Homework 3 - GVPT 729A

Answer the following questions. Include your code, and report all the results you used to answer the questions. https://raw.githubusercontent.com/Neilblund/729A/master/data/voterid.csv

The link above contains data from "Hicks et al. 2015: A Principle or a Strategy? Voter Identification Laws and Partisan Competition in the American States"

- *photo* is equal to 1 if a state has legislation that requires voters to show photo ID at the polling booth, and 0 of they do not have this requirement.
- *fraud* is the average number of voter fraud cases prosecuted in a given state since 2001.
- *election_margin* is the average partisan vote margin (%Republican-%Democratic) in a state since 2001.
- *gopleg* is the average % of a state's legislature that is Republican.

Questions

1. Use OLS to estimate a linear probability model using photo as the dependent variable, and *fraud*, *election_margin* and *gopleg* as independent variables. Obtain predicted probabilities that *photo* = 1 under two or more hypothetical scenarios. Discuss your results.
2. Run a logistic regression model using photo as the dependent variable, and *fraud*, *election_margin* and *gopleg* as independent variables. Obtain predicted probabilities that *photo* = 1 under two or more hypothetical scenarios. Discuss your results, and compare with your results from model 1.

notes

1. There's no "right" or "wrong" scenario here, but you should think about comparing scenarios that are plausible reflections of the real world.
2. We used the code below to get predicted probabilities from the probit model. You will need to modify this slightly to get predictions from a logit, just remember that:

$$Pr(Y = 1|X1, X2) = \frac{1}{1 + e^{-(B_0 + B_1 X1 + B_2 X2)}}$$

.

```r
  n <- 1000
  x1 <- rnorm(n)
  x2 <- rnorm(n)
  X <- cbind(1, x1, x2)
  b <- c(1, -.5, .5)
  p <- pnorm(X%*%b)
  y <- rbinom(n, 1, p)
  model <- glm(y ~ x1 + x2, family = "binomial"(link="probit"))

#with x1 and x2 set at their mean
meanx1<-mean(x1)
meanx2<-mean(x2)
pnorm(model$coefficients[1]+
    meanx1*model$coefficients[2]+
    meanx2*model$coefficients[3])

## (Intercept)
##   0.8393224

#with x1 at its 25%th percentile and x2 at its mean

q.x1<-quantile(x1, .25)

pnorm(model$coefficients[1]+
    q.x1*model$coefficients[2]+
    meanx2*model$coefficients[3])

## (Intercept)
##   0.8945215
```