PE-PSet2

Drazzel Feliu - 12174100

For this assignment, provide a write–up where you answer the questions below, selectively cutting and pasting output where needed. Be concise in your write-up; excess wordiness will be penalized. Also, submit a log file that includes commands and results for your entire analysis. The assignment makes use of almond_etal_2008.dta, which you can find on Canvas.

```
# load data set
data <- read_dta("almond_etal_2008.dta")
# create data table identifying class and labels for each variable
datainfo <- data.frame(variable=colnames(data),
              class=sapply(data, class),
              label=sapply(data, function (x) attr(x, "label")))
datainfo <- `rownames<-`(datainfo, 1:44)
datainfo
# summary statistics of variables
summary(data)
```

Motivation

A key policy question in health economics is whether the benefits of additional medical expenditures exceed their cost. The question is particularly relevant since medical expenditures in the United States have been on the rise for a long time. To analyze this question Almond et al (2008), use a RDD design and compare health outcomes of newborns around the threshold of very low birth weight (1500 grams).They argue that the threshold is commonly used as a rule of thumb to prescribe medical treatment, which is followed mainly by convention, and does not reflect biological criteria. In this problem set we will reproduce some of their basic results,so start by reading their paper, which you can find in Canvas.

Questions:

1

Start by getting the descriptive statistics of birth weight in the sample, what is the mean, standard deviation, minimum, and maximum?

```
mean(data$bweight)
sd(data$bweight)
min(data$bweight)
max(data$bweight)
```

Answer

The mean birth weight is 1511.58 grams. The standard deviation is 89.02. The minimum birth weight is 1350 grams. The maximum birth weight is 1650 grams.

2

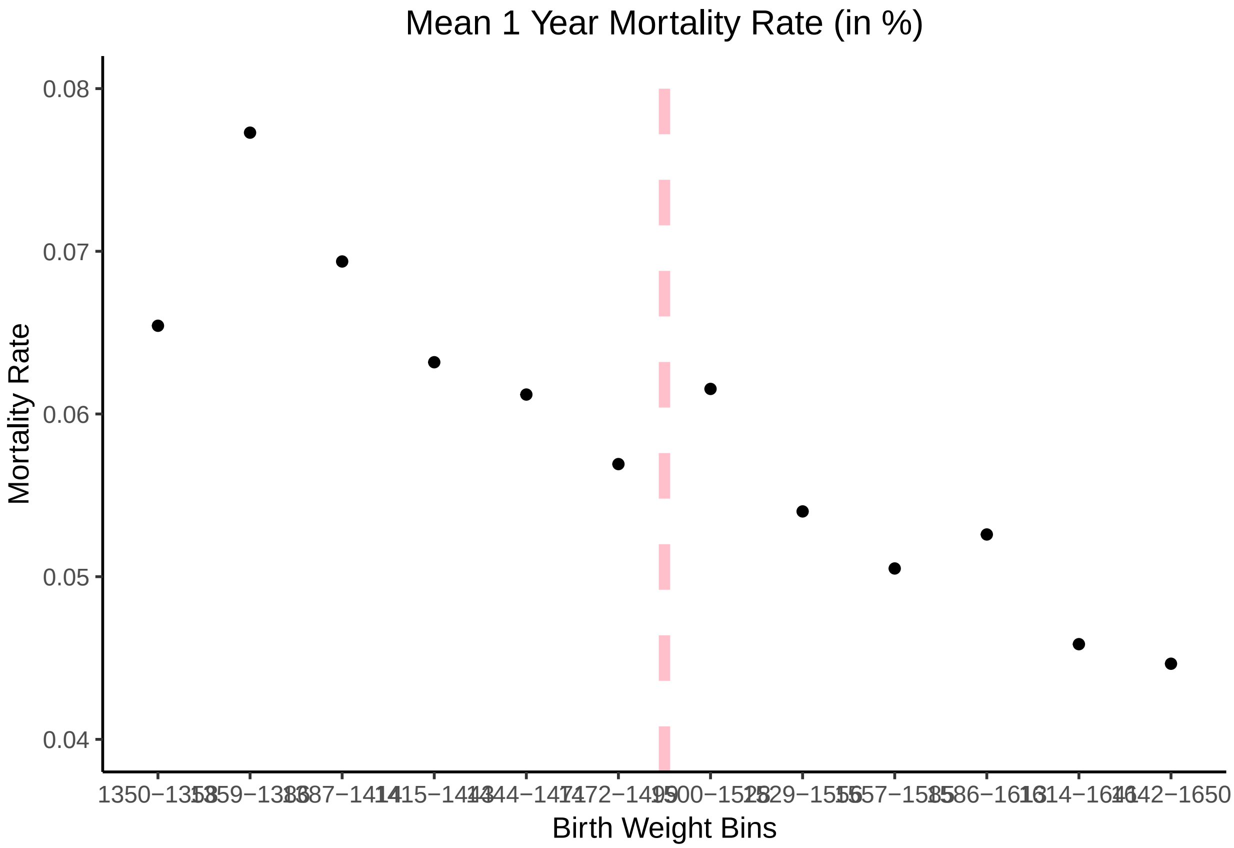
Now plot one year and 28 day mortality rates against our running variable, birth weight. To do so, make bins of one ounce (28.35 grams) around the 1500 grams threshold, and get the mean mortality rate on each bin. Make a separate graph for each outcome. Describe the relationship between birth weight and mortality. Does it appear to be a discontinuity of mortality around the very low birth weight threshold? How does the number of observations in each bin affect your mean estimates?

Answer

```
#create data bins
data$bins <- ((data$bweight-1500)/28.35) %>% floor()

#label bins
data$bins = factor(data$bins,labels=c("1350-1358","1359-1386","1387-1414","1415-1443","1444-1471","1472-

#produce plots
plot1 <- data %>% group_by(bins) %>% summarize(mean = mean(agedth5)) %>% ggplot(aes(bins,mean)) + geom_p

plot2 <- data %>% group_by(bins) %>% summarize(mean = mean(agedth4)) %>% ggplot(aes(bins,mean)) + geom_p
```

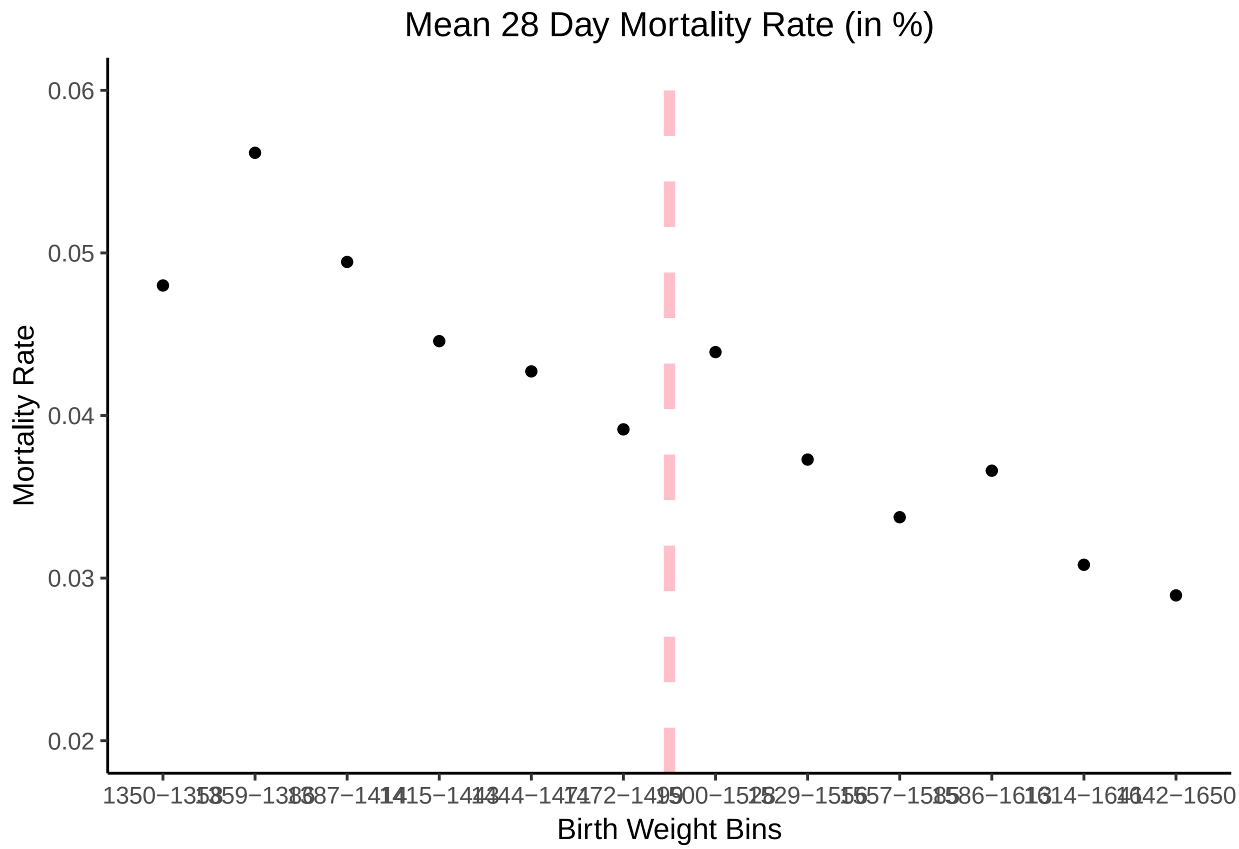
Mean 1 Year Mortality Rate by Birth Weight

```
plot1
```



Mean 28 Day Mortality Rate by Birth Weight

```r
plot2
```



The relationship between birth weight and both of the mortality measures we analyzed in our plots are negatively correlated and highly statistically significant, indicating that higher birth weights are correlated with a minimized probability of mortality, both 28 days out and 1 year out from date of birth.

```r
cor.test(data$agedth4,data$bweight)
```

```
##
##  Pearson's product-moment correlation
##
## data:  data$agedth4 and data$bweight
## t = -22.761, df = 376410, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.04026415 -0.03388369
## sample estimates:
##        cor
## -0.0370743
```

```r
cor.test(data$agedth5,data$bweight)
```

```
##
##  Pearson's product-moment correlation
##
## data:  data$agedth5 and data$bweight
## t = -23.633, df = 376410, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
```

```
## 95 percent confidence interval:
##  -0.04168074 -0.03530097
## sample estimates:
##         cor
## -0.03849125
```

The regression discontinuity at 1500 grams seems to be sharp. It indicates a marked increase in the probability of a premature death at either of the date cutoffs specified previously (28 days or 1 year) above the running variable.

The mean estimate for the lowest bin (1350-1358) is sensitive because the number of observations is ~10% the size of the other bins. Any outliers in either direction are likely to bias the mean estimate of that bin.

```
data %>% count(bins)
```

```
## # A tibble: 12 x 2
##    bins          n
##    <fct>     <int>
##  1 1350-1358  3271
##  2 1359-1386 30236
##  3 1387-1414 29002
##  4 1415-1443 31386
##  5 1444-1471 31653
##  6 1472-1499 32185
##  7 1500-1528 35830
##  8 1529-1556 35344
##  9 1557-1585 35680
## 10 1586-1613 39966
## 11 1614-1641 39649
## 12 1642-1650 32206
```

3

A key assumption for an RDD to provide a causal estimate is that individuals are not able to sort according to the running variable, i.e., they should not be able to manipulate its value. Discuss in your own words whether this is a reasonable assumption in this case. (Include tables with the relevant info (Coefficients of interest, standard errors and sample size).)

Answer

This assumption is reasonable. While there may be some financial incentive to induce early births on behalf of the practicing obstetrician (in order to charge more in future services), there seems likely to be enough practical resistance to inducing early births around the cutoff. Especially when considering that attempting to fall below the cutoff only endangers the life of the baby even more. No parent, under this scenario, would attempt to induce an early birth. The cutoff of 1500 grams is also so low that no parent would reasonably attempt to have a baby born at so low a weight (3.3 lbs). Additionally, any selection across the birth weight variable would have to be reflected in all the other covariates variables around the margin.

I have conducted a balance test between the two bins immediately above and below the 1500 gram threshold. A series of statistically significant differences amongst the two would indicate that there was a likelihood of some manipulation around the boundary of the regression discontinuity.

```
databal <- data
databal <- databal %>% filter(., bins==c("1472-1499") | bins==("1500-1528"))
```

```r
baltest <- lm(data = databal,as.numeric(bins) ~ mom_age + mom_race + mom_ed + gest + nprenatal + sex + a
```

```r
chart1 <- stargazer(baltest, type = "latex", title = "Balance Test Across The Threshold Using a 28.5g Ca
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Oct 30, 2018 - 01:31:20

Table 1: Balance Test Across The Threshold Using a 28.5g Caliper

	Dependent variable:
Mother's Age	0.002968*
	(0.001526)
Mother's Race	0.000133
	(0.016792)
Mother's Education	−0.004343
	(0.003867)
Gestational Weeks	0.001067
	(0.002376)
Prenatal Visits	−0.000849
	(0.001731)
Child's Sex	−0.029315
	(0.018679)
Infant Age of Death	−0.000399
	(0.006728)
Constant	6.521725***
	(0.098360)
Observations	2,950
R^2	0.002234
Adjusted R^2	−0.000140
Residual Std. Error	0.499188 (df = 2942)
F Statistic	0.941185 (df = 7; 2942)
Note:	*p<0.1; **p<0.05; ***p<0.01

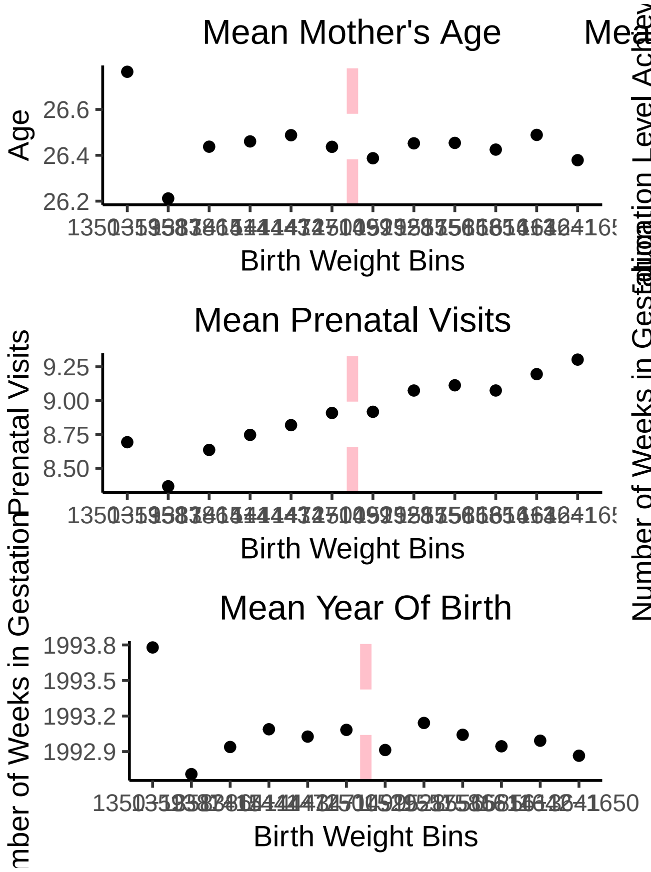
The results indicate that there is only a statistically significant difference in the ages of the mothers. We can reasonably assume that individuals are not selecting across the running variable at the threshold.

4

Assess informally whether the behavior of other covariates is smooth around the threshold, by plotting the mean of some covariates (mother's age, mother's education less than high school, gestational age, prenatal care visits, and year of birth) against birth weight as you did in point (2). Is there any evidence of discontinuities on other covariates around the very low birth weight threshold? If they were, how could these affect your RDD estimates?

Answer

```
plot3 <- data %>% group_by(bins) %>% summarize(mean = mean(mom_age)) %>% ggplot(aes(bins,mean)) + geom_
plot4 <- data %>% group_by(bins) %>% summarize(mean = mean(mom_ed1, na.rm = TRUE)) %>% ggplot(aes(bins,m
plot5 <- data %>% group_by(bins) %>% summarize(mean = mean(nprenatal, na.rm = TRUE)) %>% ggplot(aes(bin
plot6 <- data %>% group_by(bins) %>% summarize(mean = mean(gest, na.rm = TRUE)) %>% ggplot(aes(bins,mea
plot7 <- data %>% group_by(bins) %>% summarize(mean = mean(yob, na.rm = TRUE)) %>% ggplot(aes(bins,mean
```



There is evidence of discontinuities on other covariates (particularly number of prenatal visits). The problem this could likely introduce in our RDD estimates is that the supposition that our populations are similar around the threshold will not hold and thereby undermining the assumption necessary for us to estimate the RDD accurately.

5

Now get an estimate of the size of the discontinuity in one–year and 28–day mortality, around the 1500 grams threshold using a caliper of 85 grams (above and below the threshold). To do so, use the following model:

$$Y_i = \alpha_0 + \alpha_1 VLBW_i + \alpha_2 VLBW_i * (g_i - 1500) + \alpha_3 * (1 - VLBW_i) * (g_i - 1500) + \epsilon_i$$

where Y_i is the outcome of interest, $VLBW_i$ indicates that a newborn had very low birth weight (<1500 grams), g_i is birth weight, and ϵ_i a disturbance term. Interpret the coefficients α_1, α_2, and α_3.

Answer

```
data$VLBW <- ifelse(data$bweight >= 1500, 0,1)
data$gidiff <- data$bweight - 1500
data$VLBWInter <- data$VLBW*data$gidiff
data$NegVLBWInter <- (1-data$VLBW)*data$gidiff

reg1 <- data %>% filter(., (bweight >= (1500-85)) & (bweight <= (1500+85))) %>% lm(agedth4~VLBW + VLBWIn

reg2 <- data %>% filter(., (bweight >= (1500-85)) & (bweight <= (1500+85))) %>% lm(agedth5~VLBW + VLBWIn
```

```
stargazer(reg1,reg2, type = "latex", title = "Infant Mortality by Very Low Birth Weight Status (85g Cal
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Oct 30, 2018 - 01:31:23

Table 2: Infant Mortality by Very Low Birth Weight Status (85g Caliper)

	Dependent variable:	
	28 Day Mortality	1 Year Mortality
	(1)	(2)
Birth Weight < 1,500g	−0.008781*** (0.001811)	−0.009510*** (0.002153)
Birth Weight < 1500g x grams from cutoff	−0.000113*** (0.000027)	−0.000135*** (0.000032)
Birth Weight >= 1500g x grams from cutoff	−0.000200*** (0.000024)	−0.000225*** (0.000029)
Constant	0.045244*** (0.001037)	0.063156*** (0.001233)
Observations	202,078	202,078
R^2	0.000512	0.000502
Adjusted R^2	0.000497	0.000487
Residual Std. Error (df = 202074)	0.196176	0.233195
F Statistic (df = 3; 202074)	34.512520***	33.816930***
Note:		*p<0.1; **p<0.05; ***p<0.01

In a sharp RDD (Table 2), α_1 represents the local average treatment effect, in particular indicating the amount of impact treatment has on the mortality rate. Crossing above the 1500 gram threshold has the impact of increasing the likelihood of infant mortality at both 28 days and 1 year time horizons. A likely reason for this impact is that as a baby crosses that threshold, they receive less intensive medical attention but are still at similar risks compared to babies just below that threshold, thereby increasing the risk of death for children born above the cutoff weight.

Subsequently in α_2, the rate of mortality decreases the further away you are from the cutoff rate, as you receive more intensive medical care given how low the birth weight is even compared to the range directly around the cutoff. α_3 shows that as you trend further above the cutoff, you are also likely to experience a lower mortality rate, as you'd expect those children, who are further from the cutoff to be healthier than the children who are cluster around the cutoff weight.

6

Now add covariates to the model in (5). Include mother's age, indicators for mother's education and race, indicators for year of birth, indicators for gestational age and prenatal care visits. Use the dummies provided in the data for gestational age and prenatal care visits. Compare your estimates to those obtained in (5) and explain the difference if any.

Answer

```
data$mom_age_bins <- cut(data$mom_age, breaks = seq(10,55, by = 5), right = FALSE, labels = c(0:8))

reg3 <- data %>% filter(., (bweight >= (1500-85)) & (bweight <= (1500+85))) %>% lm(agedth4~VLBW + VLBWIr

reg4 <- data %>% filter(., (bweight >= (1500-85)) & (bweight <= (1500+85))) %>% lm(agedth5~VLBW + VLBWIr
```

In comparison to the previous model (Table 3 versus Table 2), the impacts of α_1, α_2, and α_3 are all muted partially, after controlling for the year and demographic characteristics across the threshold, in both the 28 day and 1 year analysis. After controlling, the LATE or α_1 is -0.007424 (28 day, controlled) versus -0.008781 (28 day general). Similarly, the α_2 is -0.000103 (28 day, controlled) versus -0.000113 (28 day general). α_3 exhibits a similar relationship where the controlled estimate (-0.000183) is smaller in intensity than the general estimate (-0.000200). The same tendency exists in the 1 year mortality rates across all three variables, given the controls.

This relationship is valid as as you'd expect to not have statistically significant difference across the populations as evidenced earlier by our balance tests. So controlling for demographic characteristics and underlying variables doesn't have a significant impact on the value of our estimates.

7

Use the model in (6) to assess the sensitivity of the estimates to the use of different calipers. Use calipers of 30 and 120 grams (above and below the 1500 threshold). Are the estimates any different to those obtained in (6)? What is the tradeoff that we face when increasing/decreasing the caliper?

Answer

As we begin to adjust the bandwidth of our analysis (Table 3 & Table 4), we see the very obvious tradeoffs associated with the decisions to tighten or expand the bandwidth. As we shrink the bandwidth from 85g to 30g, we see estimates with larger magnitudes (28 day - 85g & 30g: $\alpha_1 = -0.007424$ vs $\alpha_1 = -0.014398$, $\alpha_2 = -0.000103$ vs $\alpha_2 = -0.000278$, $\alpha_3 = -0.000183$ vs $\alpha_3 = -0.000552$) but standard errors increase across the board (28 day - 85g & 30g: $\alpha_1 = 0.001805$ vs $\alpha_1 = 0.004375$, $\alpha_2 = 0.000027$ vs $\alpha_2 = 0.000173$, $\alpha_3 = 0.000024$ vs $\alpha_3 = 0.000112$). This relationship holds across the 1 year mortality rate. As we tighten the bandwidth, we find our estimates to be less precise (the increase in standard errors) but we have less bias towards the outliers and get much closer to what the likely impact of treatment is in this case.

The inverse of this relationship is true when we expand the bandwidth (Table 3 & Table 5). We get more precise, indicated by the standard errors decreasing across the board. However, our estimates are more biased. We stray further from the likely impact of the treatment as our LATE estimator is now smaller in magnitude in both the 28 day case (28 day - 30g, 85g, 120g:$\alpha_1 = -0.014398$,$\alpha_1 = -0.007424$, $\alpha_1 = -0.005648$) and the 1 year case (1 year - 30g, 85g, 120g:$\alpha_1 = -0.014230$,$\alpha_1 = -0.007637$, $\alpha_1 = -0.006444$) compared to our 85g and 30g estimations.

```
reg5 <- data %>% filter(., (bweight >= (1500-30)) & (bweight <= (1500+30))) %>% lm(agedth4~VLBW + VLBWIr

reg6 <- data %>% filter(., (bweight >= (1500-30)) & (bweight <= (1500+30))) %>% lm(agedth5~VLBW + VLBWIr

reg7 <- data %>% filter(., (bweight >= (1500-120)) & (bweight <= (1500+120))) %>% lm(agedth4~VLBW + VLBW

reg8 <- data %>% filter(., (bweight >= (1500-120)) & (bweight <= (1500+120))) %>% lm(agedth5~VLBW + VLBW
```

8

Synthetize your findings and discuss what kind of supplementary information would you need to make a cost-benefit analysis of treatment received by newborns close to the very low birth weight threshold.

Answer

We are aware that there is a likely impact of additional treatment for children born under the threshold that minimizes the probability of infant mortality. Given that, we can expect that a broader plan of medical services for infants will be effective. In order to conduct a cost-benefit analysis of said strategy, valuable metrics to have would be the spending on medical services for both the control and the treatment group and the time spent under medical care for control and treatment groups. As we identify how costly the intervention would be both in terms of spending and time, we can prepare an analysis of how effective this strategy could be from the perspective of a cost-benefit analysis.

Appendix

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Oct 30, 2018 - 01:31:34

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Oct 30, 2018 - 01:31:35

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Oct 30, 2018 - 01:31:36

Table 3: Infant Mortality by Very Low Birth Weight Status (85g Caliper)

	Dependent variable:	
	28 Day Mortality	1 Year Mortality
	(1)	(2)
Birth Weight < 1,500g	−0.007424*** (0.001805)	−0.007637*** (0.002144)
Birth Weight < 1500g x grams from cutoff	−0.000103*** (0.000027)	−0.000122*** (0.000032)
Birth Weight >= 1500g x grams from cutoff	−0.000183*** (0.000024)	−0.000200*** (0.000029)
as.factor(mom_age_bins)1	−0.007756 (0.005956)	−0.012202* (0.007076)
as.factor(mom_age_bins)2	−0.002516 (0.005954)	−0.007897 (0.007073)
as.factor(mom_age_bins)3	−0.005058 (0.005973)	−0.013271* (0.007096)
as.factor(mom_age_bins)4	−0.007207 (0.006000)	−0.015287** (0.007128)
as.factor(mom_age_bins)5	−0.006640 (0.006078)	−0.015192** (0.007220)
as.factor(mom_age_bins)6	0.001438 (0.006604)	−0.002356 (0.007846)
as.factor(mom_age_bins)7	0.011181 (0.011785)	0.008213 (0.014001)
as.factor(mom_age_bins)8	−0.033071 (0.032681)	−0.048574 (0.038826)
mom_ed1	−0.011510*** (0.001886)	−0.009027*** (0.002241)
mom_ed2	−0.008722*** (0.001806)	−0.011897*** (0.002146)
mom_ed3	−0.012735*** (0.001977)	−0.017867*** (0.002349)
mom_ed4	−0.011467*** (0.002072)	−0.018365*** (0.002461)
mom_ed5		
as.factor(mom_race)2	−0.020106*** (0.000978)	−0.016880*** (0.001162)
as.factor(mom_race)3	−0.003626 (0.002226)	−0.000871 (0.002644)
as.factor(yob)1984	−0.005147* (0.002839)	−0.008100** (0.003373)
as.factor(yob)1985	−0.009244*** (0.002766)	−0.009919*** (0.003286)
as.factor(yob)1986	−0.012888*** (0.002762)	−0.016420*** (0.003282)
as.factor(yob)1987	−0.016306*** (0.002748)	−0.021337*** (0.003265)
as.factor(yob)1988	−0.014686*** (0.002725)	−0.017621*** (0.003237)
as.factor(yob)1989	−0.015468*** (0.002697)	−0.020413*** (0.003204)
as.factor(yob)1990	−0.021852*** (0.002687)	−0.029073*** (0.003192)
as.factor(yob)1991	−0.024371*** (0.002689)	−0.030340*** (0.003195)
as.factor(yob)1995	−0.033311*** (0.002699)	−0.043087*** (0.003206)
as.factor(yob)1996	−0.033893*** (0.002698)	−0.043759*** (0.003206)
as.factor(yob)1997	−0.035376*** (0.002685)	−0.045407*** (0.003190)
as.factor(yob)1998	−0.034734*** (0.002672)	−0.046312*** (0.003175)
as.factor(yob)1999	−0.032986*** (0.002657)	−0.045196*** (0.003157)
as.factor(yob)2000	−0.036089*** (0.002651)	−0.049076*** (0.003150)
as.factor(yob)2001	−0.035809*** (0.002658)	−0.046734*** (0.003157)
as.factor(yob)2002	−0.037207*** (0.002647)	−0.048649*** (0.003144)
gest_wks1	−0.007298*** (0.001611)	−0.009418*** (0.001914)
gest_wks2	0.012823*** (0.002241)	0.017995*** (0.002663)
gest_wks3	−0.002238 (0.006037)	−0.000905 (0.007172)
gest_wks4		
nprenatal_1	−0.001339 (0.001645)	−0.001665 (0.001954)
nprenatal_2	−0.008279*** (0.001710)	−0.011007*** (0.002031)
nprenatal_3	−0.013365*** (0.002010)	−0.015327*** (0.002388)
nprenatal_4		
Constant	0.100441*** (0.006622)	0.136327*** (0.007867)
Observations	202,078	202,078
R^2	0.009191	0.010333
Adjusted R^2	0.008999	0.010142
Residual Std. Error (df = 202038)	0.195340	0.232066
F Statistic (df = 39; 202038)	48.053810***	54.090630***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 4: Infant Mortality by Very Low Birth Weight Status (30g Caliper)

	Dependent variable:	
	28 Day Mortality	1 Year Mortality
	(1)	(2)
Birth Weight < 1,500g	−0.014398*** (0.004375)	−0.014230*** (0.005174)
Birth Weight < 1500g x grams from cutoff	−0.000278 (0.000173)	−0.000252 (0.000205)
Birth Weight >= 1500g x grams from cutoff	−0.000552*** (0.000112)	−0.000641*** (0.000132)
as.factor(mom_age_bins)1	−0.019095* (0.010110)	−0.022012* (0.011956)
as.factor(mom_age_bins)2	−0.014124 (0.010106)	−0.017855 (0.011952)
as.factor(mom_age_bins)3	−0.019487* (0.010139)	−0.026050** (0.011990)
as.factor(mom_age_bins)4	−0.017770* (0.010182)	−0.022516* (0.012041)
as.factor(mom_age_bins)5	−0.018763* (0.010314)	−0.024645** (0.012197)
as.factor(mom_age_bins)6	−0.014740 (0.011169)	−0.018522 (0.013209)
as.factor(mom_age_bins)7	−0.015220 (0.020109)	−0.017205 (0.023781)
as.factor(mom_age_bins)8	−0.047315 (0.053949)	−0.061365 (0.063800)
mom_ed1	−0.011046*** (0.003178)	−0.006436* (0.003759)
mom_ed2	−0.009734*** (0.003050)	−0.013381*** (0.003607)
mom_ed3	−0.011383*** (0.003346)	−0.016387*** (0.003957)
mom_ed4	−0.012444*** (0.003499)	−0.020341*** (0.004138)
mom_ed5		
as.factor(mom_race)2	−0.019268*** (0.001649)	−0.017599*** (0.001950)
as.factor(mom_race)3	−0.001191 (0.003702)	0.001198 (0.004378)
as.factor(yob)1984	−0.000801 (0.004780)	−0.002219 (0.005653)
as.factor(yob)1985	−0.006156 (0.004670)	−0.005825 (0.005523)
as.factor(yob)1986	−0.019763*** (0.004682)	−0.024319*** (0.005537)
as.factor(yob)1987	−0.023530*** (0.004636)	−0.027484*** (0.005483)
as.factor(yob)1988	−0.016244*** (0.004615)	−0.017707*** (0.005458)
as.factor(yob)1989	−0.019821*** (0.004551)	−0.021897*** (0.005382)
as.factor(yob)1990	−0.019766*** (0.004526)	−0.026766*** (0.005352)
as.factor(yob)1991	−0.025827*** (0.004570)	−0.030576*** (0.005404)
as.factor(yob)1995	−0.033033*** (0.004574)	−0.043495*** (0.005409)
as.factor(yob)1996	−0.032346*** (0.004578)	−0.039088*** (0.005414)
as.factor(yob)1997	−0.033846*** (0.004528)	−0.043414*** (0.005355)
as.factor(yob)1998	−0.035202*** (0.004517)	−0.044658*** (0.005342)
as.factor(yob)1999	−0.033369*** (0.004505)	−0.044519*** (0.005328)
as.factor(yob)2000	−0.036095*** (0.004484)	−0.048491*** (0.005302)
as.factor(yob)2001	−0.035013*** (0.004493)	−0.044510*** (0.005313)
as.factor(yob)2002	−0.035571*** (0.004475)	−0.043834*** (0.005292)
gest_wks1	−0.011844*** (0.002736)	−0.015179*** (0.003236)
gest_wks2	0.005768 (0.003812)	0.011204** (0.004508)
gest_wks3	−0.009220 (0.010122)	−0.011989 (0.011970)
gest_wks4		
nprenatal_1	−0.004548 (0.002773)	−0.007298** (0.003279)
nprenatal_2	−0.010922*** (0.002888)	−0.015423*** (0.003416)
nprenatal_3	−0.018796*** (0.003392)	−0.022131*** (0.004011)
nprenatal_4		
Constant	0.122940*** (0.011151)	0.159286*** (0.013187)
Observations	72,941	72,941
R^2	0.009081	0.010601
Adjusted R^2	0.008551	0.010072
Residual Std. Error (df = 72901)	0.198117	0.234296
F Statistic (df = 39; 72901)	17.130030***	20.029180***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 5: Infant Mortality by Very Low Birth Weight Status (120g Caliper)

	Dependent variable:	
	28 Day Mortality	1 Year Mortality
	(1)	(2)
Birth Weight < 1,500g	−0.005648*** (0.001483)	−0.006444*** (0.001761)
Birth Weight < 1500g x grams from cutoff	−0.000113*** (0.000017)	−0.000141*** (0.000020)
Birth Weight >= 1500g x grams from cutoff	−0.000097*** (0.000012)	−0.000120*** (0.000014)
as.factor(mom_age_bins)1	−0.005293 (0.004896)	−0.012082** (0.005813)
as.factor(mom_age_bins)2	−0.000632 (0.004894)	−0.008019 (0.005810)
as.factor(mom_age_bins)3	−0.002448 (0.004909)	−0.012780** (0.005828)
as.factor(mom_age_bins)4	−0.004060 (0.004931)	−0.015344*** (0.005855)
as.factor(mom_age_bins)5	−0.003658 (0.004994)	−0.014533** (0.005929)
as.factor(mom_age_bins)6	0.003870 (0.005415)	−0.001730 (0.006430)
as.factor(mom_age_bins)7	0.010419 (0.009511)	0.003137 (0.011292)
as.factor(mom_age_bins)8	−0.031201 (0.027428)	−0.048974 (0.032566)
mom_ed1	−0.010522*** (0.001526)	−0.007324*** (0.001812)
mom_ed2	−0.007866*** (0.001460)	−0.010203*** (0.001734)
mom_ed3	−0.010751*** (0.001600)	−0.014865*** (0.001900)
mom_ed4	−0.010383*** (0.001676)	−0.016339*** (0.001990)
mom_ed5		
as.factor(mom_race)2	−0.019107*** (0.000793)	−0.017054*** (0.000942)
as.factor(mom_race)3	−0.003880** (0.001813)	−0.002070 (0.002153)
as.factor(yob)1984	−0.003302 (0.002286)	−0.005025* (0.002715)
as.factor(yob)1985	−0.009889*** (0.002225)	−0.010551*** (0.002641)
as.factor(yob)1986	−0.013227*** (0.002221)	−0.016935*** (0.002637)
as.factor(yob)1987	−0.018498*** (0.002215)	−0.022822*** (0.002630)
as.factor(yob)1988	−0.016806*** (0.002192)	−0.020187*** (0.002602)
as.factor(yob)1989	−0.015097*** (0.002177)	−0.019520*** (0.002584)
as.factor(yob)1990	−0.021840*** (0.002166)	−0.027977*** (0.002572)
as.factor(yob)1991	−0.024683*** (0.002171)	−0.030817*** (0.002578)
as.factor(yob)1995	−0.033704*** (0.002184)	−0.042550*** (0.002593)
as.factor(yob)1996	−0.034846*** (0.002184)	−0.045610*** (0.002593)
as.factor(yob)1997	−0.036242*** (0.002175)	−0.046186*** (0.002583)
as.factor(yob)1998	−0.035247*** (0.002158)	−0.046645*** (0.002563)
as.factor(yob)1999	−0.035506*** (0.002147)	−0.047449*** (0.002550)
as.factor(yob)2000	−0.035802*** (0.002141)	−0.048298*** (0.002542)
as.factor(yob)2001	−0.036493*** (0.002146)	−0.047470*** (0.002548)
as.factor(yob)2002	−0.037754*** (0.002140)	−0.049461*** (0.002541)
gest_wks1	−0.007671*** (0.001317)	−0.009676*** (0.001564)
gest_wks2	0.011903*** (0.001808)	0.017752*** (0.002146)
gest_wks3	−0.001676 (0.004867)	0.001936 (0.005778)
gest_wks4		
nprenatal_1	−0.002878** (0.001340)	−0.002806* (0.001591)
nprenatal_2	−0.009149*** (0.001392)	−0.011398*** (0.001653)
nprenatal_3	−0.014003*** (0.001635)	−0.016073*** (0.001942)
nprenatal_4		
Constant	0.096514*** (0.005428)	0.133822*** (0.006445)
Observations	304,636	304,636
R^2	0.009209	0.010727
Adjusted R^2	0.009082	0.010600
Residual Std. Error (df = 304596)	0.194519	0.230958
F Statistic (df = 39; 304596)	72.591450***	84.686390***
Note:		*p<0.1; **p<0.05; ***p<0.01