# **MS Contribution to:** Lee, C. (2019). China's Energy Diplomacy: Does Chinese Foreign Policy Favor Oil-Producing Countries? Foreign Policy Analysis, 15(4), 570–588. https://doi.org/10.1093/fpa/orz011

Applied Stats II

**Maiia Skrypnyk 23371609**

Due: March 31, 2024

## 1 Twist 1: Plots

**Loading the data:**

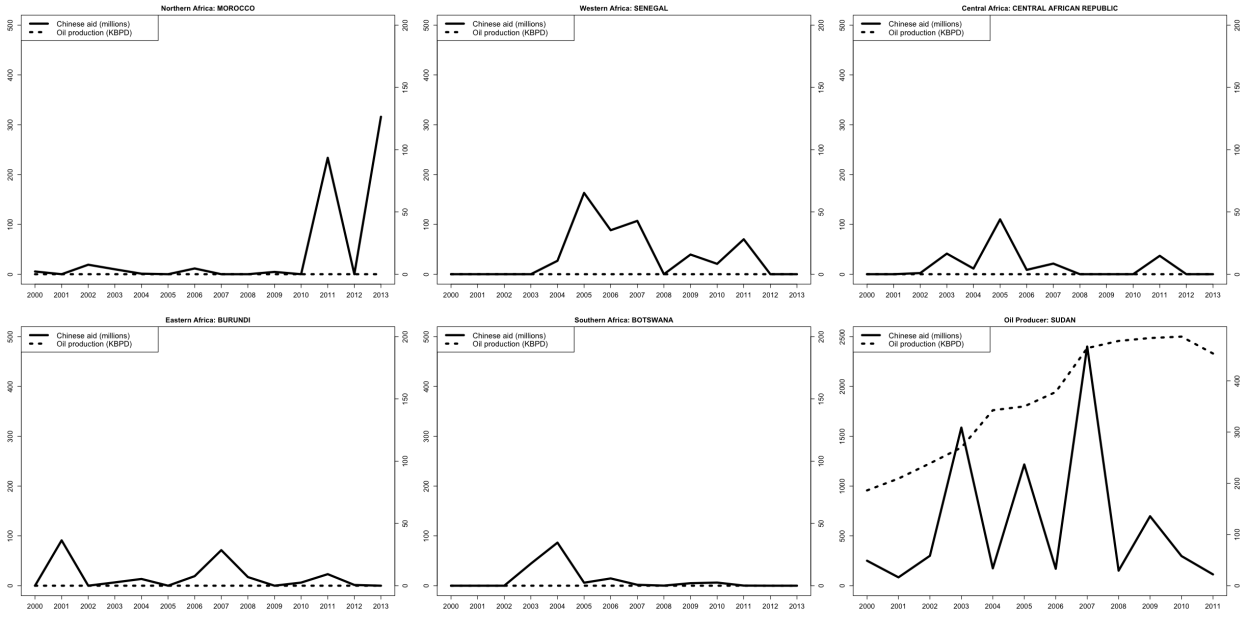
```
data_part <- read.csv("data_partner.csv") # Data on China's diplomatic
    partnerships (cross-national data)
data_TSCS <- read.csv("data_TSCS.csv") # Data on China's aid to Africa (TSCS
    data)
data_aid <- read.csv("data_aid.csv") # Data on China's aid to Africa (cross-
    national data)
data_visit <- read.csv("data_visit.csv") # Data on China's leadership visits (
    cross-national data)
```

**Figure 2 Twist:**

```
africa <- subset(data_TSCS, africa==1 & year>1999)
africa <- cbind(africa["country"], africa["year"], africa["china_aid"], africa
    ["production2"])
africa$china_aid <- ifelse(is.na(africa$china_aid)==T, 0, africa$china_aid)
africa <- africa[order(africa$country, africa$year), ]

Morocco <- subset(africa, country == "Morocco")
Senegal <- subset(africa, country == "Senegal")
CAR <- subset(africa, country == "Central African Republic")
Burundi <- subset(africa, country == "Burundi")
Botswana <- subset(africa, country == "Botswana")
Sudan <- subset(africa, country == "Sudan")

par(mar=c(4,3,2,3), mfrow=c(2, 3))

#Northern Africa: Morocco
```

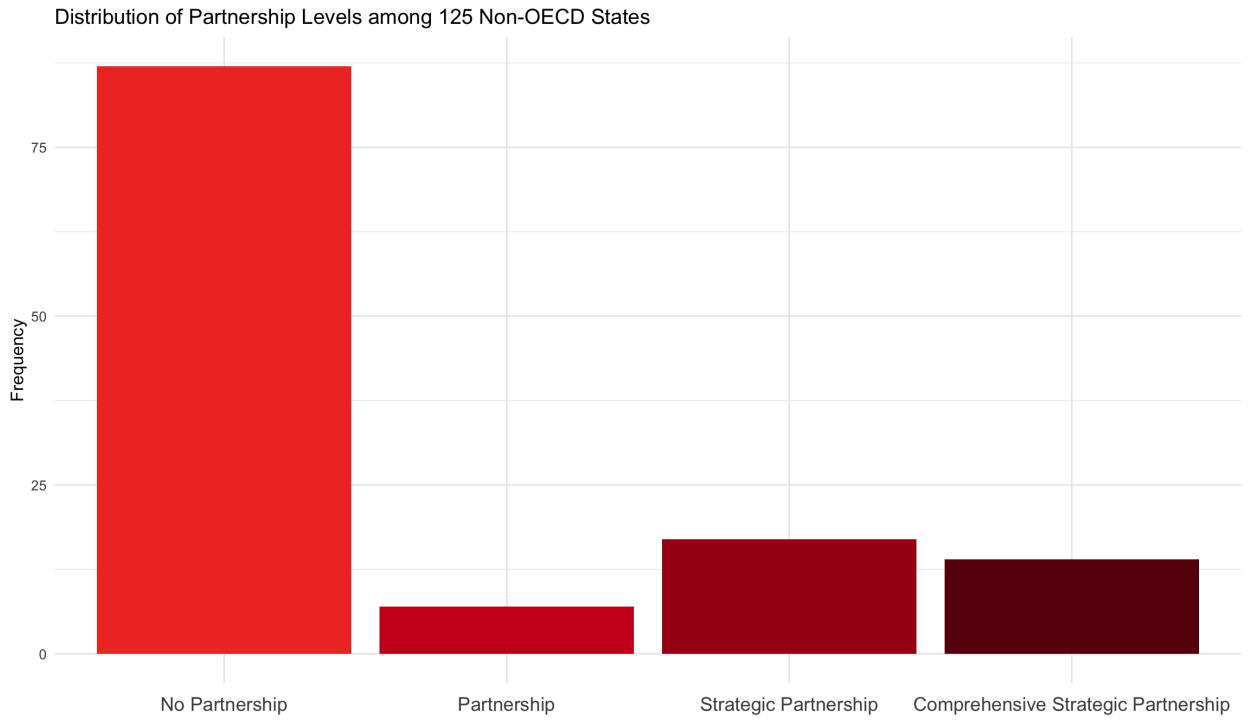
```
plot(Morocco$china_aid/1000000 ~ Morocco$year, type="l", lwd=4, xlab="", ylab="", ylim=c(0, 500), axes=F, main="Northern Africa: MOROCCO")
axis(1, at= Morocco$year, label= Morocco$year, cex.axis=1.2)
axis(2, at=seq(0, 500, by=100), label=seq(0, 500, by=100), cex.axis=1.2)
par(new=T)
plot(Morocco$production2/1000 ~ Morocco$year, type="l", lwd=4, lty=9, xlab="", ylab="", ylim=c(0, 200), axes=F)
axis(4, at=seq(0, 200, by=50), label=seq(0, 200, by=50), cex.axis=1.2)
box()
legend("topleft", c("Chinese aid (millions)", "Oil production (KBPD)"), cex=1.3, lwd=4, lty=c(1, 9))

#Western Africa: Senegal
plot(Senegal$china_aid/1000000 ~ Senegal$year, type="l", lwd=4, xlab="", ylab="", ylim=c(0, 500), axes=F, main="Western Africa: SENEGAL")
axis(1, at= Senegal$year, label= Senegal$year, cex.axis=1.2)
axis(2, at=seq(0, 500, by=100), label=seq(0, 500, by=100), cex.axis=1.2)
par(new=T)
plot(Senegal$production2/1000 ~ Senegal$year, type="l", lwd=4, lty=9, xlab="", ylab="", ylim=c(0, 200), axes=F)
axis(4, at=seq(0, 200, by=50), label=seq(0, 200, by=50), cex.axis=1.2)
box()
legend("topleft", c("Chinese aid (millions)", "Oil production (KBPD)"), cex=1.3, lwd=4, lty=c(1, 9))

#Central Africa: Central African Republic
plot(CAR$china_aid/1000000 ~ CAR$year, type="l", lwd=4, xlab="", ylab="", ylim=c(0, 500), axes=F, main="Central Africa: CENTRAL AFRICAN REPUBLIC")
axis(1, at= CAR$year, label= CAR$year, cex.axis=1.2)
axis(2, at=seq(0, 500, by=100), label=seq(0, 500, by=100), cex.axis=1.2)
par(new=T)
plot(CAR$production2/1000 ~ CAR$year, type="l", lwd=4, lty=9, xlab="", ylab="", ylim=c(0, 200), axes=F)
axis(4, at=seq(0, 200, by=50), label=seq(0, 200, by=50), cex.axis=1.2)
box()
legend("topleft", c("Chinese aid (millions)", "Oil production (KBPD)"), cex=1.3, lwd=4, lty=c(1, 9))

#Eastern Africa: Burundi
plot(Burundi$china_aid/1000000 ~ Burundi$year, type="l", lwd=4, xlab="", ylab="", ylim=c(0, 500), axes=F, main="Eastern Africa: BURUNDI")
axis(1, at= Burundi$year, label= Burundi$year, cex.axis=1.2)
axis(2, at=seq(0, 500, by=100), label=seq(0, 500, by=100), cex.axis=1.2)
par(new=T)
plot(Burundi$production2/1000 ~ Burundi$year, type="l", lwd=4, lty=9, xlab="", ylab="", ylim=c(0, 200), axes=F)
axis(4, at=seq(0, 200, by=50), label=seq(0, 200, by=50), cex.axis=1.2)
box()
legend("topleft", c("Chinese aid (millions)", "Oil production (KBPD)"), cex=1.3, lwd=4, lty=c(1, 9))

```

```r
#Southern Africa: Botswana
plot(Botswana$china_aid/1000000 ~ Botswana$year, type="l", lwd=4, xlab="", ylab="", ylim=c(0, 500), axes=F, main="Southern Africa: BOTSWANA")
axis(1, at= Botswana$year, label= Botswana$year, cex.axis=1.2)
axis(2, at=seq(0, 500, by=100), label=seq(0, 500, by=100), cex.axis=1.2)
par(new=T)
plot(Botswana$production2/1000 ~ Botswana$year, type="l", lwd=4, lty=9, xlab="", ylab="", ylim=c(0, 200), axes=F)
axis(4, at=seq(0, 200, by=50), label=seq(0, 200, by=50), cex.axis=1.2)
box()
legend("topleft", c("Chinese aid (millions)", "Oil production (KBPD)"), cex=1.3, lwd=4, lty=c(1, 9))

#Bonus: oil-producing Sudan
plot(Sudan$china_aid/1000000 ~ Sudan$year, type = "l", lwd = 4, xlab = "", ylab = "", ylim = c(0, 2500), axes = FALSE, main = "Oil Producer: SUDAN")
axis(1, at = Sudan$year, label = Sudan$year, cex.axis = 1.2)
axis(2, at = seq(0, 2500, by = 500), label = seq(0, 2500, by = 500), cex.axis = 1.2)
par(new = TRUE)
plot(Sudan$production2 ~ Sudan$year, type = "l", lwd = 4, lty = 9, xlab = "", ylab = "", ylim = c(0, max(Sudan$production2, na.rm = TRUE)), axes = FALSE)
axis(4, at = seq(0, max(Sudan$production2, na.rm = TRUE), by = 100), label = seq(0, max(Sudan$production2, na.rm = TRUE), by = 100), cex.axis = 1.2)
box()
legend("topleft", c("Chinese aid (millions)", "Oil production (KBPD)"), cex = 1.3, lwd = 4, lty = c(1, 9))
```



**Model 1 Twist:**

```r
#PREPARING THE DATA

#1
#Model 1 dropped 4 observations from the 'subset(data_part, oecd==0)' (129),
#so there were 125 states left. After re-checking with the Table A2, and manually,
#we know that the 4 states dropped are: Cuba, Kyrgyzstan, Somalia, South Sudan.

#Subsetting 125 non-OECD members
data_part <- subset(data_part, !(country %in% c("Cuba", "Kyrgyzstan", "Somalia", "South Sudan")))
data_part125 <- subset(data_part, oecd==0)

#2
#Creating new column for the partnership level
data_part125$partlevel <- NA

#Combining partnership levels (dummy variables) into one column
data_part125$partlevel[data_part125$part == 1] <- "Partnership"
data_part125$partlevel[data_part125$part2 == 1] <- "Strategic Partnership"
data_part125$partlevel[data_part125$part3 == 1] <- "Comprehensive Strategic Partnership"
data_part125$partlevel[is.na(data_part125$partlevel)] <- "No Partnership"

#Recoding 'partlevel' into a categorical variable (as factor), unordered (for now),
#with reference category being "No Partnership"
data_part125$partlevel <- relevel(factor(data_part125$partlevel),
                                  ref = "No Partnership")

#Bar plot for 'partlevel'
color_palette <- brewer.pal(9, "Reds")[6:9] #Custom color palette in dark red colours
ggplot(data_part125, aes(x = factor(partlevel, levels = c("No Partnership", "Partnership", "Strategic Partnership", "Comprehensive Strategic Partnership")))) +
  geom_bar(fill = color_palette) +
  labs(title = "Distribution of Partnership Levels among 125 Non-OECD States",
       x = "",
       y = "Frequency") +
  theme_minimal() +
  theme(axis.text.x = element_text(size = 12, vjust = -0.5))
```




```r
#ORIGINAL MODEL 1 (DV: Partnership: 0 for no partnership,
#1 for partnership (any kind))
#Binomial logit regression. OECD countries excluded
model1 <- glm(partner~prod2+GDPpc2+growth2+FDI2+trade.de2+polity2+dom2+usally2
    , family=binomial, data = data_part125)
summary(model1)

#Exponentiating coefficients (to OR).
#Checking again by adding confidence intervals;
#CI should not contain 1 for a significant result (odds either below or above
    1)
exp_model1 <- exp(cbind(OR = coef(model1), confint(model1)))
print(round(exp_model1, 3))
stargazer(exp_model1, type="latex", t.auto=F); logLik(model1)
```

Table 1: Original Model 1 (exponentiated)

| | *Dependent variable:* |
|---|---|
| | Partnership |
| Oil production | 1.139** |
| | (0.058) |
| GDP per capita | 0.715 |
| | (0.270) |
| GDP growth | 0.968 |
| | (0.089) |
| FDI inflows | 1.299*** |
| | (0.081) |
| Trade importance | 3.678 |
| | (0.826) |
| Level of democracy | 0.993 |
| | (0.051) |
| Domestic conflict | 1.223 |
| | (0.138) |
| US ally | 0.398 |
| | (0.777) |
| Constant | 0.012** |
| | (2.226) |
| Observations | 125 |
| Log Likelihood | −55.546 |
| Residual Deviance | 111.09 |
| AIC | 129.091 |
| BIC | 154.5459 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

```r
#Calculating predicted probabilities and converting it to a factor,
#where a probability <0.05 = 0 (no partnership), >0.5 = 1 (partnership)
predprob_model1 <- factor(ifelse(predict(model1, type = "response") > 0.5, 1,
    0), levels = c(0, 1))
#Converting the dependent variable into factor
data_part125$partner <- factor(data_part125$partner)
#Confusion matrix
cm_model1 <- confusionMatrix(predprob_model1, data_part125$partner)
cm_model1 #Accuracy score 0.784
```

```
Confusion Matrix and Statistics

          Reference
Prediction  0  1
         0 78 18
         1  9 20

               Accuracy : 0.784
                 95% CI : (0.7015, 0.8526)
    No Information Rate : 0.696
    P-Value [Acc > NIR] : 0.01828

                  Kappa : 0.4531

 Mcnemar's Test P-Value : 0.12366

            Sensitivity : 0.8966
            Specificity : 0.5263
         Pos Pred Value : 0.8125
         Neg Pred Value : 0.6897
             Prevalence : 0.6960
         Detection Rate : 0.6240
   Detection Prevalence : 0.7680
      Balanced Accuracy : 0.7114

       'Positive' Class : 0
```

```r
#UNORDERED MODEL. OECD countries excluded.
unordmodel <- multinom(partlevel~prod2+GDPpc2+growth2+FDI2+
                         trade.de2+trade.de2+polity2+dom2+usally2,
                       data_part125)
summary(unordmodel)

#Calculating p-values to estimate significance of the coefficients
coefs <- coef(unordmodel)
std_errors <- summary(unordmodel)$standard.errors
t_values <- coefs/std_errors
df <- nrow(data_part125) - length(coef(unordmodel))
```

```r
p_values <- 2 * pnorm(abs(t_values), lower.tail = FALSE)
p_values < 0.05
```

```
                                (Intercept) prod2 GDPpc2 growth2  FDI2 trade.de2 polity2  dom2 usally2
Comprehensive Strategic Partnership       FALSE FALSE   TRUE   FALSE  TRUE     FALSE   FALSE FALSE   FALSE
Partnership                               FALSE FALSE  FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE
Strategic Partnership                     FALSE  TRUE  FALSE   FALSE  TRUE     FALSE   FALSE FALSE   FALSE
```

```r
#Exponentiating coefficients (to OR).
#Checking again by adding confidence intervals;
#CI should not contain 1 for a significant result (odds either below or above
    1)
exp_unordmodel <- exp(coef(unordmodel))
print(round(exp_unordmodel, 3))
stargazer(exp_unordmodel, type="latex", t.auto=F); logLik(unordmodel)
```

Table 2: Unordered Model (exponentiated)

| | *Dependent variable:* | | |
|---|---|---|---|
| | Comprehensive Strategic Partnership | Partnership | Strategic Partnership |
| | (1) | (2) | (3) |
| Oil production | 1.221** | 1.013 | 1.197** |
| | (0.104) | (0.105) | (0.083) |
| GDP per capita | 0.347** | 1.004 | 0.756 |
| | (0.520) | (0.490) | (0.363) |
| GDP growth | 1.003 | 0.864 | 0.980 |
| | (0.157) | (0.227) | (0.103) |
| FDI inflows | 1.885** | 1.133 | 1.266** |
| | (0.284) | (0.148) | (0.097) |
| Trade importance | 10.218** | 0.153 | 2.253 |
| | (1.189) | (4.021) | (1.083) |
| Level of democracy | 0.944 | 1.095 | 0.973 |
| | (0.087) | (0.106) | (0.067) |
| Domestic conflict | 0.957 | 1.308 | 1.337 |
| | (0.239) | (0.236) | (0.177) |
| US ally | 1.536 | 0.230 | 0.288 |
| | (1.268) | (1.281) | (1.047) |
| Constant | 0.000 | 0.008 | 0.003** |
| | (5.576) | (3.694) | (2.928) |
| Observations | 125 | | |
| Log Likelihood | -83.263 | | |
| Residual Deviance | 166.526 | | |
| AIC | 220.526 | | |
| BIC | 296.89 | | |
| *Note:* | | | *p<0.1; **p<0.05; ***p<0.01 |

```
#Calculating predicted probabilities
```

```r
predprob_unord <- predict(unordmodel, newdata = data_part125, type = "class")

#Confusion matrix
cm_unordmodel <- confusionMatrix(predprob_unord, data_part125$partlevel)
cm_unordmodel #Accuracy score 0.744
```

```
Confusion Matrix and Statistics

                                 Reference
Prediction                        No Partnership Comprehensive Strategic Partnership Partnership Strategic Partnership
  No Partnership                              81                                   5           7                    12
  Comprehensive Strategic Partnership          3                                   8           0                     1
  Partnership                                  0                                   0           0                     0
  Strategic Partnership                        3                                   1           0                     4


Overall Statistics

               Accuracy : 0.744
                 95% CI : (0.6582, 0.8178)
    No Information Rate : 0.696
    P-Value [Acc > NIR] : 0.1419

                  Kappa : 0.3534

 Mcnemar's Test P-Value : NA

Statistics by Class:

                     Class: No Partnership Class: Comprehensive Strategic Partnership Class: Partnership
Sensitivity                         0.9310                                     0.5714              0.000
Specificity                         0.3684                                     0.9640              1.000
Pos Pred Value                      0.7714                                     0.6667                NaN
Neg Pred Value                      0.7000                                     0.9469              0.944
Prevalence                          0.6960                                     0.1120              0.056
Detection Rate                      0.6480                                     0.0640              0.000
Detection Prevalence                0.8400                                     0.0960              0.000
Balanced Accuracy                   0.6497                                     0.7677              0.500
                     Class: Strategic Partnership
Sensitivity                            0.2353
Specificity                            0.9630
Pos Pred Value                         0.5000
Neg Pred Value                         0.8889
Prevalence                             0.1360
Detection Rate                         0.0320
Detection Prevalence                   0.0640
Balanced Accuracy                      0.5991
```

```r
#ORDERED MODEL. OECD countries excluded.
#Recoding 'partlevel' into a categorical ORDERED variable (as factor)
data_part125$partlevel <- factor(data_part125$partlevel,
                                 ordered = TRUE,
                                 levels = c("No Partnership",
                                            "Partnership",
                                            "Strategic Partnership",
```

```
                                                    "Comprehensive Strategic
    Partnership"))
#Fitting an ordered model
ordmodel<- polr(partlevel~prod2+GDPpc2+growth2+FDI2+
                trade.de2+trade.de2+polity2+dom2+usally2,
                data_part125)
summary(ordmodel)

#Calculating p-values to estimate significance of the coefficients
t_values <- coef(summary(ordmodel))[, "t value"]
df <- nrow(data_part125) - length(coef(ordmodel))
p_values <- 2 * pt(abs(t_values), df = df, lower.tail = FALSE)
p_values < 0.05
```

```
                                                    prod2
                                                     TRUE
                                                  growth2
                                                    FALSE
                                                 trade.de2
                                                     TRUE
                                                     dom2
                                                    FALSE
                                  No Partnership|Partnership
                                                     TRUE
Strategic Partnership|Comprehensive Strategic Partnership
                                                     TRUE
```

```
                                                    GDPpc2
                                                     FALSE
                                                     FDI2
                                                     TRUE
                                                  polity2
                                                     FALSE
                                                  usally2
                                                     FALSE
                                  Partnership|Strategic Partnership
                                                     TRUE
```

```
##Exponentiating coefficients (to OR).
#Checking again by adding confidence intervals;
#CI should not contain 1 for a significant result (odds either below or above
    1)
exp_ordmodel <- exp(cbind(OR = coef(ordmodel), confint(ordmodel)))
print(round(exp_ordmodel, 3))
stargazer(exp_ordmodel, type="latex", t.auto=F); logLik(ordmodel)
```

Table 3: Ordered Model (exponentiated)

| | *Dependent variable:* |
|---|---|
| | Partnership Level |
| Oil production | 1.124** |
| | (0.057) |
| GDP per capita | 0.694 |
| | (0.256) |
| GDP growth | 0.958 |
| | (0.086) |
| FDI inflows | 1.324*** |
| | (0.078) |
| Trade importance | 6.095** |
| | (0.745) |
| Level of democracy | 0.953 |
| | (0.050) |
| Domestic conflict | 1.155 |
| | (0.131) |
| US ally | 0.841 |
| | (0.678) |
| *Intercepts* | |
| No Partnership—Partnership | 77.138** |
| Partnership—Strategic Partnership | 118.238** |
| Strategic Partnership—Comprehensive Strategic Partnership | 513.783** |
| Observations | 125 |
| Log Likelihood | −90.19376 |
| Residual Deviance | 180.3875 |
| AIC | 202.3875 |
| BIC | 233.499 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

```r
#Calculating predicted probabilities
predprob_ord <- predict(ordmodel, newdata = data_part125, type = "class")

#Confusion matrix
cm_ordmodel <- confusionMatrix(data = predprob_ord, reference = data_part125$
    partlevel)
cm_ordmodel #Accuracy score 0.752
```

```
Confusion Matrix and Statistics

                                  Reference
Prediction                         No Partnership Partnership Strategic Partnership
  No Partnership                               82           7                     11
  Partnership                                   0           0                      0
  Strategic Partnership                         2           0                      4
  Comprehensive Strategic Partnership           3           0                      2
                                  Reference
Prediction                         Comprehensive Strategic Partnership
  No Partnership                                                   4
  Partnership                                                      0
  Strategic Partnership                                            2
  Comprehensive Strategic Partnership                              8
```

```
Overall Statistics

               Accuracy : 0.752
                 95% CI : (0.6668, 0.8249)
    No Information Rate : 0.696
    P-Value [Acc > NIR] : 0.1016

                  Kappa : 0.3809

 Mcnemar's Test P-Value : NA

Statistics by Class:

                     Class: No Partnership Class: Partnership Class: Strategic Partnership
Sensitivity                         0.9425              0.000                       0.2353
Specificity                         0.4211              1.000                       0.9630
Pos Pred Value                      0.7885                NaN                       0.5000
Neg Pred Value                      0.7619              0.944                       0.8889
Prevalence                          0.6960              0.056                       0.1360
Detection Rate                      0.6560              0.000                       0.0320
Detection Prevalence                0.8320              0.000                       0.0640
Balanced Accuracy                   0.6818              0.500                       0.5991
                     Class: Comprehensive Strategic Partnership
Sensitivity                                              0.5714
Specificity                                              0.9550
Pos Pred Value                                           0.6154
Neg Pred Value                                           0.9464
Prevalence                                               0.1120
Detection Rate                                           0.0640
Detection Prevalence                                     0.1040
Balanced Accuracy                                        0.7632
```

```r
#MODELS STATISTICS
#Residual Deviance
deviance(unordmodel)
deviance(ordmodel)
deviance(model1)
```

```
[1] 166.5255
[1] 180.3875
[1] 111.0911
```

```r
#Akaike Information Criterion (the lower value, the better)
AIC(model1)
AIC(unordmodel)
AIC(ordmodel)
```

```
[1] 129.0911
[1] 220.5255
[1] 202.3875
```

```
#Bayesian Information Criterion (the lower value, the better)
BIC(model1)
BIC(unordmodel)
BIC(ordmodel)
```

```
[1] 154.5459
[1] 296.89
[1] 233.499
```

```
#Brant test for the ordered model (testing the "proportional odds assumption")
brant(ordmodel)
```

| Test for | X2 | df | probability |
|---|---|---|---|
| Omnibus | 15.7 | 16 | 0.47 |
| prod2 | 1.74 | 2 | 0.42 |
| GDPpc2 | 2.55 | 2 | 0.28 |
| growth2 | -2.46 | 2 | 1 |
| FDI2 | 1.66 | 2 | 0.44 |
| trade.de2 | 3.99 | 2 | 0.14 |
| polity2 | 1.39 | 2 | 0.5 |
| dom2 | 3.42 | 2 | 0.18 |
| usally2 | 3.64 | 2 | 0.16 |

H0: Parallel Regression Assumption holds

# 2 References

Lee, C. (2019). China's Energy Diplomacy: Does Chinese Foreign Policy Favor Oil-Producing Countries? *Foreign Policy Analysis*, **15**(4), 570–588. https://doi.org/10.1093/fpa/orz011

Lee, C. (2023). "Replication data for 'China's Energy Diplomacy: Does Chinese Foreign Policy Favor Oil Producing Countries?"" [Data set]. Harvard Dataverse. https://doi.org/10.7910/DVN/7E3O5P, V1, UNF:6:x+exuF4WVfknqWQhUWz4MQ== [fileUNF].