4000	2.018e+01	3.112e+01	0.649	0.516657
## adequacy_preterm	-1.738e-01	2.988e+01	-0.006	0.995359
## adequacy_tobacco	4.176e+01	1.777e+01	2.350	0.018796 *
## adequacy_cigar	2.320e+00	1.195e+00	1.941	0.052254 .
## adequacy_alcohol	-2.987e+01	4.478e+01	-0.667	0.504742
## adequacy_drink	3.541e+00	9.763e+00	0.363	0.716811
## adequacy_wgain	9.685e-01	3.228e-01	3.000	0.002700 **
## dtotord_monpre	-1.870e-01	1.037e+00	-0.180	0.856961
## dtotord_nprevist	-8.938e-01	4.078e-01	-2.192	0.028403 *
## dtotord_disllb	-6.235e-03	4.975e-03	-1.253	0.210127
## dtotord_birmon	-1.230e-01	3.392e-01	-0.363	0.716949
## dtotord_dgestat	7.382e-01	4.678e-01	1.578	0.114573
## dtotord_csex	-3.535e+00	2.304e+00	-1.535	0.124882
## dtotord_dplural	2.565e+01	6.620e+00	3.874	0.000107 ***
## dtotord_anemia	-6.459e-01	9.868e+00	-0.065	0.947809
## dtotord_diabetes	7.619e+00	6.093e+00	1.251	0.211110
## dtotord_herpes	1.558e+01	1.417e+01	1.099	0.271681
## dtotord_chyper	2.330e+01	1.121e+01	2.078	0.037671 *
## dtotord_pre4000	1.583e+00	7.441e+00	0.213	0.831521
## dtotord_preterm	-2.001e+01	7.168e+00	-2.792	0.005241 **
## dtotord_tobacco	4.845e+00	5.126e+00	0.945	0.344546
## dtotord_cigar	7.493e-02	3.089e-01	0.243	0.808355
## dtotord_alcohol	8.925e+00	1.311e+01	0.681	0.495916
## dtotord_drink	4.565e+00	2.765e+00	1.651	0.098739 .
## dtotord_wgain	-8.531e-02	9.653e-02	-0.884	0.376860
## monpre_nprevist	1.232e+00	6.544e-01	1.882	0.059814 .
## monpre_disllb	-1.972e-03	4.732e-03	-0.417	0.676921
## monpre_birmon	-9.021e-01	3.958e-01	-2.279	0.022655 *
## monpre_dgestat	-6.413e+00	4.933e-01	-13.001	< 2e-16 ***
## monpre_csex	6.896e-01	2.677e+00	0.258	0.796745
## monpre_dplural	-6.678e+00	8.746e+00	-0.764	0.445110
## monpre_anemia	2.264e+01	1.117e+01	2.026	0.042743 *
## monpre_diabetes	6.680e+00	9.188e+00	0.727	0.467169
## monpre_herpes	-6.313e+00	1.810e+01	-0.349	0.727287
## monpre_chyper	-1.384e+01	1.602e+01	-0.864	0.387582

## monpre_pre4000	-1.319e+01	1.088e+01	-1.212	0.225629
## monpre_preterm	2.351e+01	1.023e+01	2.298	0.021579 *
## monpre_tobacco	-5.294e+00	5.651e+00	-0.937	0.348855
## monpre_cigar	-4.691e-01	3.737e-01	-1.255	0.209455
## monpre_alcohol	8.278e+00	1.277e+01	0.648	0.516841
## monpre_drink	-8.283e-01	2.456e+00	-0.337	0.735947
## monpre_wgain	-4.176e-01	1.099e-01	-3.798	0.000146 ***
## nprevist_disllb	8.915e-04	1.733e-03	0.514	0.607001
## nprevist_birmon	-1.149e-01	1.476e-01	-0.779	0.436096
## nprevist_dgestat	-1.450e-01	1.894e-01	-0.766	0.443832
## nprevist_csex	-3.162e-01	1.013e+00	-0.312	0.754965
## nprevist_dplural	-3.687e-01	2.014e+00	-0.183	0.854779
## nprevist_anemia	7.010e+00	4.506e+00	1.556	0.119764
## nprevist_diabetes	-4.436e+00	2.158e+00	-2.056	0.039802 *
## nprevist_herpes	1.031e+01	6.863e+00	1.502	0.133180
## nprevist_chyper	-1.627e+00	3.659e+00	-0.445	0.656634
## nprevist_pre4000	-1.904e+00	4.014e+00	-0.474	0.635249
## nprevist_preterm	1.769e+00	3.445e+00	0.513	0.607649
## nprevist_tobacco	5.162e+00	2.487e+00	2.075	0.037961 *
## nprevist_cigar	3.219e-01	1.681e-01	1.915	0.055502 .
## nprevist_alcohol	-5.125e+00	7.078e+00	-0.724	0.469030
## nprevist_drink	6.112e-01	1.574e+00	0.388	0.697758
## nprevist_wgain	-3.627e-03	3.871e-02	-0.094	0.925358
## disllb_birmon	-1.049e-03	1.351e-03	-0.777	0.437447
## disllb_dgestat	1.523e-02	1.991e-03	7.648	2.05e-14 ***
## disllb_csex	1.093e-02	9.165e-03	1.193	0.233003
## disllb_dplural	1.826e-01	2.772e-02	6.587	4.52e-11 ***
## disllb_anemia	-4.157e-02	4.611e-02	-0.902	0.367298
## disllb_diabetes	5.148e-02	2.841e-02	1.812	0.069994 .
## disllb_herpes	-2.208e-03	5.653e-02	-0.039	0.968838
## disllb_chyper	-7.294e-02	5.333e-02	-1.368	0.171435
## disllb_pre4000	7.805e-01	4.095e-01	1.906	0.056672 .
## disllb_preterm	2.282e-01	3.901e-01	0.585	0.558574
## disllb_tobacco	-2.647e-03	2.303e-02	-0.115	0.908470
## disllb_cigar	-1.185e-03	1.618e-03	-0.732	0.463897
## disllb_alcohol	4.470e-02	6.492e-02	0.689	0.491107
## disllb_drink	2.090e-02	1.555e-02	1.344	0.179033
## disllb_wgain	-2.965e-05	4.010e-04	-0.074	0.941053
## birmon_dgestat	6.554e-03	1.675e-01	0.039	0.968790
## birmon_csex	-5.434e-02	7.635e-01	-0.071	0.943260
## birmon_dplural	-3.349e+00	2.318e+00	-1.445	0.148563
## birmon_anemia	-2.731e+00	4.042e+00	-0.676	0.499223
## birmon_diabetes	-3.384e+00	2.409e+00	-1.405	0.160054
## birmon_herpes	4.432e+00	4.981e+00	0.890	0.373558
## birmon_chyper	-2.865e-01	4.468e+00	-0.064	0.948871
## birmon_pre4000	3.645e+00	3.372e+00	1.081	0.279635
## birmon_preterm	-2.002e+00	3.328e+00	-0.602	0.547484
## birmon_tobacco	-2.731e-01	1.872e+00	-0.146	0.884024
## birmon_cigar	-1.761e-02	1.289e-01	-0.137	0.891334
## birmon_alcohol	1.138e+01	5.469e+00	2.080	0.037501 *
## birmon_drink	1.878e+00	1.302e+00	1.442	0.149276
## birmon_wgain	-2.980e-02	3.324e-02	-0.897	0.369932
## dgestat_csex	-7.761e+00	1.144e+00	-6.783	1.18e-11 ***
## dgestat_dplural	-2.909e+01	2.495e+00	-11.658	< 2e-16 ***

## dgestat_anemia	1.275e+00	5.069e+00	0.251	0.801435
## dgestat_diabetes	1.352e+01	3.636e+00	3.720	0.000199 ***
## dgestat_herpes	-8.764e+00	7.821e+00	-1.121	0.262457
## dgestat_chyper	-2.919e+01	5.007e+00	-5.829	5.59e-09 ***
## dgestat_pre4000	-4.454e+00	5.829e+00	-0.764	0.444831
## dgestat_preterm	-4.015e+00	3.782e+00	-1.061	0.288491
## dgestat_tobacco	1.135e+01	2.556e+00	4.441	8.97e-06 ***
## dgestat_cigar	-1.420e-02	1.758e-01	-0.081	0.935592
## dgestat_alcohol	3.985e+00	6.474e+00	0.616	0.538167
## dgestat_drink	2.123e-02	1.249e+00	0.017	0.986443
## dgestat_wgain	-5.210e-02	4.861e-02	-1.072	0.283895
## csex_dplural	-1.360e+01	1.567e+01	-0.868	0.385588
## csex_anemia	-2.552e+01	2.636e+01	-0.968	0.332939
## csex_diabetes	-1.757e+01	1.626e+01	-1.080	0.279957
## csex_herpes	1.738e+01	3.345e+01	0.519	0.603441
## csex_chyper	-1.781e-01	3.016e+01	-0.006	0.995289
## csex_pre4000	4.506e+01	2.209e+01	2.040	0.041354 *
## csex_preterm	-2.111e+00	2.249e+01	-0.094	0.925218
## csex_tobacco	-1.165e+01	1.283e+01	-0.908	0.364012
## csex_cigar	-1.213e+00	8.802e-01	-1.378	0.168165
## csex_alcohol	-6.384e+00	3.590e+01	-0.178	0.858844
## csex_drink	-9.513e+00	8.452e+00	-1.126	0.260380
## csex_wgain	-4.440e-01	2.256e-01	-1.968	0.049056 *
## dplural_anemia	5.956e+01	5.878e+01	1.013	0.310885
## dplural_diabetes	1.705e+02	4.082e+01	4.177	2.96e-05 ***
## dplural_herpes	-2.872e+02	1.490e+02	-1.927	0.053971 .
## dplural_chyper	-2.137e+02	8.082e+01	-2.644	0.008193 **
## dplural_pre4000	-9.197e+00	7.716e+01	-0.119	0.905120
## dplural_preterm	-1.975e+02	6.098e+01	-3.238	0.001204 **
## dplural_tobacco	-3.462e+01	4.279e+01	-0.809	0.418487
## dplural_cigar	-3.010e-01	3.018e+00	-0.100	0.920538
## dplural_alcohol	4.043e+01	1.325e+02	0.305	0.760288
## dplural_drink	1.698e+00	2.440e+01	0.070	0.944532
## dplural_wgain	-3.332e+00	5.918e-01	-5.631	1.80e-08 ***
## anemia_diabetes	-2.207e+00	7.875e+01	-0.028	0.977636
## anemia_herpes	-7.372e+00	1.137e+02	-0.065	0.948304
## anemia_chyper	4.578e+01	1.210e+02	0.378	0.705231
## anemia_pre4000	-1.391e+02	7.100e+01	-1.960	0.050050 .
## anemia_preterm	2.009e+02	7.106e+01	2.827	0.004706 **
## anemia_tobacco	4.351e+01	5.496e+01	0.792	0.428544
## anemia_cigar	3.809e+00	4.033e+00	0.945	0.344895
## anemia_alcohol	-1.758e+02	1.893e+02	-0.929	0.353043
## anemia_drink	-3.453e+01	3.764e+01	-0.917	0.358934
## anemia_wgain	1.394e+00	1.088e+00	1.281	0.200061
## diabetes_herpes	-9.523e+01	8.770e+01	-1.086	0.277561
## diabetes_chyper	4.479e+01	4.657e+01	0.962	0.336099
## diabetes_pre4000	-6.115e+01	4.564e+01	-1.340	0.180361
## diabetes_preterm	4.940e+00	5.811e+01	0.085	0.932264
## diabetes_tobacco	6.609e+01	4.024e+01	1.642	0.100506
## diabetes_cigar	6.371e-02	2.654e+00	0.024	0.980847
## diabetes_alcohol	-2.335e+02	1.227e+02	-1.902	0.057117 .
## diabetes_drink	-2.534e+01	2.748e+01	-0.922	0.356457
## diabetes_wgain	1.674e-02	6.319e-01	0.026	0.978867
## herpes_chyper	-1.589e+01	2.048e+02	-0.078	0.938144

```

## herpes_pre4000 -6.954e+01 1.332e+02 -0.522 0.601586
## herpes_preterm -1.160e+02 1.728e+02 -0.672 0.501779
## herpes_tobacco 1.214e+00 7.709e+01 0.016 0.987433
## herpes_cigar -3.141e+00 5.444e+00 -0.577 0.563899
## herpes_alcohol -1.716e+01 2.254e+02 -0.076 0.939315
## herpes_drink -3.777e+01 8.386e+01 -0.450 0.652424
## herpes_wgain -2.077e+00 1.434e+00 -1.449 0.147274
## chyper_pre4000 2.510e+02 1.188e+02 2.112 0.034686 *
## chyper_preterm 2.690e+01 8.362e+01 0.322 0.747649
## chyper_tobacco 1.279e+02 7.940e+01 1.611 0.107168
## chyper_cigar 2.552e+00 5.243e+00 0.487 0.626434
## chyper_alcohol 1.191e+03 5.268e+02 2.261 0.023762 *
## chyper_drink 4.062e+02 2.700e+02 1.505 0.132386
## chyper_wgain 2.020e+00 1.111e+00 1.818 0.069053 .
## pre4000_preterm -2.371e+01 9.308e+01 -0.255 0.798958
## pre4000_tobacco -7.841e+01 7.275e+01 -1.078 0.281123
## pre4000_cigar -3.301e+00 5.169e+00 -0.639 0.523101
## pre4000_alcohol 3.660e+02 2.209e+02 1.657 0.097523 .
## pre4000_drink 2.740e+02 1.089e+02 2.515 0.011912 *
## pre4000_wgain 1.555e+00 8.985e-01 1.731 0.083504 .
## preterm_tobacco 4.582e+01 4.530e+01 1.011 0.311783
## preterm_cigar 8.391e+00 2.866e+00 2.928 0.003412 **
## preterm_alcohol -3.929e+01 9.449e+01 -0.416 0.677598
## preterm_drink -1.900e+01 1.517e+01 -1.252 0.210473
## preterm_wgain 4.866e-01 9.805e-01 0.496 0.619700
## tobacco_cigar NA NA NA NA
## tobacco_alcohol 2.557e+01 5.388e+01 0.475 0.635125
## tobacco_drink -1.491e+01 1.177e+01 -1.267 0.204985
## tobacco_wgain -7.249e-01 4.973e-01 -1.457 0.144991
## cigar_alcohol 4.744e+00 2.792e+00 1.699 0.089300 .
## cigar_drink -3.668e-02 4.070e-01 -0.090 0.928185
## cigar_wgain -5.638e-02 3.405e-02 -1.656 0.097807 .
## alcohol_drink NA NA NA NA
## alcohol_wgain 2.036e+00 1.537e+00 1.325 0.185203
## drink_wgain 4.974e-01 3.462e-01 1.437 0.150852
## e2_mod1 8.741e-05 3.993e-06 21.888 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 437 on 114243 degrees of freedom
## Multiple R-squared: 0.4441, Adjusted R-squared: 0.4423
## F-statistic: 249.4 on 366 and 114243 DF, p-value: < 2.2e-16

```

Equal quantile knots were assessed with 3, 4, and 5 knots. All produced almost identical results, so 3 was used.

```

# Find knots based on equal quantiles of data
nknuts <- 3
knots <- attr(bs(dat_drop$dbrwt, df = nknuts+3), "knots")

cspline.mod <- lm(dbrwt ~ bs(stresfip,knots = knots)+bs(dimage,knots = knots)+bs(ormoth,knots = knots)+bs(summary(cspline.mod)

##
## Call:

```

```

## lm(formula = dbrwt ~ bs(stresfip, knots = knots) + bs(dimage,
##   knots = knots) + bs(ormoth, knots = knots) + bs(mrace3, knots = knots) +
##   bs(dmeduc, knots = knots) + bs(dmar, knots = knots) + bs(adequacy,
##   knots = knots) + bs(dtotord, knots = knots) + bs(monpre,
##   knots = knots) + bs(npervist, knots = knots) + bs(disllb,
##   knots = knots) + bs(birmon, knots = knots) + bs(dgestat,
##   knots = knots) + bs(csex, knots = knots) + bs(dplural, knots = knots) +
##   bs(anemia, knots = knots) + bs(diabetes, knots = knots) +
##   bs(herpes, knots = knots) + bs(chyper, knots = knots) + bs(pre4000,
##   knots = knots) + bs(preterm, knots = knots) + bs(tobacco,
##   knots = knots) + bs(cigar, knots = knots) + bs(alcohol, knots = knots) +
##   bs(drink, knots = knots) + bs(wgain, knots = knots), data = dat_drop)
##
## Residuals:
##      Min       1Q     Median      3Q      Max
## -2885.76  -287.16   -14.85  270.34  2989.09
##
## Coefficients: (100 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1433.515   114.322 12.539 < 2e-16 ***
## bs(stresfip, knots = knots)1 -168.880   158.538 -1.065 0.286772
## bs(stresfip, knots = knots)2  41.321    63.050  0.655 0.512230
## bs(stresfip, knots = knots)3 -69.843   97.371 -0.717 0.473203
## bs(stresfip, knots = knots)4     NA        NA     NA     NA
## bs(stresfip, knots = knots)5     NA        NA     NA     NA
## bs(stresfip, knots = knots)6     NA        NA     NA     NA
## bs(dimage, knots = knots)1  52.308   42.341  1.235 0.216688
## bs(dimage, knots = knots)2 107.886   25.122  4.295 1.75e-05 ***
## bs(dimage, knots = knots)3   8.624   47.409  0.182 0.855661
## bs(dimage, knots = knots)4     NA        NA     NA     NA
## bs(dimage, knots = knots)5     NA        NA     NA     NA
## bs(dimage, knots = knots)6     NA        NA     NA     NA
## bs(ormoth, knots = knots)1 -195.409   45.477 -4.297 1.73e-05 ***
## bs(ormoth, knots = knots)2 -44.661   63.643 -0.702 0.482843
## bs(ormoth, knots = knots)3 -65.251   18.624 -3.504 0.000459 ***
## bs(ormoth, knots = knots)4     NA        NA     NA     NA
## bs(ormoth, knots = knots)5     NA        NA     NA     NA
## bs(ormoth, knots = knots)6     NA        NA     NA     NA
## bs(mrace3, knots = knots)1 -429.623   25.036 -17.160 < 2e-16 ***
## bs(mrace3, knots = knots)2     NA        NA     NA     NA
## bs(mrace3, knots = knots)3 -124.671   4.568 -27.292 < 2e-16 ***
## bs(mrace3, knots = knots)4     NA        NA     NA     NA
## bs(mrace3, knots = knots)5     NA        NA     NA     NA
## bs(mrace3, knots = knots)6     NA        NA     NA     NA
## bs(dmeduc, knots = knots)1 -229.804   82.433 -2.788 0.005308 **
## bs(dmeduc, knots = knots)2 -154.830   48.623 -3.184 0.001451 **
## bs(dmeduc, knots = knots)3 -139.209   54.987 -2.532 0.011353 *
## bs(dmeduc, knots = knots)4     NA        NA     NA     NA
## bs(dmeduc, knots = knots)5     NA        NA     NA     NA
## bs(dmeduc, knots = knots)6     NA        NA     NA     NA
## bs(dmar, knots = knots)1     NA        NA     NA     NA
## bs(dmar, knots = knots)2     NA        NA     NA     NA
## bs(dmar, knots = knots)3 -34.004   4.014 -8.472 < 2e-16 ***
## bs(dmar, knots = knots)4     NA        NA     NA     NA

```

## bs(dmar, knots = knots)5	NA	NA	NA	NA
## bs(dmar, knots = knots)6	NA	NA	NA	NA
## bs(adequacy, knots = knots)1	42.002	11.608	3.618	0.000297 ***
## bs(adequacy, knots = knots)2	NA	NA	NA	NA
## bs(adequacy, knots = knots)3	57.037	13.515	4.220	2.44e-05 ***
## bs(adequacy, knots = knots)4	NA	NA	NA	NA
## bs(adequacy, knots = knots)5	NA	NA	NA	NA
## bs(adequacy, knots = knots)6	NA	NA	NA	NA
## bs(dtotord, knots = knots)1	48.861	27.456	1.780	0.075140 .
## bs(dtotord, knots = knots)2	265.385	99.929	2.656	0.007915 **
## bs(dtotord, knots = knots)3	-160.186	294.553	-0.544	0.586561
## bs(dtotord, knots = knots)4	NA	NA	NA	NA
## bs(dtotord, knots = knots)5	NA	NA	NA	NA
## bs(dtotord, knots = knots)6	NA	NA	NA	NA
## bs(monpre, knots = knots)1	-5.257	25.655	-0.205	0.837647
## bs(monpre, knots = knots)2	2.497	22.620	0.110	0.912097
## bs(monpre, knots = knots)3	55.034	26.240	2.097	0.035969 *
## bs(monpre, knots = knots)4	NA	NA	NA	NA
## bs(monpre, knots = knots)5	NA	NA	NA	NA
## bs(monpre, knots = knots)6	NA	NA	NA	NA
## bs(nprevist, knots = knots)1	398.348	50.872	7.830	4.90e-15 ***
## bs(nprevist, knots = knots)2	184.457	54.354	3.394	0.000690 ***
## bs(nprevist, knots = knots)3	-21.182	102.138	-0.207	0.835706
## bs(nprevist, knots = knots)4	NA	NA	NA	NA
## bs(nprevist, knots = knots)5	NA	NA	NA	NA
## bs(nprevist, knots = knots)6	NA	NA	NA	NA
## bs(disllb, knots = knots)1	-15.566	43.457	-0.358	0.720207
## bs(disllb, knots = knots)2	-694.249	161.723	-4.293	1.77e-05 ***
## bs(disllb, knots = knots)3	-142.069	6.095	-23.310	< 2e-16 ***
## bs(disllb, knots = knots)4	NA	NA	NA	NA
## bs(disllb, knots = knots)5	NA	NA	NA	NA
## bs(disllb, knots = knots)6	NA	NA	NA	NA
## bs(birmon, knots = knots)1	-1.853	11.811	-0.157	0.875312
## bs(birmon, knots = knots)2	-3.058	8.608	-0.355	0.722424
## bs(birmon, knots = knots)3	-7.276	5.955	-1.222	0.221808
## bs(birmon, knots = knots)4	NA	NA	NA	NA
## bs(birmon, knots = knots)5	NA	NA	NA	NA
## bs(birmon, knots = knots)6	NA	NA	NA	NA
## bs(dgestat, knots = knots)1	-1172.652	102.426	-11.449	< 2e-16 ***
## bs(dgestat, knots = knots)2	3642.377	48.898	74.490	< 2e-16 ***
## bs(dgestat, knots = knots)3	2169.008	66.484	32.625	< 2e-16 ***
## bs(dgestat, knots = knots)4	NA	NA	NA	NA
## bs(dgestat, knots = knots)5	NA	NA	NA	NA
## bs(dgestat, knots = knots)6	NA	NA	NA	NA
## bs(csex, knots = knots)1	NA	NA	NA	NA
## bs(csex, knots = knots)2	NA	NA	NA	NA
## bs(csex, knots = knots)3	-133.070	2.590	-51.371	< 2e-16 ***
## bs(csex, knots = knots)4	NA	NA	NA	NA
## bs(csex, knots = knots)5	NA	NA	NA	NA
## bs(csex, knots = knots)6	NA	NA	NA	NA
## bs(dplural, knots = knots)1	-1090.733	73.959	-14.748	< 2e-16 ***
## bs(dplural, knots = knots)2	-362.195	153.810	-2.355	0.018533 *
## bs(dplural, knots = knots)3	-1344.040	126.548	-10.621	< 2e-16 ***
## bs(dplural, knots = knots)4	NA	NA	NA	NA

```

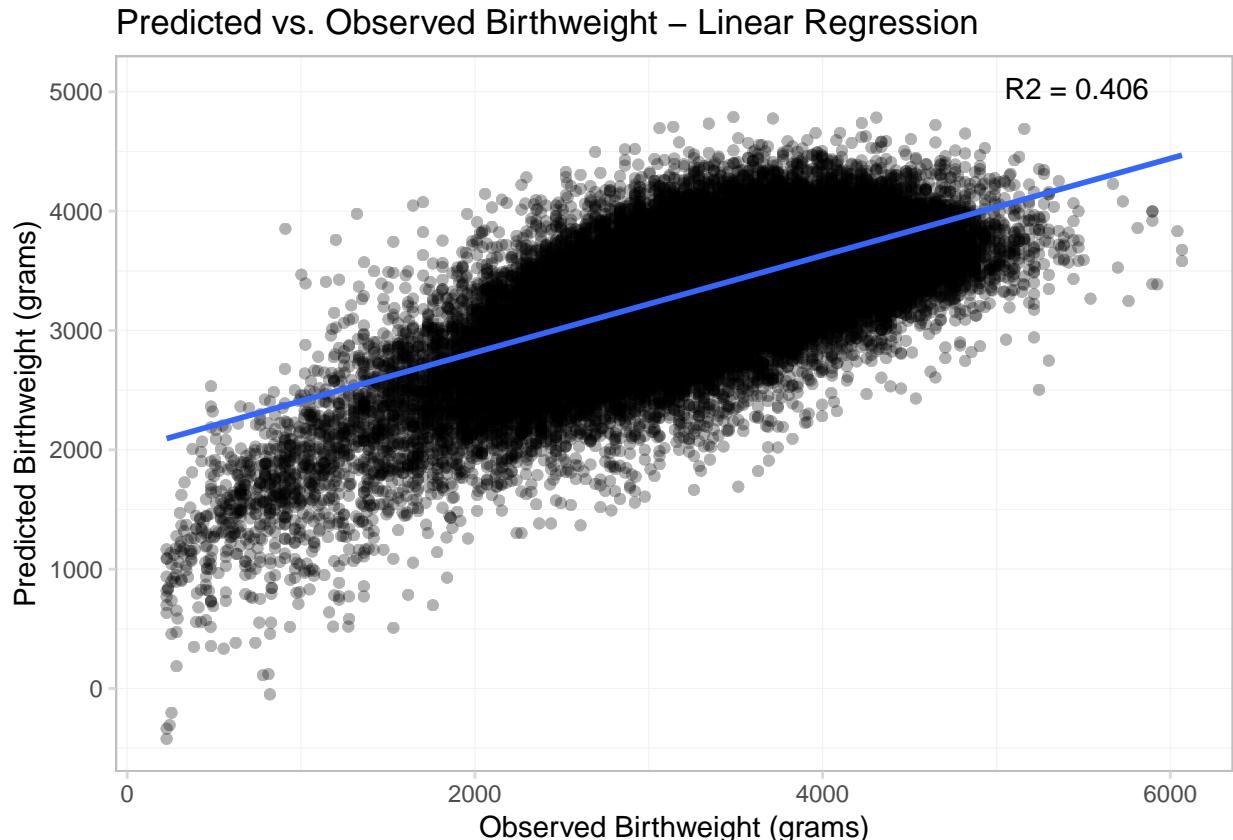
## bs(dplural, knots = knots)5      NA      NA      NA
## bs(dplural, knots = knots)6      NA      NA      NA
## bs(anemia, knots = knots)1      NA      NA      NA
## bs(anemia, knots = knots)2      NA      NA      NA
## bs(anemia, knots = knots)3     -25.175   13.057  -1.928  0.053853 .
## bs(anemia, knots = knots)4      NA      NA      NA
## bs(anemia, knots = knots)5      NA      NA      NA
## bs(anemia, knots = knots)6      NA      NA      NA
## bs(diabetes, knots = knots)1    NA      NA      NA
## bs(diabetes, knots = knots)2    NA      NA      NA
## bs(diabetes, knots = knots)3   -157.465   8.104  -19.431 < 2e-16 ***
## bs(diabetes, knots = knots)4    NA      NA      NA
## bs(diabetes, knots = knots)5    NA      NA      NA
## bs(diabetes, knots = knots)6    NA      NA      NA
## bs(herpes, knots = knots)1      NA      NA      NA
## bs(herpes, knots = knots)2      NA      NA      NA
## bs(herpes, knots = knots)3    -8.711   16.574  -0.526  0.599185
## bs(herpes, knots = knots)4      NA      NA      NA
## bs(herpes, knots = knots)5      NA      NA      NA
## bs(herpes, knots = knots)6      NA      NA      NA
## bs(chyper, knots = knots)1      NA      NA      NA
## bs(chyper, knots = knots)2      NA      NA      NA
## bs(chyper, knots = knots)3   104.762   14.881   7.040  1.93e-12 ***
## bs(chyper, knots = knots)4      NA      NA      NA
## bs(chyper, knots = knots)5      NA      NA      NA
## bs(chyper, knots = knots)6      NA      NA      NA
## bs(pre4000, knots = knots)1      NA      NA      NA
## bs(pre4000, knots = knots)2      NA      NA      NA
## bs(pre4000, knots = knots)3   -375.513   10.925  -34.371 < 2e-16 ***
## bs(pre4000, knots = knots)4      NA      NA      NA
## bs(pre4000, knots = knots)5      NA      NA      NA
## bs(pre4000, knots = knots)6      NA      NA      NA
## bs(preterm, knots = knots)1      NA      NA      NA
## bs(preterm, knots = knots)2      NA      NA      NA
## bs(preterm, knots = knots)3   228.203   11.084   20.588 < 2e-16 ***
## bs(preterm, knots = knots)4      NA      NA      NA
## bs(preterm, knots = knots)5      NA      NA      NA
## bs(preterm, knots = knots)6      NA      NA      NA
## bs(tobacco, knots = knots)1      NA      NA      NA
## bs(tobacco, knots = knots)2      NA      NA      NA
## bs(tobacco, knots = knots)3   -109.105   10.790  -10.112 < 2e-16 ***
## bs(tobacco, knots = knots)4      NA      NA      NA
## bs(tobacco, knots = knots)5      NA      NA      NA
## bs(tobacco, knots = knots)6      NA      NA      NA
## bs(cigar, knots = knots)1     -406.491   60.121  -6.761  1.38e-11 ***
## bs(cigar, knots = knots)2     307.799   156.863   1.962  0.049739 *
## bs(cigar, knots = knots)3    -420.037   338.071  -1.242  0.214072
## bs(cigar, knots = knots)4      NA      NA      NA
## bs(cigar, knots = knots)5      NA      NA      NA
## bs(cigar, knots = knots)6      NA      NA      NA
## bs(alcohol, knots = knots)1      NA      NA      NA
## bs(alcohol, knots = knots)2      NA      NA      NA
## bs(alcohol, knots = knots)3   -15.498   21.579  -0.718  0.472641
## bs(alcohol, knots = knots)4      NA      NA      NA

```

```

## bs(alcohol, knots = knots)5      NA      NA      NA
## bs(alcohol, knots = knots)6      NA      NA      NA
## bs(drink, knots = knots)1     -796.275 242.324 -3.286 0.001017 **
## bs(drink, knots = knots)2      963.025 706.753 1.363 0.173010
## bs(drink, knots = knots)3     751.333 362.511 2.073 0.038214 *
## bs(drink, knots = knots)4      NA      NA      NA
## bs(drink, knots = knots)5      NA      NA      NA
## bs(drink, knots = knots)6      NA      NA      NA
## bs(wgain, knots = knots)1    -115.327 25.471 -4.528 5.97e-06 ***
## bs(wgain, knots = knots)2      846.225 26.244 32.245 < 2e-16 ***
## bs(wgain, knots = knots)3     225.123 44.980 5.005 5.60e-07 ***
## bs(wgain, knots = knots)4      NA      NA      NA
## bs(wgain, knots = knots)5      NA      NA      NA
## bs(wgain, knots = knots)6      NA      NA      NA
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 437.7 on 114553 degrees of freedom
## Multiple R-squared: 0.4409, Adjusted R-squared: 0.4406
## F-statistic: 1613 on 56 and 114553 DF, p-value: < 2.2e-16
ggscat(x = dat_drop$dbrwt, y = lm.mod$fitted.values, ylab = "Predicted Birthweight (grams)", xlab = "Observe

```

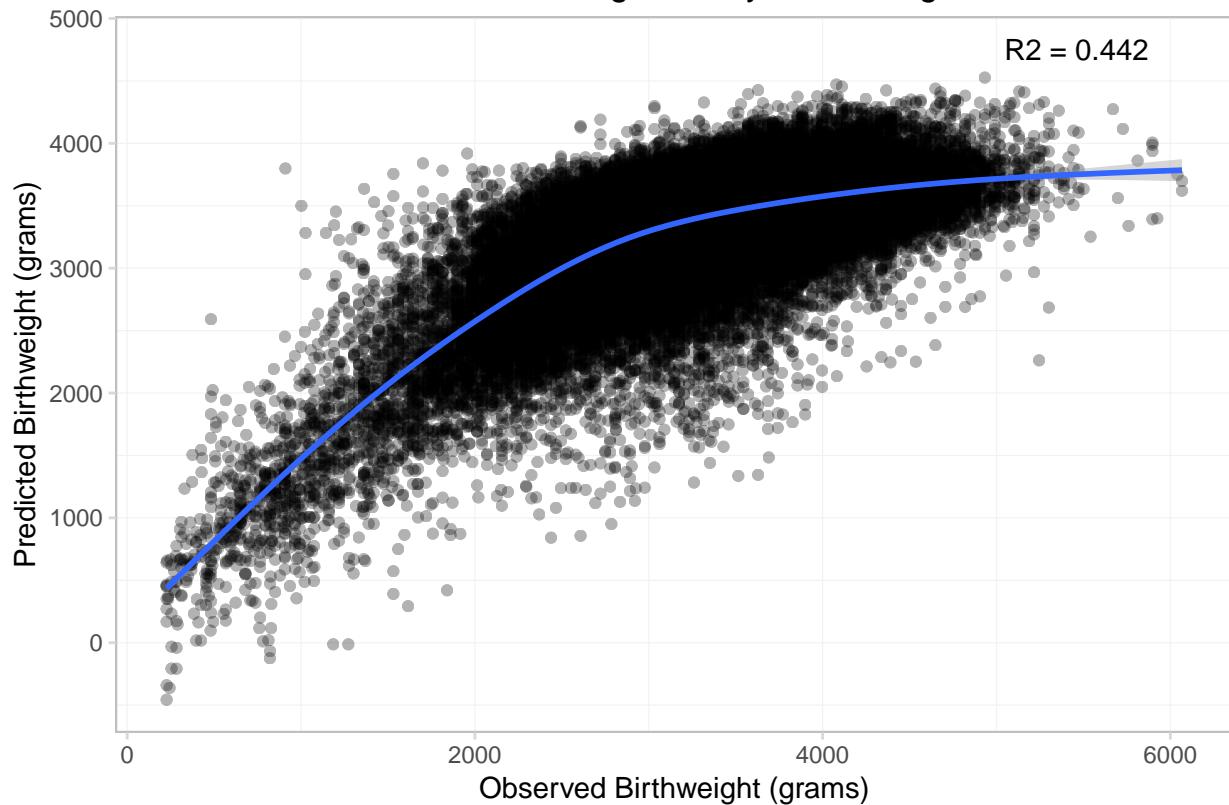


```

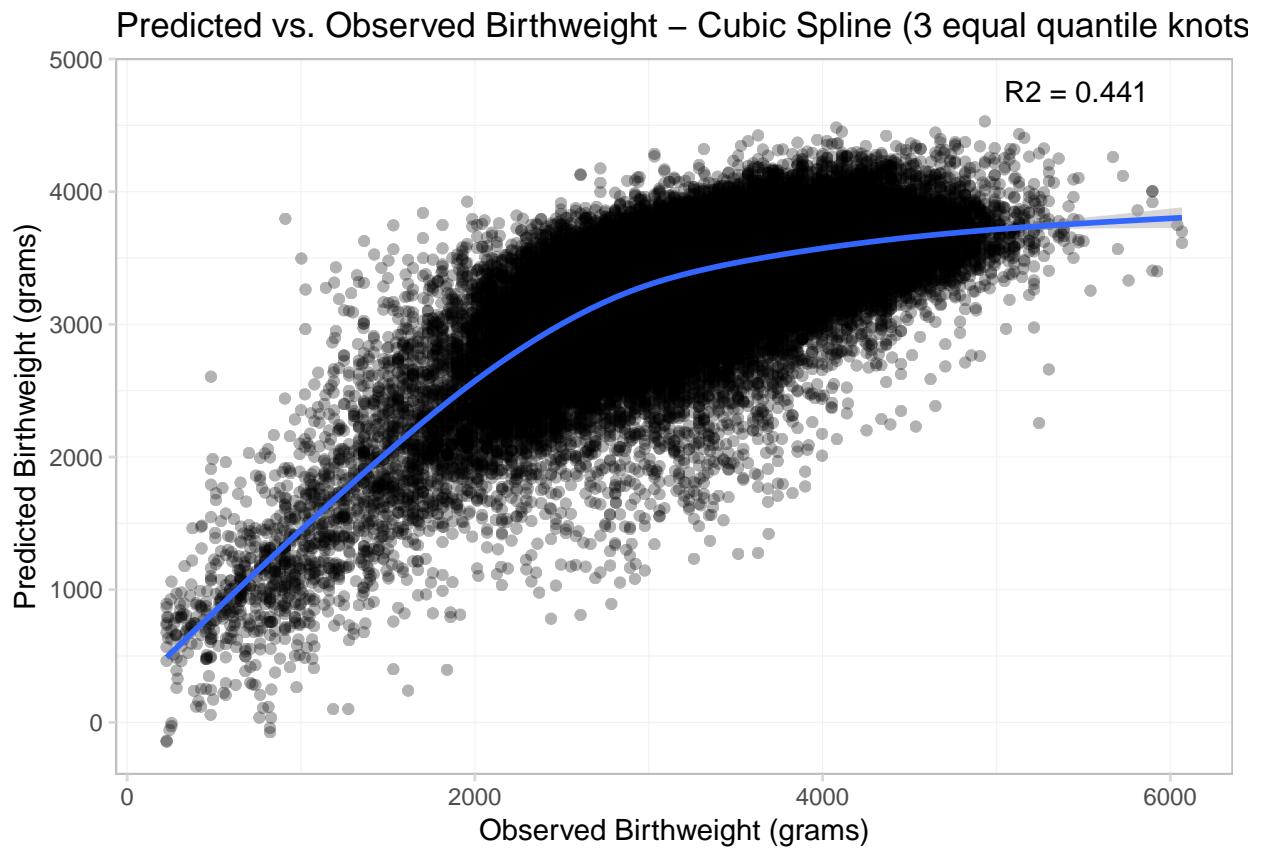
ggscat(x = dat_drop$dbrwt, y = poly.mod$fitted.values, ylab = "Predicted Birthweight (grams)", xlab = "Observe
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'

```

Predicted vs. Observed Birthweight – Polynomial Regression



```
ggscat(x = dat_drop$dbrwt, y =cspline.mod$fitted.values, ylab = "Predicted Birthweight (grams)", xlab =  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



1c - Using LASSO

```

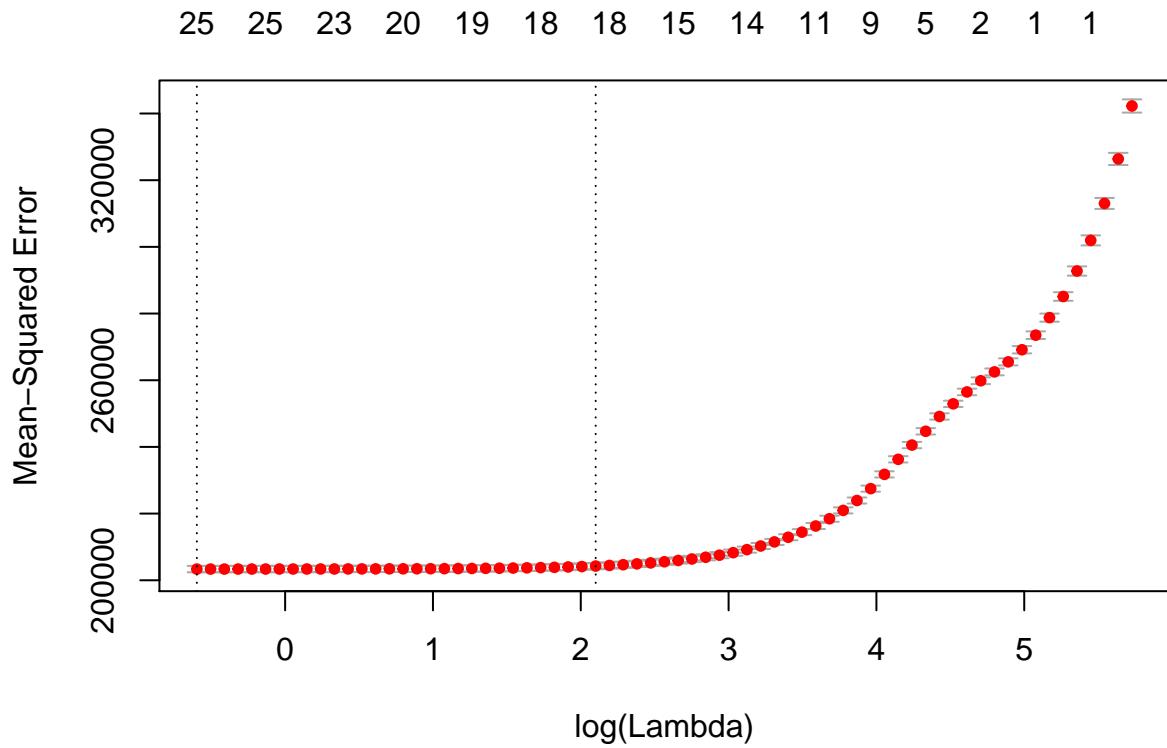
n <- nrow(dat.mod)
k <- ncol(dat.mod)
y <- dat.mod$dbrwt
D <- dat.mod$tobacco

# X-matrix for the 1st step of Lasso with treatment as the outcome variable
xmat0 <- model.matrix(tobacco ~., data = dat.mod %>% select(-dbrwt))[, -1]

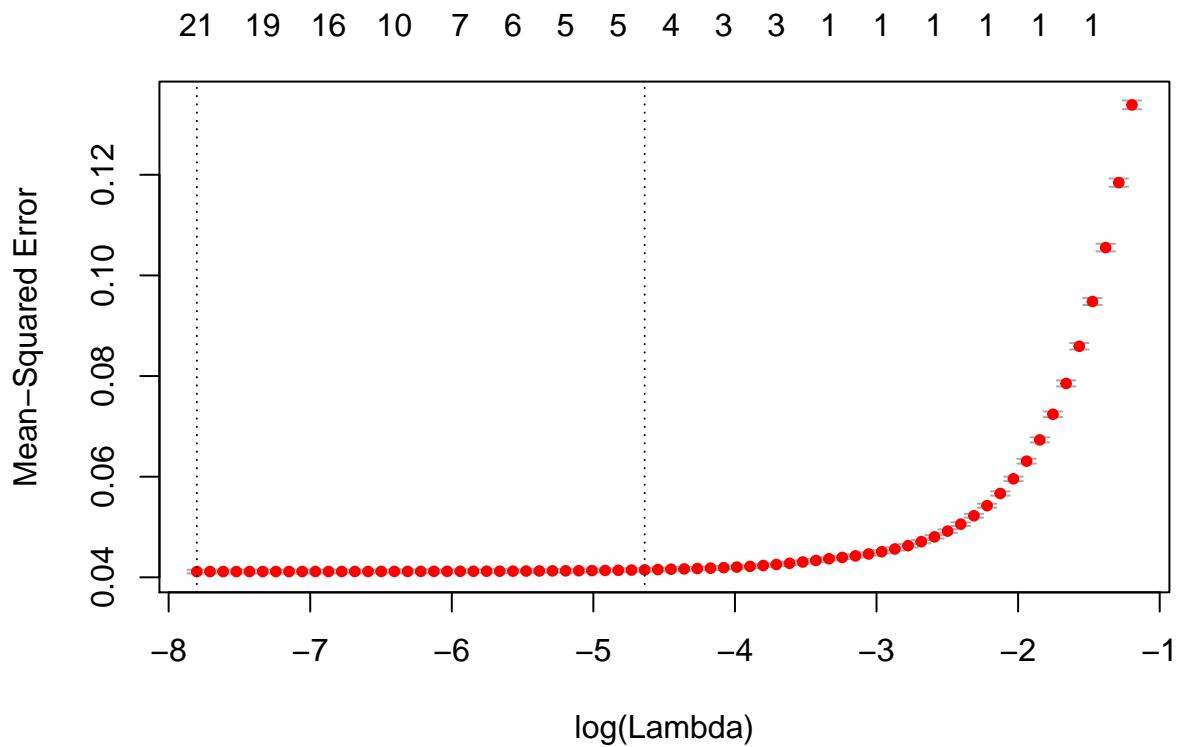
# X-matrix for the 2nd step of Lasso with original outcome as the outcome variable
xmat1 <- model.matrix(dbrwt ~., data = dat.mod) [, -1]

# Test which lambdas to use generally
cv.out <- cv.glmnet(xmat1, y, alpha = 1)
plot(cv.out)

```



```
# Treatment as Outcome
# Cross-Validation to Identify Lambdas to Use
cv.out <- cv.glmnet(xmat0, D, alpha = 1)
plot(cv.out)
```

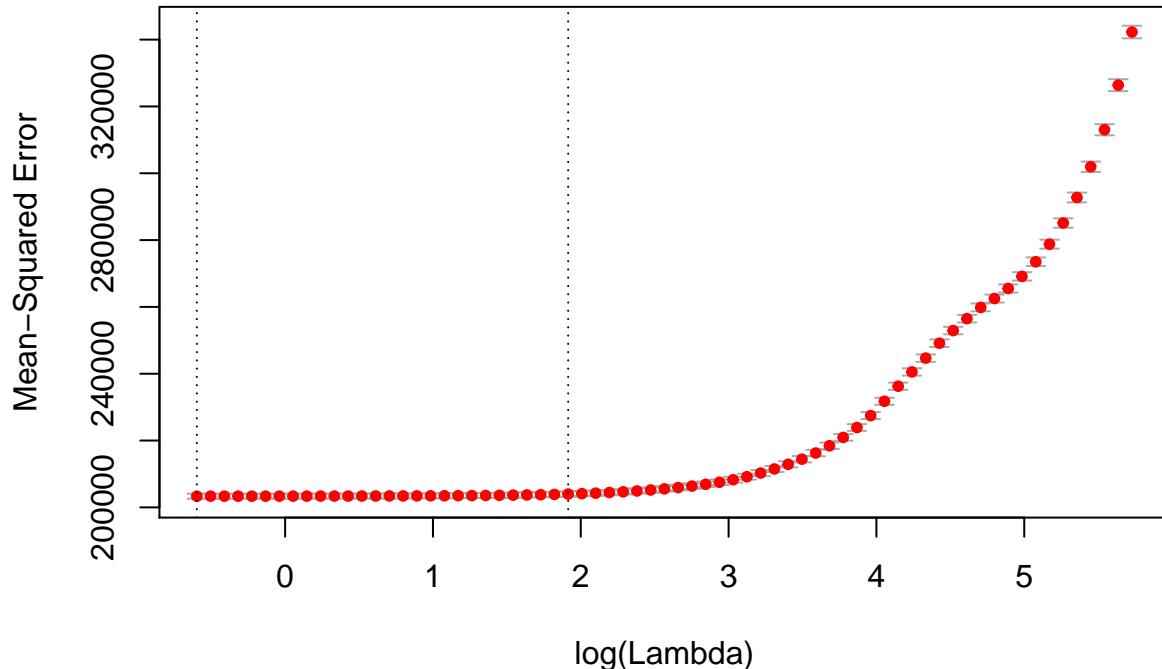


```
lasso.mod0 <- glmnet(xmat0, D, alpha = 1, lambda = cv.out$lambda.1se)
Dvars_save <- lasso.mod0$beta[abs(lasso.mod0$beta[,1])>0,1]
```

```
Dvars_save
```

```
##      dmage      dmeduc      dmar      cigar      alcohol
## -0.0001861988 -0.0044465717  0.0406821785  0.0542266793 -0.0218221446
# Treatment as Outcome
# Cross-Validation to Identify Lambdas to Use
cv.out <- cv.glmnet(xmat1, y, alpha = 1)
plot(cv.out)
```

25 25 23 20 19 18 18 15 14 11 9 5 2 1 1



```
lasso.mod1 <- glmnet(xmat1, y, alpha = 1, lambda = cv.out$lambda.1se)
Ovars_save <- lasso.mod1$beta[abs(lasso.mod1$beta[,1])>0,1]
Ovars_save
```

```
##      dmage      ormoth      mrace3      dmeduc      dmar
## 1.8295865 -14.0165257 -58.6392998  2.6493479 -48.6331403
##      dtotord      nprevist      disllb      dgestat      csex
## 4.3179275  6.3337875 -0.1808722 105.9293884 -119.5123683
##      dplural      diabetes      chyper      pre4000      preterm
## -600.5104409 -115.9360966  44.3949148 -336.7837430  194.1447484
##      tobacco      cigar      wgain
## -153.3579309 -2.8427989   7.6797993
c(names(Dvars_save)[!which(names(Dvars_save) %in% names(Ovars_save))],names(Ovars_save))

## [1] "dmage"      "orthoth"    "mrace3"     "dmeduc"     "dmar"       "dtotord"
## [7] "nprevist"   "disllb"     "dgestat"    "csex"       "dplural"    "diabetes"
## [13] "chyper"     "pre4000"    "preterm"    "tobacco"    "cigar"      "wgain"
```

2 - Propensity score description

2a - Create propensity score

Predetermined variables used for propensity score calculation using logit.

```
pd.vars <- c("tobacco", "dimage", "ormoth", "mrace3", "dmeduc", "dmar", "dtotord", "disllb", "anemia", "diabetes")
kable(vartab[row.names(vartab) %in% pd.vars[-1],], caption = "Predetermined Variables Selected as Covariates for the logit Propensity Score Calculation")
```

Table 1: Predetermined Variables Selected as Covariates for the logit Propensity Score Calculation

	Variables	Mean	SD	Min	Max
dimage	Age of Mother	27.757	5.699	12	49
ormoth	Hispanic Origin of Mother	0.091	0.522	0	5
mrace3	Race of Mother Recode	1.259	0.657	1	3
dmeduc	Education of Mother Detail	13.211	2.272	0	17
dmar	Marital Status of Mother	1.251	0.434	1	2
dtotord	Detail Total Birth Order	2.420	1.520	1	24
disllb	Interval Since Last Live Birth	350.412	362.325	0	777
anemia	Anemia	1.990	0.099	1	2
diabetes	Diabetes	1.973	0.162	1	2
herpes	Genital Herpes	1.994	0.078	1	2
chyper	Chronic Hypertension	1.992	0.087	1	2
pre4000	Previous Infant 4000+ Grams	1.986	0.119	1	2

Using logit

```
##
## Call:
## glm(formula = tobacco ~ ., family = binomial, data = dat.logit)
##
## Deviance Residuals:
##      Min        1Q     Median        3Q       Max
## -2.6077  -0.5892  -0.4362  -0.3012   3.2660
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.122e+00 3.955e-01 -2.837 0.00455 ***
## dimage      -4.123e-03 1.970e-03 -2.093 0.03633 *
## ormoth     -5.566e-01 2.248e-02 -24.763 < 2e-16 ***
## mrace3     -5.059e-01 1.404e-02 -36.020 < 2e-16 ***
## dmeduc     -2.372e-01 4.699e-03 -50.482 < 2e-16 ***
## dmar        1.311e+00 2.168e-02  60.447 < 2e-16 ***
## dtotord     1.262e-01 6.782e-03  18.608 < 2e-16 ***
## disllb     -3.696e-04 3.047e-05 -12.128 < 2e-16 ***
## anemia     -3.238e-02 7.785e-02 -0.416 0.67743
## diabetes    -8.303e-02 5.315e-02 -1.562 0.11823
## herpes      -1.502e-01 1.067e-01 -1.408 0.15903
## chyper      6.957e-02 1.036e-01  0.672 0.50174
## pre4000     8.578e-01 8.937e-02  9.598 < 2e-16 ***
## ---
```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 100543  on 114609  degrees of freedom
## Residual deviance:  88626  on 114597  degrees of freedom
## AIC: 88652
##
## Number of Fisher Scoring iterations: 5

```

Now only with significant predetermined covariates from first logit model above.

```

pd.vars2 <- names(ps.mod$coefficients[which(summary(ps.mod)$coefficients[,4]<0.05)])
pd.vars2 <- pd.vars2[pd.vars2 != "(Intercept)"]
kable(vartab[row.names(vartab) %in% pd.vars2], caption = "Significant Predetermined Variables Used for"

```

Table 2: Significant Predetermined Variables Used for Second Propensity Score Calculation

	Variables	Mean	SD	Min	Max
dmage	Age of Mother	27.757	5.699	12	49
ormoth	Hispanic Origin of Mother	0.091	0.522	0	5
mrace3	Race of Mother Recode	1.259	0.657	1	3
dmeduc	Education of Mother Detail	13.211	2.272	0	17
dmar	Marital Status of Mother	1.251	0.434	1	2
dtotord	Detail Total Birth Order	2.420	1.520	1	24
disllb	Interval Since Last Live Birth	350.412	362.325	0	777
pre4000	Previous Infant 4000+ Grams	1.986	0.119	1	2

Propensity score calculation using only the significant variables

```

dat.logit2 <- dat.logit[,c("tobacco",pd.vars2)]
ps.mod2 <- glm(tobacco ~ ., data = dat.logit2, family = binomial)
ps2 <- ps.mod2$fitted.values
summary(ps.mod2)

```

```

##
## Call:
## glm(formula = tobacco ~ ., family = binomial, data = dat.logit2)
##
## Deviance Residuals:
##      Min        1Q     Median        3Q       Max
## -2.5797  -0.5870  -0.4361  -0.3017   3.2643
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.508e+00  1.944e-01 -7.760  8.5e-15 ***
## dmage       -3.974e-03  1.960e-03 -2.027  0.0426 *
## ormoth      -5.560e-01  2.246e-02 -24.756 < 2e-16 ***
## mrace3      -5.054e-01  1.403e-02 -36.011 < 2e-16 ***
## dmeduc      -2.373e-01  4.697e-03 -50.520 < 2e-16 ***
## dmar        1.311e+00  2.168e-02  60.491 < 2e-16 ***
## dtotord     1.263e-01  6.781e-03  18.632 < 2e-16 ***
## disllb      -3.686e-04  3.046e-05 -12.100 < 2e-16 ***

```

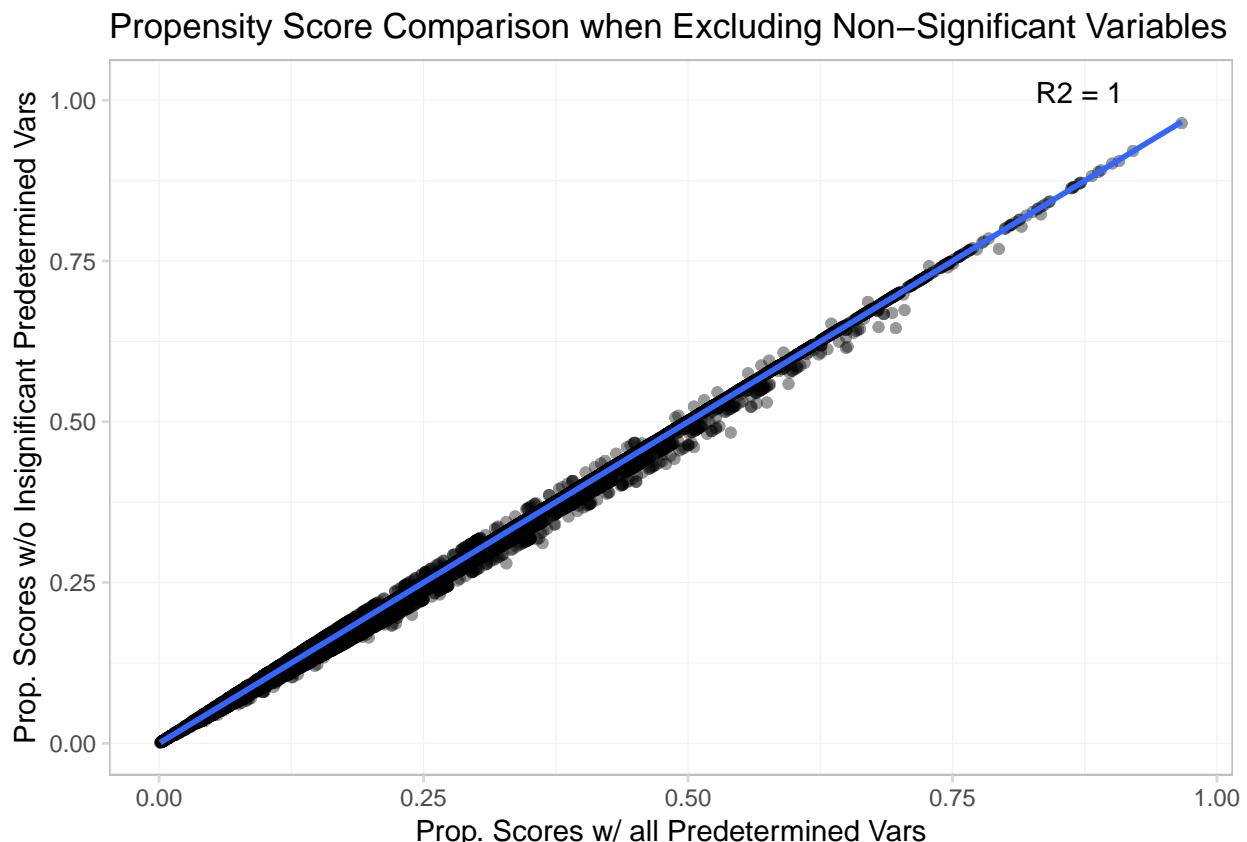
```

## pre4000      8.536e-01  8.932e-02   9.557  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 100543  on 114609  degrees of freedom
## Residual deviance: 88631  on 114601  degrees of freedom
## AIC: 88649
##
## Number of Fisher Scoring iterations: 5

```

We can see that excluding the non-significant covariates in the first propensity score estimation has a little overall effect on the propensity scores of being treated ($R^2 = 1$).

```
ggscat(ps1,ps2, xlab = "Prop. Scores w/ all Predetermined Vars", ylab = "Prop. Scores w/o Insignificant Vars")
```



Since these values of propensity scores are so similar, we can conclude more or less that we are including the “correct” set of predetermined variables in the smaller set of just those that enter significantly ($p < 0.05$).

2b - Regression Adjusting Propensity Score Estimation

```

dat.ps.reg <- dat.mod %>% select(c("dbrwt","tobacco")) %>% mutate(PropScore = ps2)
mod.2b <- lm(dbrwt ~., data = dat.ps.reg)
summary(mod.2b)

```

```
##
```

```

## Call:
## lm(formula = dbrwt ~ ., data = dat.ps.reg)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3200.2  -309.1    29.8  361.9 2752.5
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3435.035    2.775 1237.79 <2e-16 ***
## tobacco     -221.924    4.943 -44.90 <2e-16 ***
## PropScore   -165.491   14.537 -11.38 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 578.2 on 114607 degrees of freedom
## Multiple R-squared:  0.02373, Adjusted R-squared:  0.02371
## F-statistic: 1393 on 2 and 114607 DF, p-value: < 2.2e-16

```

In this case of conditioning on the propensity scores and the treatment, the ATE of smoking while pregnant has risen to roughly -222 grams (i.e. smoking while pregnant will reduce infant birthweight by 222g, on average). However, for this approach to be sound, we assume conditions of unconfoundedness, overlap of treatment and controls in the covariate space, and a constant/homogenous treatment effect.

2c - Propensity Score Reweighting

Average treatment effect with reweighting propensity score:

```

#Redefine final propensity score variable
pX <- ps2

# Define ATE with reweighted propensity score scheme
pATE <- sum(D*y/pX)/sum(D/pX) - sum((1-D)*y/(1-pX))/sum((1-D)/(1-pX))

## [1] -222.9179

```

We can see that the ATE of -222.9179012 using this reweighted propensity score scheme returns a very similar value to the regression adjusted propensity score process above, suggesting that reweighting in this scenario may not make a large difference in the ATE estimate.

Average effect of the treatment on the treated

```

TOT <- sum(D*y)/sum(D) - sum((pX*(1-D)*y)/(1-pX))/sum((pX*(1-D))/(1-pX))

## [1] -220.9535

```

2d&e - kernel density estimator

For treatment

```

#Number of Control Obs
nC <- length(D[D==0])

```

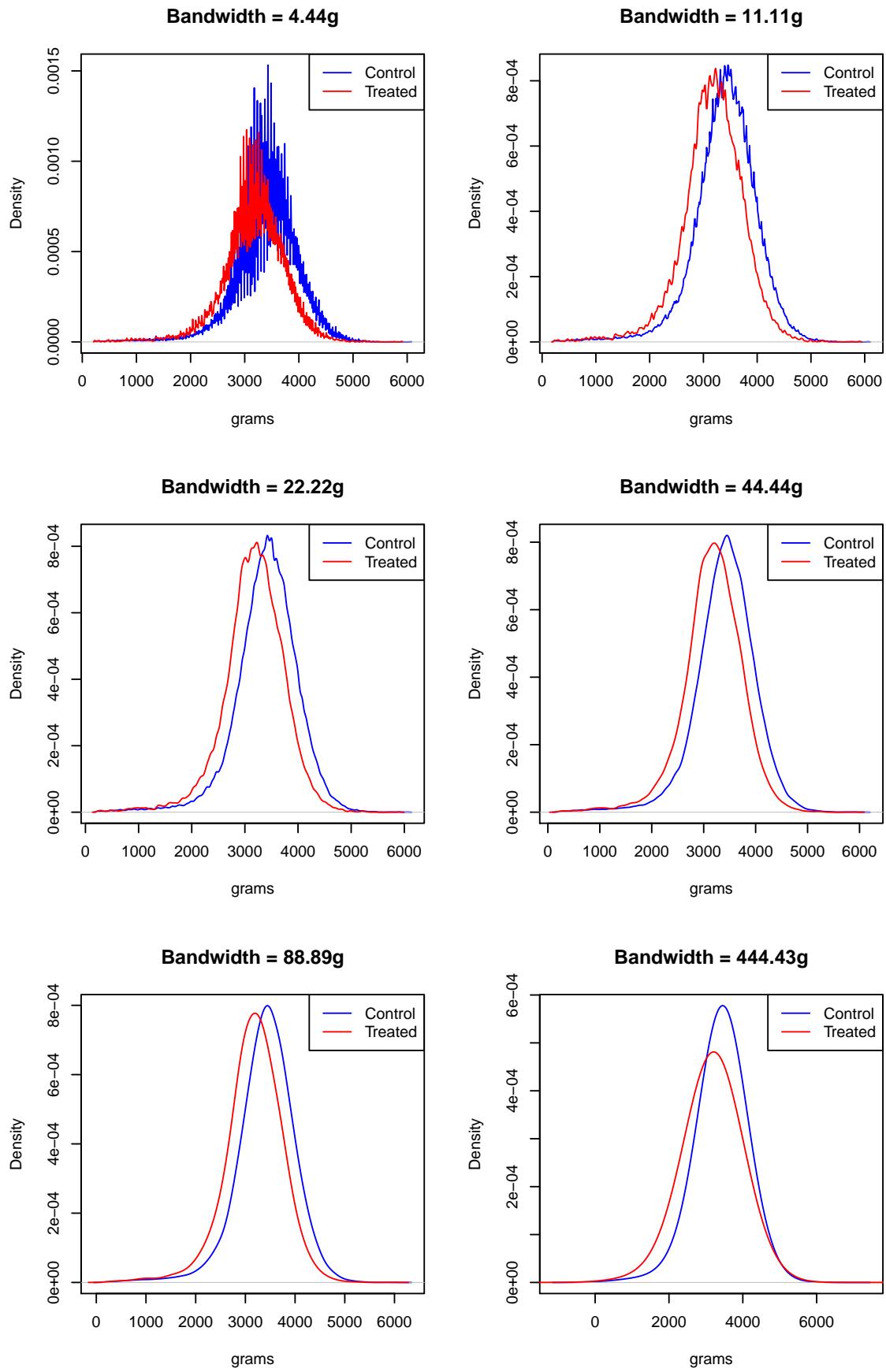
```

nT <- length(D[D==1])
par(mfrow = c(3,2))
for(adj in c(0.1,0.25,0.5,1,2,10)){
# Define Bandwidth
h0 <- density(y)$bw
bw_adjust <- adj
h <- h0*bw_adjust

# Control & Treatment Densities
Cdens <- density(y[D==0], adjust = bw_adjust)
Tdens <- density(y[D==1], adjust = bw_adjust)

#Plot densities
plot(Cdens, col = "blue", main = paste0("Bandwidth = ",round(h,2),"g"), xlab = "grams")
lines(Tdens, col = "red")
legend('topright', legend = c("Control", "Treated"), col = c("blue", "red"), lty = 1)
}

```



The default bandwidth of ~44 grams seems to produce a fairly smooth kernel density, and was selected as the preferred bandwidth.

```

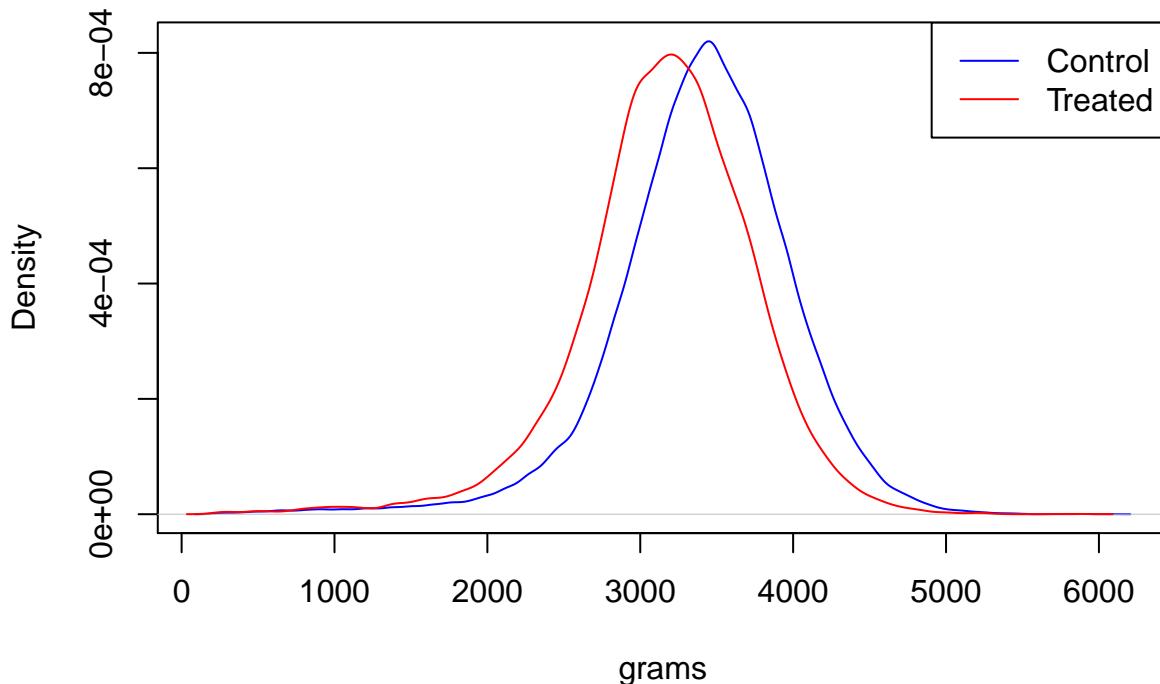
h0 <- density(y)$bw
bw_adjust <- 1
h <- h0*bw_adjust

# Control & Treatment Densities
Cdens <- density(y[D==0], adjust = bw_adjust)
Tdens <- density(y[D==1], adjust = bw_adjust)

#Plot densities
plot(Cdens, col = "blue", main = paste0("Kernel Densities of Birthweight (Bandwidth = ",round(h,2),"g)"))
lines(Tdens, col = "red")
legend('topright', legend = c("Control", "Treated"), col = c("blue", "red"), lty = 1)

```

Kernel Densities of Birthweight (Bandwidth = 44.44g)



Solving for kernel estimator at birthweight = 3000 grams.

```

# Uniform kernel function
k.uni <- function(u){
  return(ifelse(abs(u) <= 1,0.5,0))
}

# Triangular kernel function
k.tri <- function(u){
  return(ifelse(abs(u) < 1,1-abs(u),0))
}

# Epanechnikov kernel function
k.epan <- function(u){
  return(ifelse(abs(u) < 1,0.75*(1-u^2),0))
}

```

```

}

# Gaussian kernel function
k.gauss <- function(u){
  return((1/sqrt(2*pi))*exp(-0.5*u^2))
}

# Kernel Density Estimator for Control Group
k.dens.C <- function(ystar, y, h, N, k.func){
  ks <- rep(NA,N)
  for(i in 1:N){
    u <- (ystar - y[i])/h
    ks[i] <- k.func(u)
  }
  return((1/(N*h))*(sum(ks)))
}

# Kernel Density Estimator for Treatment Group
k.dens.C <- function(ystar, k.func){
  ks <- rep(NA,nC)
  for(i in 1:nC){
    u <- (ystar - y[i])/h
    ks[i] <- k.func(u)*D[i]/pX[i]
  }
  return((1/(nC*h))*(sum(ks)))
}

# Kernel Density Estimator for Treatment Group
k.dens.T <- function(ystar, k.func){
  ks <- rep(NA,nT)
  for(i in 1:nT){
    u <- (ystar - y[i])/h
    ks[i] <- k.func(u)*(1-D[i])/(1-pX[i])
  }
  return((1/(nT*h))*(sum(ks)))
}

```

Kernal Estimates at 3000g

```

h <- density(y)$bw

est.uni.C <- k.dens.C(3000,k.func = k.uni) %>% round(6)
est.tri.C <- k.dens.C(3000,k.func = k.tri) %>% round(6)
est.epan.C <- k.dens.C(3000,k.func = k.epan) %>% round(6)
est.gauss.C <- k.dens.C(3000,k.func = k.gauss) %>% round(6)

est.uni.T <- k.dens.T(3000,k.func = k.uni) %>% round(6)
est.tri.T <- k.dens.T(3000,k.func = k.tri) %>% round(6)
est.epan.T <- k.dens.T(3000,k.func = k.epan) %>% round(6)
est.gauss.T <- k.dens.T(3000,k.func = k.gauss) %>% round(6)

est.df <- data.frame(Kernal = c("Uniform","Triangular","Epanechnikov","Gaussian"), NonSmoker = c(est.uni.C, est.tri.C, est.epan.C, est.gauss.C), Smoker = c(est.uni.T, est.tri.T, est.epan.T, est.gauss.T))
kable(est.df, caption = "Kernal Estimates for Birthweight of 3000g")

```

Table 3: Kernal Estimates for Birthweight of 3000g

Kernal	NonSmoker	Smoker
Uniform	0.000735	0.000488
Triangular	0.000835	0.000506
Epanechnikov	0.000814	0.000508
Gaussian	0.000744	0.000510

2e - bandwidth adjustments

In the above figure, we could see that with smaller bandwidths, the kernal densities get much more jagged, or variable. However, as we increase the bandwidth and the curves become smoother, it's also evident that there is more bias in the estimates because the magnitudes of the curves begin to change relative to one another (i.e. the peak of the control curve starts to decline as the bandwidth is increased).

2f - benefits and drawbacks of propensity method

The benefits of the propensity weighting approach in part c is that it allows for conditioning solely on the likelihood of selecting into treatment, rather than on all predetermined control variables, which rectifies issues that arise when matching across many variables (Curse of Dimensionality). However, some drawbacks are that doing this process requires a number of assumptions to hold, such as unconfoundedness, overlap between treated and controls, and homogenous treatment effects.

2g - Present and discuss results

Assumptions 1 & 4 should hold I think (?)

3 - Blocking non-parametric approach

```

bins <- seq(0,1,by = 0.01)
vals <- rep(NA, 100)

for(K in 1:100){
  bATE <- mean(y[pX >= bins[K] & pX < bins[K+1] & dat.mod$tobacco == 1]) - mean(y[pX >= bins[K] & pX <
  vals[K] <- bATE*length(y[pX >= bins[K] & pX < bins[K+1]])/length(y)
}
sum(vals[!is.nan(vals)]))

## [1] -220.2786

```

The ATE from this blocking approach is very similar to that found using propensity scores (very slightly smaller).

4 - low birth weights

```
low.brwt <- dat.mod$dbrwt
low.brwt[low.brwt < 2500] <- 1
low.brwt[low.brwt >= 2500] <- 0

vals <- rep(NA, 100)
for(K in 1:100){
  bATE <- mean(low.brwt[pX >= bins[K] & pX < bins[K+1] & dat.mod$tobacco == 1]) - mean(low.brwt[pX >= b
  vals[K] <- bATE*length(low.brwt[pX >= bins[K] & pX < bins[K+1]]))/length(low.brwt)
}
sum(vals[!is.nan(vals)])
```

[1] 0.03899595

This shows that smoking while pregnant increases the likelihood of having a low birthweight (<2500g) roughly 3.9 percent.

5 - Summarize results