Problem Set 2

Onno Steenweg

Due: February 18, 2026

Instructions

We're interested in what types of international environmental agreements or policies people support (Bechtel and Scheve 2013). So, we asked 8,500 individuals whether they support a given policy, and for each participant, we vary the (1) number of countries that participate in the international agreement and (2) sanctions for not following the agreement.

Load in the data labeled `climateSupport.RData` on GitHub, which contains an observational study of 8,500 observations.

- Response variable:
  - `choice`: 1 if the individual agreed with the policy; 0 if the individual did not support the policy
- Explanatory variables:
  - `countries`: Number of participating countries [20 of 192; 80 of 192; 160 of 192]
  - `sanctions`: Sanctions for missing emission reduction targets [None, 5%, 15%, and 20% of the monthly household costs given 2% GDP growth]

Question 1

Remember, we are interested in predicting the likelihood of an individual supporting a policy based on the number of countries participating and the possible sanctions for non-compliance.

Fit an additive model. Provide the summary output, the global null hypothesis, and p-value. Please describe the results and provide a conclusion.

Answers Question 1

To answer this question, I first verified that the structure of the dataset was appropriate. This allowed me to confirm that the variables were already correctly formatted, and in particular that the independent variables, countries and sanctions, were stored as categorical variables. Therefore, no additional conversion to factors was necessary before fitting the model.

Additive Model

To predict the likelihood of an individual supporting an environmental policy based on the number of participating countries and the level of sanctions, I fitted an additive logistic regression model. The model treats both predictors as ordered factors, allowing us to assess systematic trends in support as the number of countries and sanctions increase. The model is specified as:

$$\text{logit}(\Pr(\text{choice} = 1)) = \beta_0 + \beta_1 \text{countries} + \beta_2 \text{sanctions}$$

Here, the term "additive" means that the model includes the main effects of the explanatory variables, countries and sanctions, but does not include an interaction between them. In other words, the model estimates the independent effect of the number of participating countries and the level of sanctions on the log-odds of supporting the environmental policy, assuming that the effect of one variable does not depend on the level of the other. The results of the additive model can be seen in Table 1 below.

Table 1: Additive Logistic Regression Model

	Dependent variable:
	Support for Policy (1 = Yes)
Countries (Linear)	0.458^{***}
	(0.038)
Countries (Quadratic)	-0.010
	(0.038)
Sanctions (Linear)	-0.276^{***}
	(0.044)
Sanctions (Quadratic)	-0.181^{***}
	(0.044)
Sanctions (Cubic)	0.150^{***}
	(0.044)
Constant	-0.006
	(0.022)
Observations	8,500
Log Likelihood	$-5{,}784.130$
Akaike Inf. Crit.	11,580.260
Note:	*p<0.1; **p<0.05; ***p<0.01

Global Null Hypothesis

The global null hypothesis states that none of the predictors have any effect on the probability of supporting the policy:

$$H_0 : \beta_{\text{countries}} = \beta_{\text{sanctions}} = 0$$

I tested the global null hypothesis, which states that neither the number of participat-

ing countries nor the level of sanctions affects support for the policy, by comparing the additive model to a null (intercept-only) model using a likelihood ratio (Chi-square) test. This comparison showed a deviance reduction of 215.15 with 5 degrees of freedom and a p-value less than 2.2×10^{-16}, providing extremely strong evidence to reject the global null hypothesis and conclude that at least one of the predictors significantly influences support for the policy.

Interpretation and Conclusion

The results indicate that both the number of participating countries and the level of sanctions are significant predictors of whether individuals support the environmental agreement:

- The linear term for countries is positive and highly significant, suggesting that increasing the number of participating countries systematically increases support for the policy.
- The sanctions terms show a significant linear and higher-order trend, indicating that support varies non-linearly across different levels of sanctions.

Overall, I conclude that international participation levels and sanctions both play an important role in shaping public support for environmental agreements. In particular, the positive and highly significant linear term for countries indicates that support increases as more countries participate. For sanctions, the significant linear, quadratic, and cubic terms suggest a non-linear relationship, meaning support varies depending on the level of sanctions. These individual effects, together with the results of the global null hypothesis test which rejected the null model in favor of the additive model with a deviance reduction of 215.15 on 5 degrees of freedom and a p-value $< 2.2 \times 10^{-16}$, provide strong evidence that the predictors collectively explain a substantial portion of variation in support. Overall, the results indicate that both international participation and sanction levels play an important role in shaping public support for environmental agreements.

All of the above was done using the code below:

```r
## Check structure
str(climateSupport) #all good

## Fit additive model
additive_model <- glm(choice ~ countries + sanctions,
                      data = climateSupport,
```

```
                        family = binomial)

summary(additive_model)

## Create LaTeX table for Overleaf
library(stargazer)
stargazer(additive_model,
          type = "latex",
          title = "Additive Logistic Regression Model",
          label = "tab:additive_model",
          dep.var.labels = "Support for Policy (1 = Yes)",
          covariate.labels = c("Countries (Linear)",
                               "Countries (Quadratic)",
                               "Sanctions (Linear)",
                               "Sanctions (Quadratic)",
                               "Sanctions (Cubic)"),
          digits = 3,
          out = "additive_model.tex")

## Fit null model for global null hypothesis
null_model <- glm(choice ~ 1, data = climateSupport, family = binomial)
anova(null_model, additive_model, test="Chisq")
```

Question 2

If any of the explanatory variables are statistically significant in this model, then:

1. For the policy in which nearly all countries participate [160 of 192], how does increasing sanctions from 5% to 15% change the odds that an individual will support the policy? (Interpretation of a coefficient)

2. For the policy in which very few countries participate [20 of 192], how does increasing sanctions from 5% to 15% change the odds that an individual will support the policy? (Interpretation of a coefficient)

3. What is the estimated probability that an individual will support a policy if there are 80 of 192 countries participating with no sanctions?

Answers Question 2

The additive logistic regression model from Question 1 indicated that both explanatory variables are statistically significant, meaning they meaningfully predict whether an individual supports an environmental policy. To interpret these effects, I calculated the predicted log-odds and convert them into odds ratios using the original dataset. The odds ratio for a change in a predictor is computed as

$$\text{odds ratio} = \exp(\text{mean logit at new level} - \text{mean logit at original level})$$

and the corresponding percentage change in odds is $(\text{odds ratio} - 1) \times 100\%$. Using this approach, I find the following results:

1. For a policy with nearly all countries participating (160 of 192), increasing sanctions from 5% to 15% reduces the odds of an individual supporting the policy by approximately 27.8% (odds ratio = 0.722). This means higher sanctions significantly decrease support even when international participation is very high.

2. For a policy with very few countries participating (20 of 192), increasing sanctions from 5% to 15% also reduces the odds of support by the same 27.8%, indicating that the negative effect of sanctions on support is consistent across low and high levels of international participation.

3. For a policy with 80 of 192 countries participating and no sanctions, the estimated probability that an individual will support the policy is approximately 51.6%, showing that moderate participation with no sanctions yields roughly even likelihood of support.

These results indicate that increasing sanctions reduces the odds of supporting the policy in both cases. For moderate participation (80 of 192 countries) without sanctions, the predicted probability of support is roughly 51.6%, indicating near-even levels of support. Overall, this seems to indicate that sanctions have a negative effect on support, while international participation levels influence baseline probabilities of support.

The calculations were done using the code below:

```
library(dplyr)

# 1. Odds ratio: 160 countries, sanctions 5% -> 15%
```

```r
rows_160_5 <- climateSupport %>%
  filter(countries == "160 of 192", sanctions == "5%")
rows_160_15 <- climateSupport %>%
  filter(countries == "160 of 192", sanctions == "15%")

logit_160_5 <- predict(additive_model, newdata = rows_160_5, type="link")
logit_160_15 <- predict(additive_model, newdata = rows_160_15, type="link")

# Average logit for each group, then compute odds ratio
odds_ratio_160 <- exp(mean(logit_160_15) - mean(logit_160_5))

# 2. Odds ratio: 20 countries, sanctions 5% -> 15%
rows_20_5 <- climateSupport %>%
  filter(countries == "20 of 192", sanctions == "5%")
rows_20_15 <- climateSupport %>%
  filter(countries == "20 of 192", sanctions == "15%")

logit_20_5 <- predict(additive_model, newdata = rows_20_5, type="link")
logit_20_15 <- predict(additive_model, newdata = rows_20_15, type="link")

odds_ratio_20 <- exp(mean(logit_20_15) - mean(logit_20_5))

# 3. Estimated probability: 80 countries, no sanctions
rows_80_none <- climateSupport %>%
  filter(countries == "80 of 192", sanctions == "None")

prob_80_none <- mean(predict(additive_model, newdata = rows_80_none, type="response"))

# 4. Results
odds_ratio_160
odds_ratio_20
prob_80_none
```

Question 3

Would the answers to 2a and 2b potentially change if we included an interaction term in this model? Why?

Perform a test to see if including an interaction is appropriate.

Answers Question 3

Including an interaction term between countries and sanctions would allow the effect of sanctions on support to differ depending on the number of participating countries. If such an interaction is significant, the odds ratios I computed in Task 2a and 2b could change, because the effect of sanctions would no longer be constant across country participation levels.

To examine whether the effect of sanctions on individual support depends on the number of participating countries, I extended the additive logistic regression model by including an interaction term between countries and sanctions. I then compared the additive model to this interaction model using a likelihood ratio test. This approach tests whether adding the interaction significantly improves model fit.

The results of the test showed a reduction in residual deviance of 6.293 with 6 degrees of freedom, corresponding to a p-value of 0.3912. This indicates that including the interaction does not significantly improve the model.

Based on this outcome, I concluded that the effect of sanctions on support does not significantly vary by the number of participating countries. Therefore, the additive model is sufficient, and the odds ratios calculated in Task 2a and 2b (a 27.8% reduction in odds when sanctions increase from 5% to 15%) remain valid across both low (20 of 192) and high (160 of 192) participation scenarios. In other words, sanctions consistently reduce support regardless of international participation levels.

The calculations were done using the code below:

```r
interaction_model <- glm(choice ~ countries * sanctions,
                         data = climateSupport,
                         family = binomial)

anova(additive_model, interaction_model, test = "Chisq")
```