Homework 2 Solution Script

Ling Zhu and Songeun Emily Lee

10/07/2025

0. Loading Packages

```r
rm(list=ls()) # Start clean: remove all objects from the workspace
## Libraries
library(tinytex); library(ggplot2); library(ggpubr); library(stargazer)
library(haven);   library(faraway); library(MASS);    library(aod)
library(lmtest);  library(dplyr);   library(AER);     library(glmx)
library(heatmapFit); library(maxLik); library(reshape2); library(foreign)

## Working directory
setwd("/Users/songeunlee/Desktop/HW2")
#setwd("/Users/lingzhu/Dropbox/UH Teaching/POLS6382_2025/HW Assignments/2025 HW Review and Solution Scr

## Data
court <- read.dta("Court.dta")                    # load Stata file
court$liberal2 <- round(court$liberal * 100, 2) # 0-1 -> 0-100 (%)
head(court) # quick sanity check
```

```
  propetit ineffcou tcterror multpet usparty   liberal liberal2
1        1        0        0       0       1 0.6938889    69.39
2        1        0        1       0       0 0.6938889    69.39
3        1        0        0       0       0 0.6712500    67.12
4        1        0        0       0       0 0.7475000    74.75
5        1        0        1       0       1 0.7383333    73.83
6        0        0        0       1       0 0.7475000    74.75
```

```r
attach(court)
```

1. Analyzing Binary Outcomes

Question 1a

```r
logitmod <- glm(propetit ~ liberal2+ineffcou+tcterror+multpet+usparty, family=binomial(link=logit), cou

stargazer(logitmod, type="text")
```

```
==============================================
                  Dependent variable:
              ----------------------------
                        propetit
```

```
------------------------------------------------
liberal2                        0.079***
                                (0.018)

ineffcou                        1.443*
                                (0.827)

tcterror                       -0.048
                                (0.517)

multpet                        -1.414*
                                (0.833)

usparty                        -0.777
                                (0.784)

Constant                       -3.632***
                                (0.836)

------------------------------------------------
Observations                       109
Log Likelihood                  -58.127
Akaike Inf. Crit.               128.254
================================================
Note:               *p<0.1; **p<0.05; ***p<0.01
```

Question 1b

```
# Logistic regression coefficients are in log-odds, which are hard to interpret directly.
# By exponentiating them with exp(), we convert log-odds into odds ratios (ORs).
# Odds ratios are more intuitive:
#   OR > 1 → predictor increases the odds of the outcome
#   OR < 1 → predictor decreases the odds of the outcome
#   OR = 1 → no effect
exp(logitmod$coefficients)
```

```
(Intercept)    liberal2    ineffcou    tcterror     multpet     usparty
 0.02646956  1.08271316  4.23525563  0.95358253  0.24326808  0.45991728
```

Answer 1b

liberal2 (Court ideology):
A one-point increase in the Court's liberalism score raises the odds of granting relief by about 8% (odds ratio = 1.083), a substantively large and statistically significant effect.

ineffcou (ineffective counsel):
Petitions alleging ineffective counsel have about 4.2 times higher odds of relief, a substantively important and statistically significant effect.

multpet (multiple petitions):
Repeat petitions reduce the odds of relief by about 76% (odds ratio = 0.24), a substantively large negative and statistically significant effect.

usparty (U.S. is a party):
Facing the U.S. lowers the odds of relief by about 54% (odds ratio = 0.46), but this effect is not statistically significant.

tcterror (terrorism case):
Terrorism-related petitions have essentially no effect on the odds of relief (odds ratio = 0.95), and the result is not statistically significant.

Question 1c

```r
# Calculate median values for control variables
median_ineffcou <- median(court$ineffcou)
median_tcterror <- median(court$tcterror)
median_multpet <- median(court$multpet)
median_usparty <- median(court$usparty)

# Create a sequence of values for liberal2
newcourt <- data.frame(liberal2 = seq(from = 0, to = 100, length.out = 109))

# Include median values for control variables
newcourt <- cbind(newcourt, ineffcou = median_ineffcou, tcterror = median_tcterror,
multpet = median_multpet, usparty = median_usparty)

# Use predict with newcourt to calculate predicted probabilities and confidence intervals
newcourt <- cbind(newcourt, predict(logitmod, newdata = newcourt,
se = TRUE, type = "response"))

# Calculate lower and upper bounds for confidence intervals
newcourt <- within(newcourt, {
LL <- fit - (1.96 * se.fit)
UL <- fit + (1.96 * se.fit)
PredictedProb <- fit
})

# Plot predicted probabilities with CIs
ggplot(newcourt, aes(x = liberal2, y = PredictedProb, colour = "Mean Prediction")) +
geom_line() + geom_errorbar(aes(ymin = LL, ymax = UL, colour = "gray"),
colour = "gray", size = 0.3, width = 0.01) + labs(x = "Court Ideology (Liberalism)",
y = "Predicted Probability of Granting Relief", colour = "Predicted Probabilities") +
ggtitle("The Effect of Preferences on Granting Relief") +
theme_bw() + theme(legend.justification = c(0.7, 1), legend.position = c(0.9,
0.3))
```
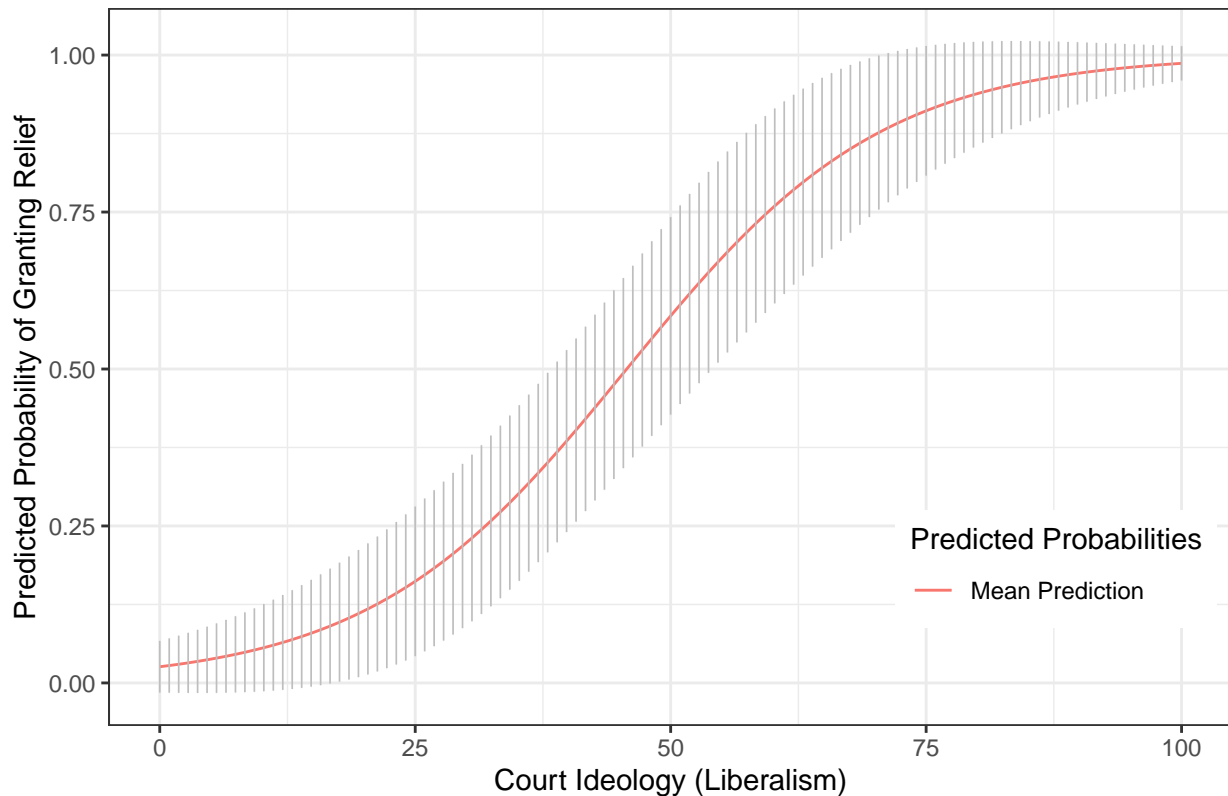
The Effect of Preferences on Granting Relief

Answer 1c

- The graph shows that as the Court's ideology score becomes more liberal, the probability of granting relief increases for the petitioner. At the most conservative end (0), the predicted probability is close to zero, while at the most liberal end (100) it rises to about 0.8–0.9. The confidence intervals are narrower at the extremes and wider around the middle, reflecting greater uncertainty near 0.5 and more certainty when probabilities approach 0 or 1. Overall, the results strongly support the conclusion that court ideology is a central determinant of habeas relief decisions.

Question 1d

```
# logitmod2: ideology + US party + interaction
logitmod2 <- glm(propetit~liberal2+liberal2*usparty, family=binomial(link=logit))

# logitmod3: ideology + legal vars + US party + interaction
logitmod3 <- glm(propetit~liberal2+ineffcou+tcterror+multpet+usparty+liberal2*usparty, family=binomial()

# Compare models
stargazer(logitmod, logitmod2, logitmod3, type="text")
```

```
===============================================
                  Dependent variable:
                 ------------------------------
                           propetit
                  (1)        (2)        (3)
-----------------------------------------------
```

```
liberal2            0.079***  0.064***  0.070***
                    (0.018)   (0.017)   (0.019)


ineffcou            1.443*              1.388*
                    (0.827)             (0.829)


tcterror            -0.048              -0.061
                    (0.517)             (0.522)


multpet             -1.414*             -1.512*
                    (0.833)             (0.844)


usparty             -0.777    -4.721    -4.913
                    (0.784)   (3.798)   (3.912)


liberal2:usparty              0.071     0.076
                              (0.065)   (0.067)


Constant            -3.632*** -2.997*** -3.229***
                    (0.836)   (0.769)   (0.864)


-------------------------------------------------
Observations        109       109       109
Log Likelihood      -58.127   -60.134   -57.242
Akaike Inf. Crit.   128.254   128.267   128.484
=================================================
Note:               *p<0.1; **p<0.05; ***p<0.01
```

```r
# Convert log-odds -> odds ratios
exp(logitmod2$coefficients)
```

```
    (Intercept)          liberal2          usparty liberal2:usparty
    0.049940048       1.066466456      0.008905885      1.074097791
```

```r
exp(logitmod3$coefficients)
```

```
    (Intercept)          liberal2          ineffcou          tcterror
    0.03959809        1.07293799        4.00842945        0.94066225
        multpet           usparty liberal2:usparty
     0.22040491        0.00734764        1.07884031
```

Answer 1d

- The results show that court ideology consistently increases the likelihood of granting relief, but there is
  no evidence that this effect differs in cases where the U.S. is a party. Although petitioners face lower
  baseline odds of success when the U.S. is involved, the interaction term between ideology and U.S. party
  status is positive but far from statistically significant, indicating that the slope of ideology does not
  differ across groups. Overall, the findings suggest that liberal courts are more favorable to petitioners in
  general, and while cases against the U.S. appear harder to win, there is no statistically reliable evidence
  that the impact of ideology itself varies by whether the U.S. is a litigant.

```r
# Fit models (with interaction)
logitmod3 <- glm(propetit ~ liberal2 + ineffcou + tcterror + multpet + usparty + liberal2:usparty,
                 data = court, family = binomial(link = "logit"))

# Medians for controls
```

```r
meds <- with(court, list(ineffcou = median(ineffcou, na.rm=TRUE),
                         tcterror = median(tcterror, na.rm=TRUE),
                         multpet  = median(multpet,  na.rm=TRUE)))

# Newdata grid: liberal2 (0-100) × usparty (0/1), controls at medians
grid <- expand.grid(liberal2 = seq(0, 100, length.out = 201),
                    usparty  = c(0, 1))
grid$ineffcou <- meds$ineffcou
grid$tcterror <- meds$tcterror
grid$multpet  <- meds$multpet

# Predict on response scale with SE
# Note: Using type = "response" directly gives predicted probabilities (0-1 scale)
# along with standard errors on the probability scale.
# This avoids the problem of asymmetric CIs that occurs when predicting on the
# link (logit) scale and then transforming with plogis().
pr <- predict(logitmod3, newdata = grid, type = "response", se.fit = TRUE)

# Point probs + 95% CI on response scale (clip to [0,1] just in case)
grid$prob <- pr$fit
grid$lo   <- pmax(0, pr$fit - 1.96 * pr$se.fit)
grid$hi   <- pmin(1, pr$fit + 1.96 * pr$se.fit)

# Factor labels
grid$usparty_f <- factor(grid$usparty, labels = c("No", "Yes"))

# Plot
library(ggplot2)
ggplot(grid, aes(liberal2, prob, color = usparty_f, fill = usparty_f)) +
  geom_ribbon(aes(ymin = lo, ymax = hi), alpha = 0.18, color = NA) +
  geom_line(size = 1) +
  labs(title = "1d) Interaction Effect: Liberalism × U.S. Party",
       x = "Court Ideology (0-100 Liberalism)",
       y = "Predicted Probability of Granting Relief",
       color = "U.S. Party Involved", fill = "U.S. Party Involved") +
  theme_bw()
```
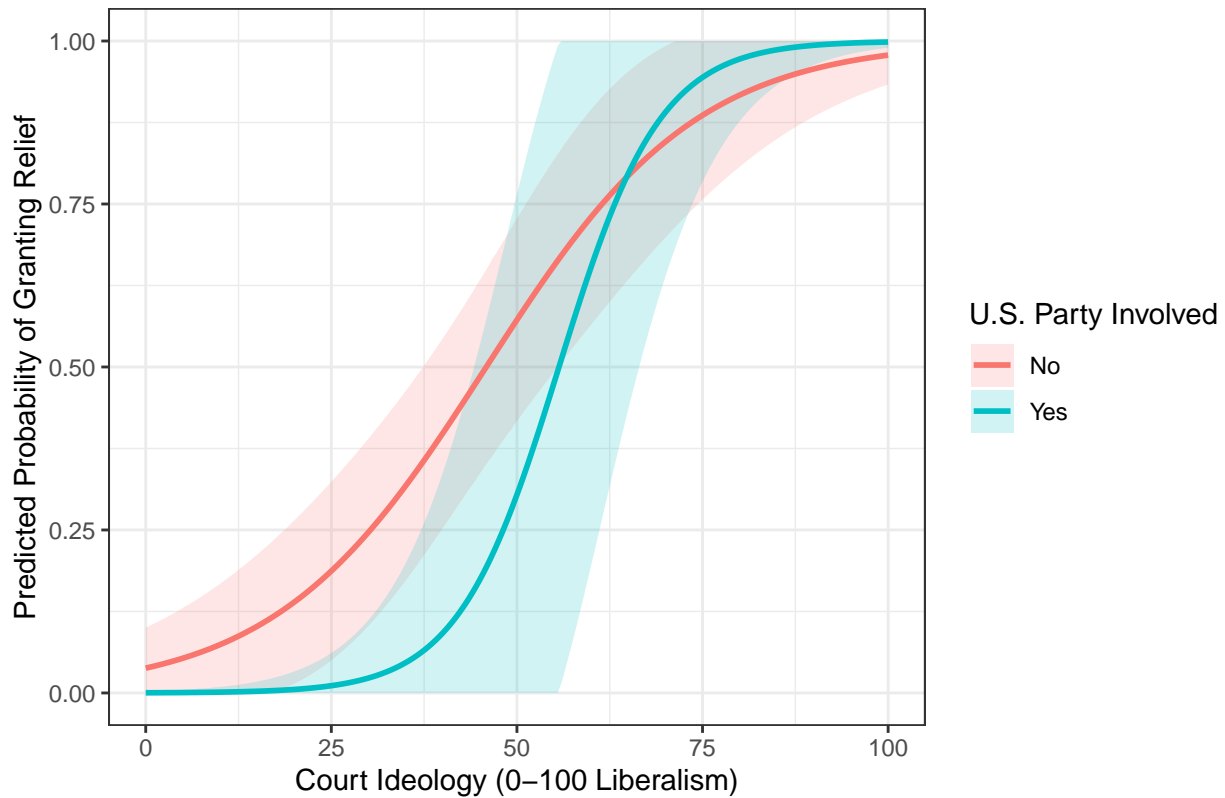
1d) Interaction Effect: Liberalism × U.S. Party

At the conservative end of the scale, the blue line (U.S. party involved) starts below the red line, suggesting that the probability of granting relief is somewhat lower when the U.S. is a party. However, as judicial ideology becomes more liberal, the two curves converge, and both approach nearly identical probabilities of granting relief at the liberal end of the scale. The shaded areas represent 95% confidence intervals, which are relatively wide—especially for the blue curve at mid-range ideology—and overlap substantially across most of the range. This overlap indicates that the observed differences are not statistically significant, reinforcing that judicial ideology, rather than U.S. party involvement, is the main driver of the Court's decisions.

2. Analyzing Binary Outcomes with Unconstant Varience

Question 2a

```
senatedata<-read_dta("SenateVote.dta")
attach(senatedata)

#Economic
hetmodel1 <- hetglm(esvote ~ buspacs + labpacs | buspacs + labpacs, family=binomial(link="probit"),data=

#Striker
hetmodel2 <- hetglm(srvote ~ buspacs + labpacs | buspacs + labpacs, family=binomial(link="probit"),data=
```

Answer 2a

- The heteroscedastic probit results show that business PAC contributions reduce the likelihood of senators supporting both the economic stimulus and striker replacement bills, while labor contributions increase it. In addition to shifting the mean probability of support, business contributions significantly raise the predictability of voting by increasing variance effects, indicating that senators receiving more business money are both less likely and more consistently opposed. Labor contributions, by contrast, strengthen support for labor-backed bills but have weaker and statistically insignificant effects on variance. Overall, these findings support the claim that campaign contributions affect not only the direction of votes but also their certainty.

Question 2b

```
summary(hetmodel1)
```

```
Call:
hetglm(formula = esvote ~ buspacs + labpacs | buspacs + labpacs, data = senatedata,
    family = binomial(link = "probit"))

Deviance residuals:
    Min      1Q  Median      3Q     Max
-1.7384 -0.1025  0.0000  0.0232  2.0125

Coefficients (binomial model with probit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   1.2885     0.9351   1.378   0.1682
buspacs      -0.7293     0.4118  -1.771   0.0766 .
labpacs       5.8501     3.4907   1.676   0.0938 .

Latent scale model coefficients (with log link):
        Estimate Std. Error z value Pr(>|z|)
buspacs  0.05669    0.01728   3.280  0.00104 **
labpacs  0.06443    0.05328   1.209  0.22659
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-likelihood: -14.6 on 5 Df
LR test for homoscedasticity: 10.26 on 2 Df, p-value: 0.00593
Dispersion: 1
Number of iterations in nlminb optimization: 22
```

```
summary(hetmodel2)
```

```
Call:
hetglm(formula = srvote ~ buspacs + labpacs | buspacs + labpacs, data = senatedata,
    family = binomial(link = "probit"))

Deviance residuals:
    Min      1Q  Median      3Q     Max
-1.8053 -0.0153  0.0000  0.0003  1.4399


Coefficients (binomial model with probit link):
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.6367     0.7677   0.829    0.407
buspacs      -1.3246     1.1238  -1.179    0.239
labpacs      11.9354    10.4389   1.143    0.253


Latent scale model coefficients (with log link):
        Estimate Std. Error z value Pr(>|z|)
buspacs  0.06389    0.02321   2.753   0.0059 **
labpacs  0.09717    0.05781   1.681   0.0928 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Log-likelihood: -11.69 on 5 Df
LR test for homoscedasticity: 11.09 on 2 Df, p-value: 0.003898
Dispersion: 1
Number of iterations in nlminb optimization: 28
```

Answer 2b

- For the economic stimulus vote, business PAC contributions significantly reduce the likelihood of support, while labor PAC contributions increase it. In the variance equation, business contributions have a strong positive and significant effect, meaning they make senators' votes more predictable, while labor contributions have no statistically significant impact on variance.

- For the striker replacement vote, the signs are similar: business PACs reduce support and labor PACs increase it, but the coefficients are less precisely estimated in the mean equation. In the variance equation, however, business contributions again significantly increase predictability, while labor contributions have only a marginal effect.

- Overall, these results confirm that business money influences both the direction and certainty of voting, while labor money shapes the direction of voting but has weaker or inconsistent effects on variance.

Question 2c

```
#Standard Probit Economic
probitmod1 <- glm(esvote ~ buspacs + labpacs, family=binomial(link=probit), senatedata)

#Standard Probit Striker
probitmod2 <- glm(srvote ~ buspacs + labpacs, family=binomial(link=probit), senatedata)

#Comparisons
stargazer(probitmod1, hetmodel1, type="text")
```

```
=================================================
                Dependent variable:
                ---------------------------------
                          esvote
                    probit    heteroskedastic
                                    GLM
                     (1)            (2)
-------------------------------------------------
buspacs            -0.103***        -0.729*
                   (0.026)          (0.412)

labpacs             0.678***         5.850*
                   (0.150)          (3.491)

Constant            0.476            1.289
                   (0.440)          (0.935)

-------------------------------------------------
Observations         98               98
Log Likelihood     -19.729          -14.601
Akaike Inf. Crit.   45.458
=================================================
Note:              *p<0.1; **p<0.05; ***p<0.01
```

`stargazer(probitmod2, hetmodel2, type="text")`

```
=================================================
                Dependent variable:
                ---------------------------------
                          srvote
                    probit    heteroskedastic
                                    GLM
                     (1)            (2)
-------------------------------------------------
buspacs            -0.115***        -1.325
                   (0.030)          (1.124)

labpacs             0.767***        11.935
                   (0.179)         (10.439)

Constant            0.428            0.637
                   (0.453)          (0.768)

-------------------------------------------------
Observations         98               98
Log Likelihood     -17.235          -11.688
Akaike Inf. Crit.   40.470
=================================================
Note:              *p<0.1; **p<0.05; ***p<0.01
```

`lrtest(probitmod1, hetmodel1)`

Likelihood ratio test

```
Model 1: esvote ~ buspacs + labpacs
Model 2: esvote ~ buspacs + labpacs | buspacs + labpacs
  #Df  LogLik Df  Chisq Pr(>Chisq)
1    3 -19.729
2    5 -14.601  2 10.256    0.00593 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

`lrtest(probitmod2, hetmodel2)`

```
Likelihood ratio test

Model 1: srvote ~ buspacs + labpacs
Model 2: srvote ~ buspacs + labpacs | buspacs + labpacs
  #Df  LogLik Df  Chisq Pr(>Chisq)
1    3 -17.235
2    5 -11.688  2 11.095    0.003898 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Answer 2c

- Comparing the standard probit and heteroscedastic probit models shows that the direction of the coefficients for business and labor contributions is consistent across specifications: business PACs are negatively associated with support, while labor PACs are positively associated. However, the magnitude of the coefficients is much larger in the heteroscedastic model, reflecting the fact that variance heterogeneity is being explicitly modeled. Moreover, likelihood ratio tests strongly reject the null of homoscedasticity for both the economic stimulus and striker replacement votes ($p < 0.01$), indicating that the heteroscedastic probit fits the data significantly better. Substantively, this means that accounting for variance is important: business money not only shifts the probability of support but also increases the predictability of votes, an effect that a standard probit model masks.