Appendix

Louisa Boulaziz

May 26, 2021

Contents

1 Loading in data and packages

In this section I am loading in the data and all the packages I use to run the analysis.

2 Structure of the data

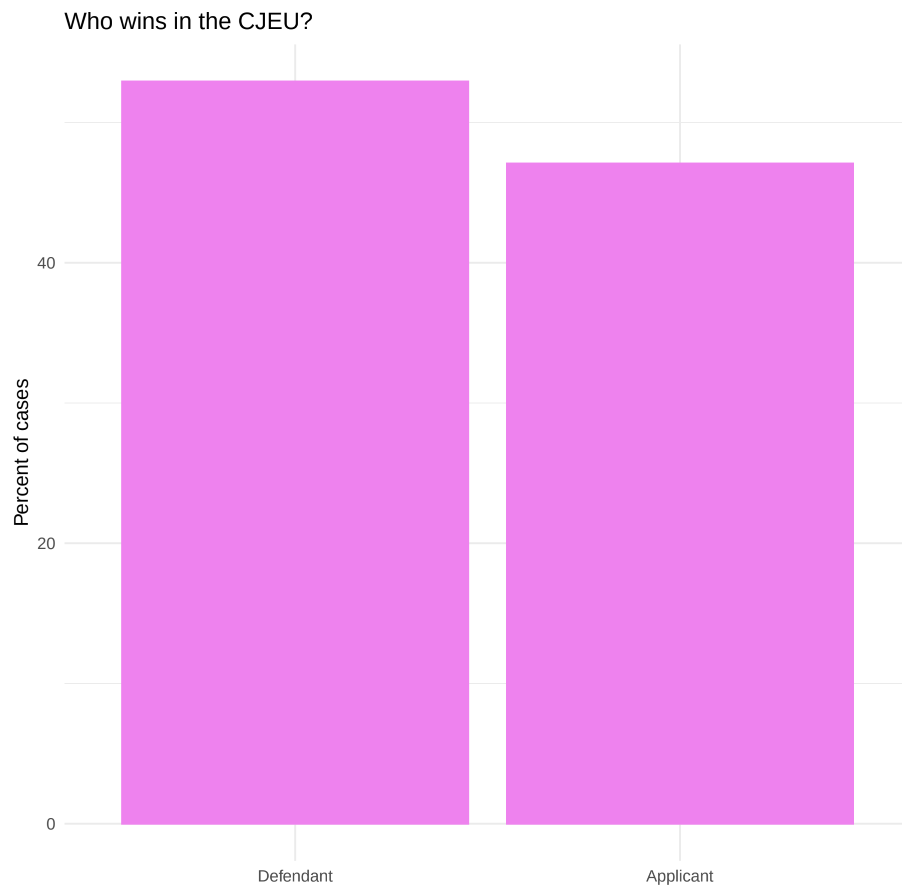
I have coded all the parties in all preliminary reference cases and their lawyers from the first judgment issued in CJEU dating to 2016. I have merged this data together with the data used in the Carrubba et al (2008) article. Thus the data consists of all cases dating from 1963 to 1994 for which Carrubba et al has coded the outcome of the case. In the sample I have 2770 observations (year when case was registered). Each new row in the dataset is an applicant or a defendant in a preliminary reference case. The applicant and the defendant are nested in a case, and they either have a lawyer or they do not (binary indicator). If they have a lawyer or lawyer team which has experience this is also flagged (binary indicator). The parties also win or loses a case which is also a binary indicator. In the data I have information about the number of member state governments who submits observations in favor of either the applicant or the defendant. I do also have information about whether or not the commission submitted observations in favor of either the applicant or the defendant. I have a case salience measure – the total number of governments who submits observations.

```
# Total number of cases in the data
length(unique(data$celex))

## [1] 1301

table(data$member_state)

##
##   10   12    6    9
##  658 1146  240  726

range(data$year)

## [1] "1963" "1994"
```

3 Dependent variable

I have two variables that measures whether the applicant or defendant won the case. One is at the case level – `ecjplaintiffagree` – and the other one is at the level of the role in the dispute – `win`. These variables measures the same thing – the cases the applicants win and the cases the defendants win.

```
prop.table(table(data$ecjplaintiffagree))*100

##
##        0        1
## 52.92419 47.07581

# Visualizing it



ggplot(data, aes(as.factor(ecjplaintiffagree))) +
 geom_bar(position = "dodge", aes(y = (..count..)/sum(..count..)*100), colour = "violet", fi
  ylab("Percent of cases") +
  xlab("")+
  ggtitle("Who wins in the CJEU?")+
  theme_minimal() +
  scale_x_discrete(labels=c("Defendant", "Applicant"))
```



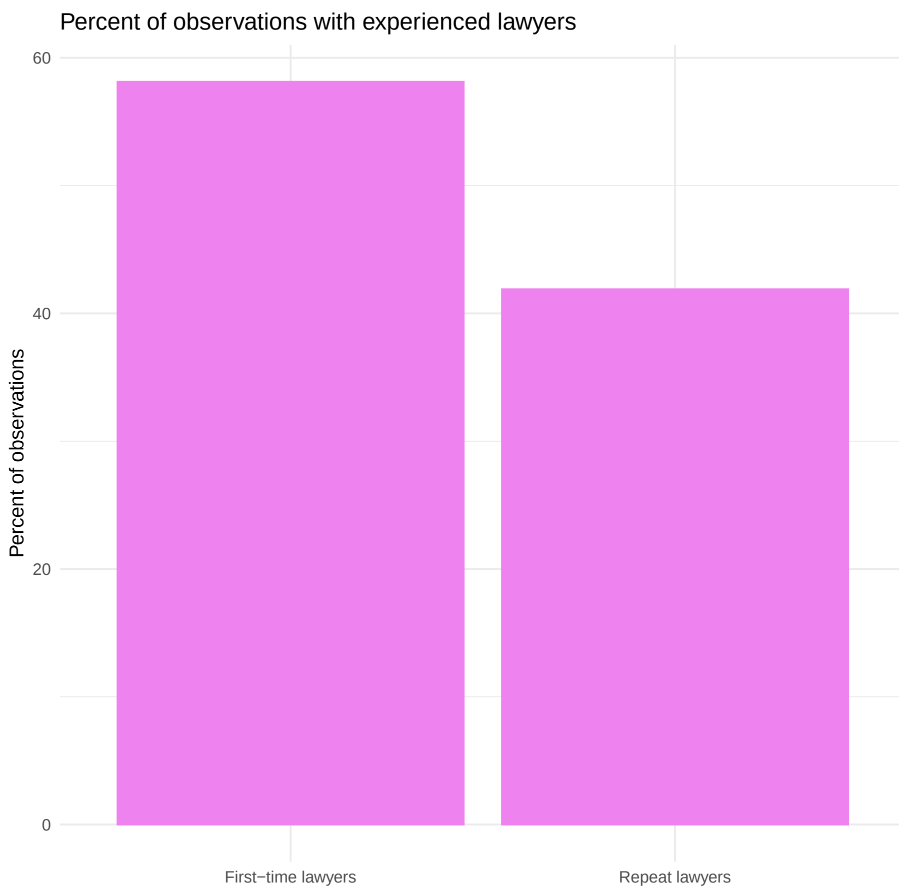
4 Main independent variables

I focus on three main explanatory variables in the analysis: `experience`, `lawyer` and `salience`. I describe them more in detail below.

4.1 Experience

The experience-variable is grouped at the level of the role in the legal dispute. This means that it measures whether or not the applicants and the defendants have experienced lawyers. The variable shows whether or not the applicant or the defendant has a lawyer or lawyer team which have prior litigation experience. Note that all 1s represent lawyers and lawyer teams that have at least argued one prior case at the CJEU. All 0s represent non-lawyers – that means when applicants and defendants come to court without representation the variable shows 0. The variable also shows 0 for applicant and defendants who come to court with lawyers without prior litigation experience.

```
# Filtering out observations with lawyers
experience <- data %>%
  filter(lawyer == 1)

experience %>%
  group_by(experience) %>%
  ggplot(aes(as.factor(experience))) +
  geom_bar(position = "dodge", aes(y = (..count..)/sum(..count..)*100), colour = "violet", f
  ylab("Percent of observations") +
  ggtitle("Percent of observations with experienced lawyers") +
  theme_minimal() +
  xlab("") +
  scale_x_discrete(labels=c("First-time lawyers", "Repeat lawyers"))
```

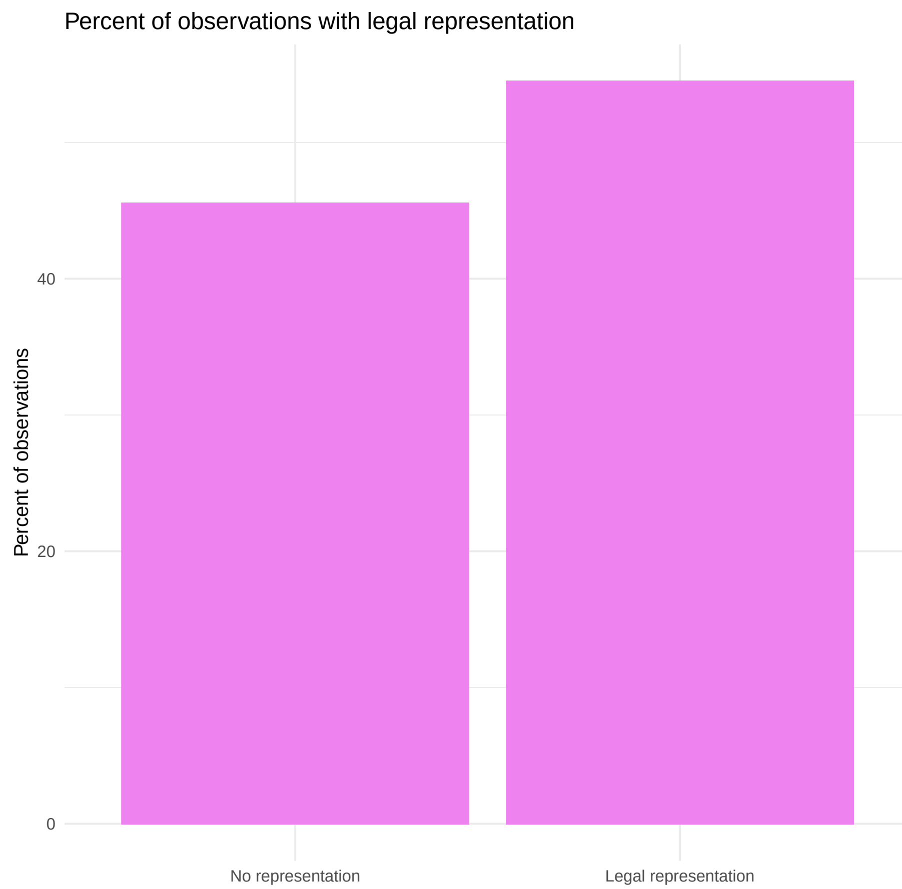


```
experience <- table(data$experience)
prop.table(experience)*100

##
##        0        1
## 77.18412 22.81588

rm(experience)

prop.table(table(experience$experience))*100
```

4.2 Lawyer

This variable is at the level of the party in the dispute.

```
# Who does not have lawyers?
no_lawyer <- data %>%
  filter(lawyer == 0)

prop.table(table(no_lawyer$type))*100

##
##             Other           Company    EU institution        Individual
##        9.91276764       15.30531324        0.07930214       15.54321967
##               NGO State institution
##        1.82394925       57.33544806

## Graphic of representation
data %>%
  group_by(lawyer) %>%
  ggplot(aes(as.factor(lawyer))) +
  geom_bar(position = "dodge", aes(y = (..count..)/sum(..count..)*100), colour="violet", fil
  ylab("Percent of observations") +
  ggtitle("Percent of observations with legal representation")+
  xlab("") +
  theme_minimal() +
  scale_x_discrete(labels=c("No representation", "Legal representation"))
```

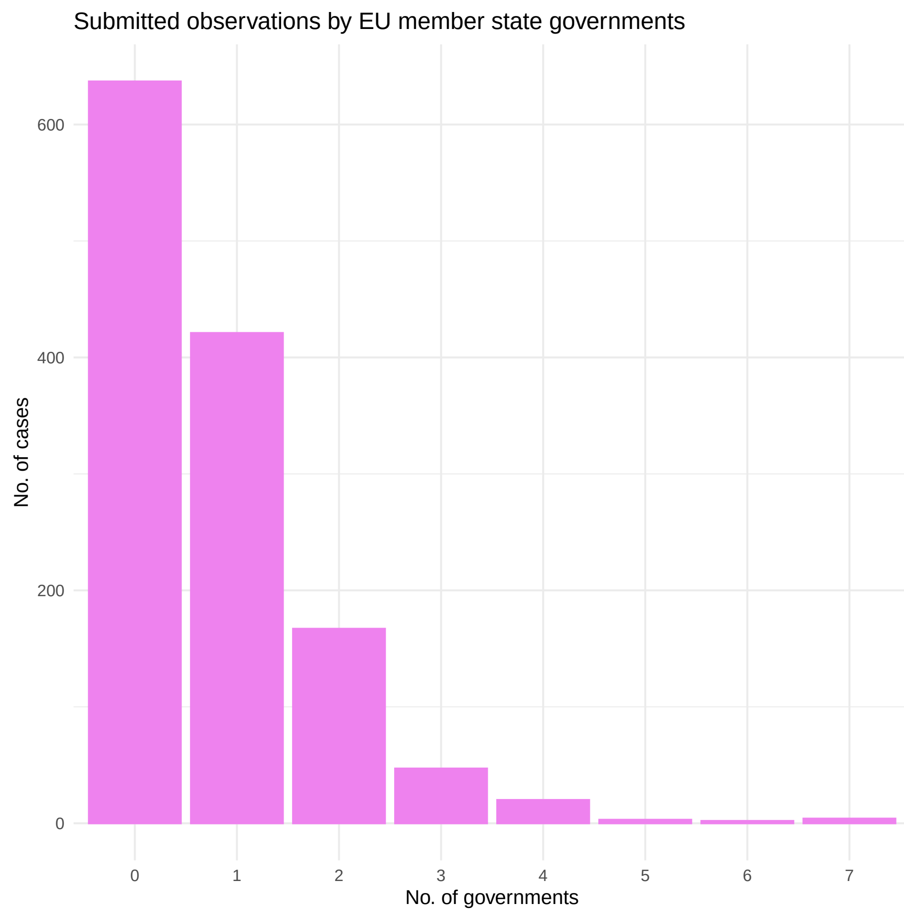



```
lawyer <- table(data$lawyer)
prop.table(lawyer)*100

##
##        0        1
## 45.52347 54.47653

rm(lawyer)
```

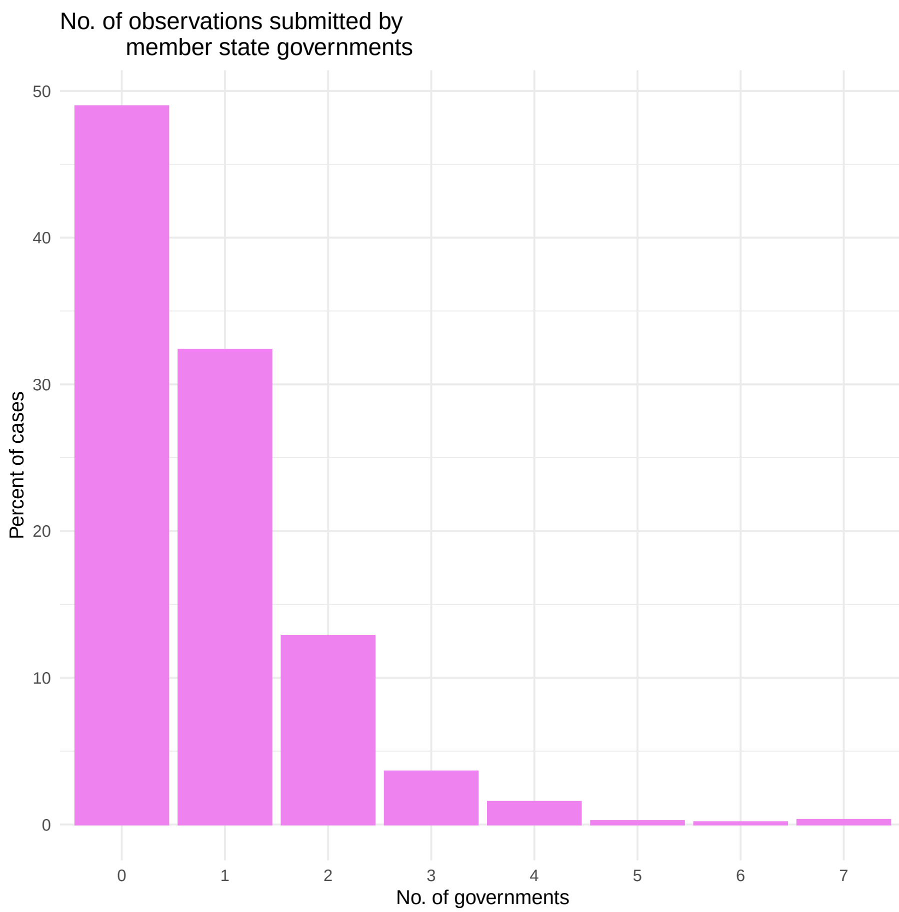
4.3 Salience

Salience is measured as number of EU governments that submit observations in cases referred to the CJEU. I create this variable by adding up the number of governments that submits observations in favor of the applicant and in favor of the defendant. I create a salience measure to seperate between salient and non salient cases. To make this variable I flag all cases that has above the average submitted observations

```
summary(data$salience) # MEAN

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0000  1.0000  0.8245  1.0000  7.0000

length(unique(data$celex)) # Total number of cases

## [1] 1301

table(data$salience)

##
##    0    1    2    3    4    5    6    7
## 1320  902  360  128   42    6    4    8

# Total number of cases
data %>%
  group_by(celex) %>%
  count(salience) %>%
  ggplot(aes(as.factor(salience))) +
  geom_bar(color = "violet", fill = "violet") +
  xlab("No. of governments") +
  ylab("No. of cases") +
  theme_minimal() +
  ggtitle("Submitted observations by EU member state governments")
```




```r
# Percentage of cases
data %>%
  group_by(celex) %>%
  count(salience) %>%
  ggplot(aes(as.factor(salience))) +
  geom_bar(color = "violet", fill = "violet", aes(y = (..count..)/sum(..count..)*100)) +
  xlab("No. of governments") +
  ylab("Percent of cases") +
  theme_minimal() +
  ggtitle("No. of observations submitted by
          member state governments")
```




```
## Binary measure of salience

table(data$binary_salience)

##
##    0    1
## 2222  548

salience <- data %>%
  filter(salience != 0)

mean(salience$salience)

## [1] 1.575172

salience <- salience %>%
  group_by(member_state) %>%
```

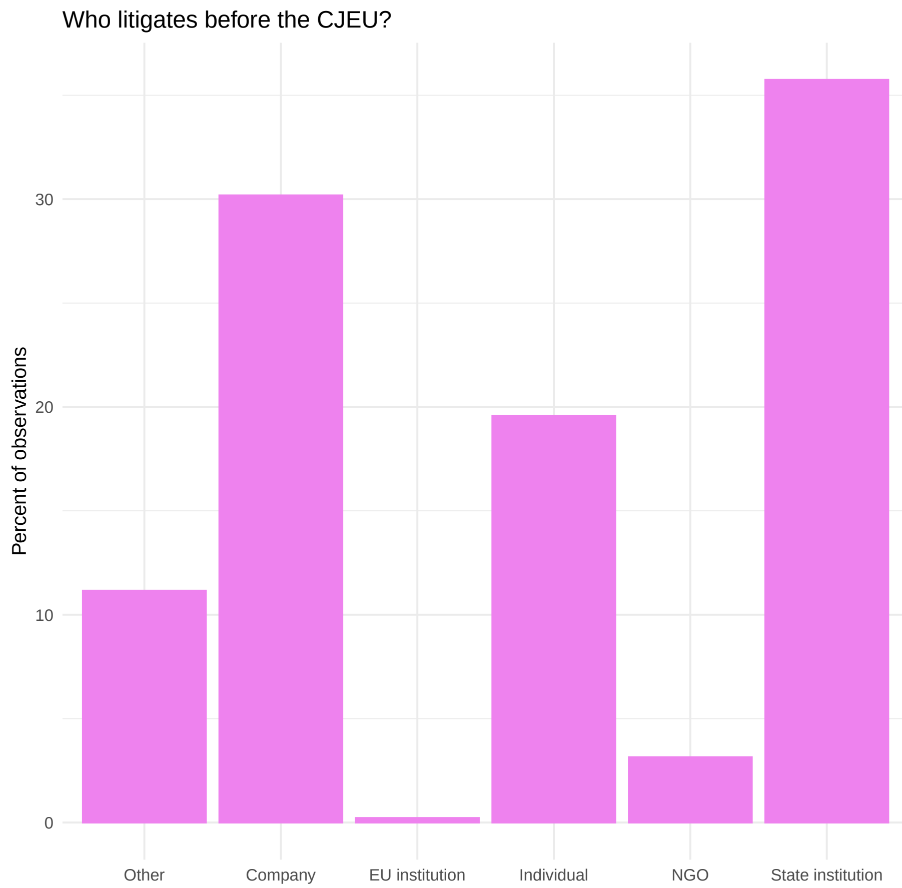
```
  summarise_at(vars(salience),
               list(name = mean))

class(data$member_state)

## [1] "character"

data$time_period_salience <- ifelse(data$member_state == "6", 1.254545, ifelse(data$member_s
              ifelse(data$member_state == "10", 1.614943,
                     ifelse(data$member_state == "12", 1.636095, ifelse(data$member_state))


data$binary_salience <-
  ifelse(data$salience > data$time_period_salience, 1, 0)

table(data$salience)

##
##    0    1    2    3    4    5    6    7
## 1320  902  360  128   42    6    4    8

table(data$binary_salience)

##
##    0    1
## 2222  548

prop.table(table(data$binary_salience))*100

##
##        0        1
## 80.21661 19.78339
```

5 Control variables

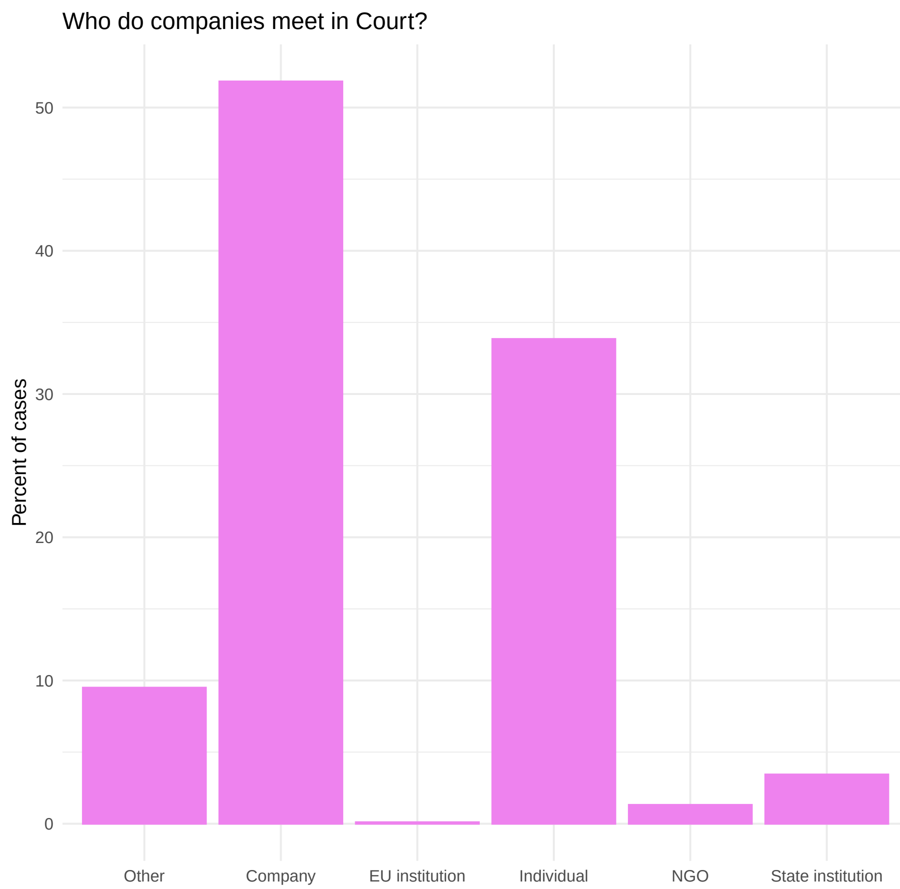
5.1 Type of actor and opponent in court

```
prop.table(table(sample1$type))*100
data %>%
  ggplot(aes(as.factor(type))) +
  geom_bar(color = "violet", fill = "violet", aes(y = (..count..)/sum(..count..)*100)) +
  xlab("") +
  ylab("Percent of observations") +
  theme_minimal() +
```

```
  ggtitle("Who litigates before the CJEU?")
```




```
data %>%
  filter(type == "State institution") %>%
   ggplot(aes(type_opponent)) +
   geom_bar(color = "violet", fill = "violet", aes(y = (..count..)/sum(..count..)*100)) +
  xlab("") +
  ylab("Percent of cases") +
  theme_minimal() +
  ggtitle("Who do companies meet in Court?")
```




```
prop.table(table(data$type))*100

##
##            Other          Company   EU institution       Individual
##       11.1552347       30.1805054        0.2166065       19.5667870
##              NGO State institution
##        3.1407942       35.7400722

prop.table(table(data$type_opponent))*100

##
##            Other          Company   EU institution       Individual
##       11.1552347       30.1805054        0.2166065       19.5667870
##              NGO State institution
##        3.1407942       35.7400722
```

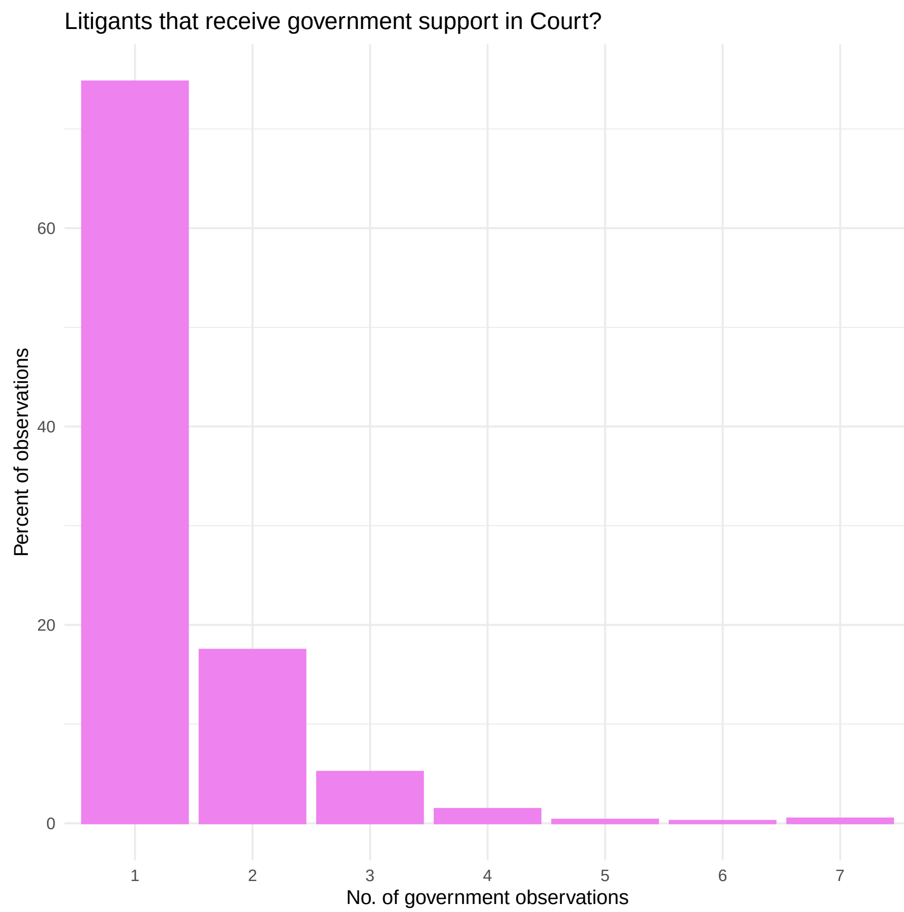
5.2 Role in legal dispute

```
data %>%
  select(celex, role) %>%
  group_by(celex) %>%
  ggplot(aes(as.factor(role))) +
  geom_bar(color = "violet", fill = "violet", aes(y = (..count..)/sum(..count..)*100)) +
  xlab("") +
  ylab("Percent of cases") +
  theme_minimal() +
  ggtitle("Who litigates before the CJEU?")
```

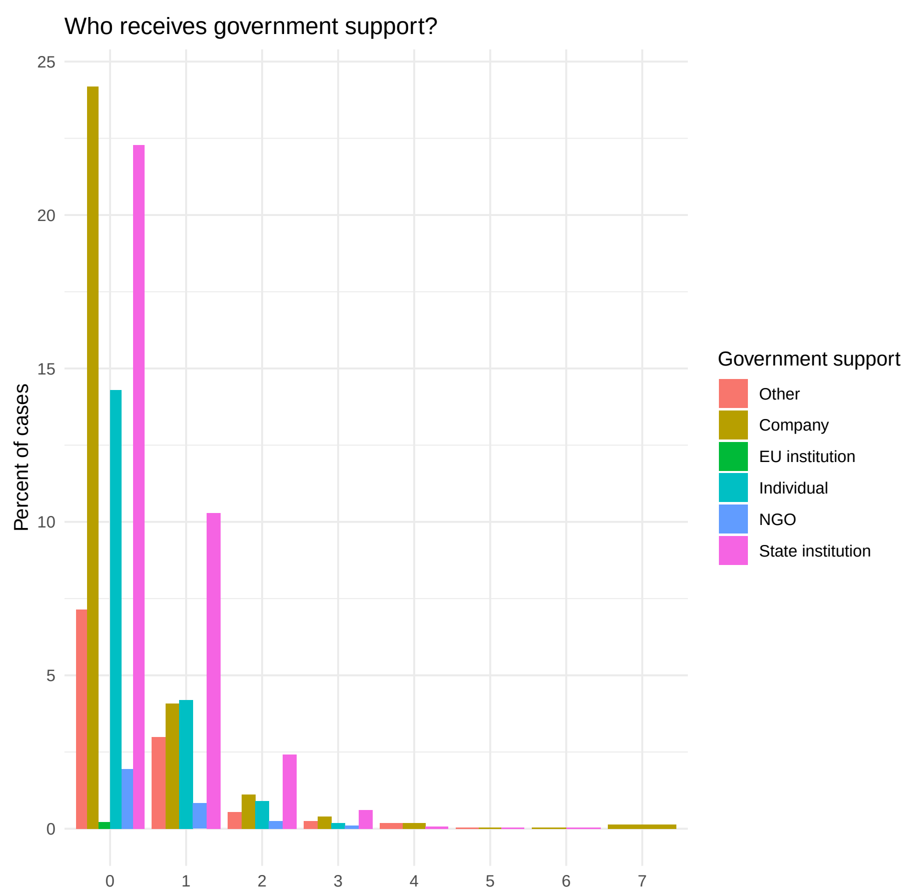
5.3 Government support

This variable measures whether or not the applicant or defendant had government support. This is measured has a binary indicator.

```
table(data$government_support)

##
##    0    1    2    3    4    5    6    7
## 1941  620  145   43   12    3    2    4

prop.table(table(data$government_support))*100

##
##           0           1           2           3           4           5
## 70.07220217 22.38267148  5.23465704  1.55234657  0.43321300  0.10830325
##           6           7
##  0.07220217  0.14440433

# Who has government support in Court?


data %>%
  filter(government_support > 0) %>%
  ggplot(aes(as.factor(government_support))) +
  geom_bar(position = "dodge", aes(y = (..count..)/sum(..count..)*100), colour = "violet", f
  ylab("Percent of observations") +
  ggtitle("Litigants that receive government support in Court?")+
  xlab("No. of government observations") +
  theme_minimal()
```

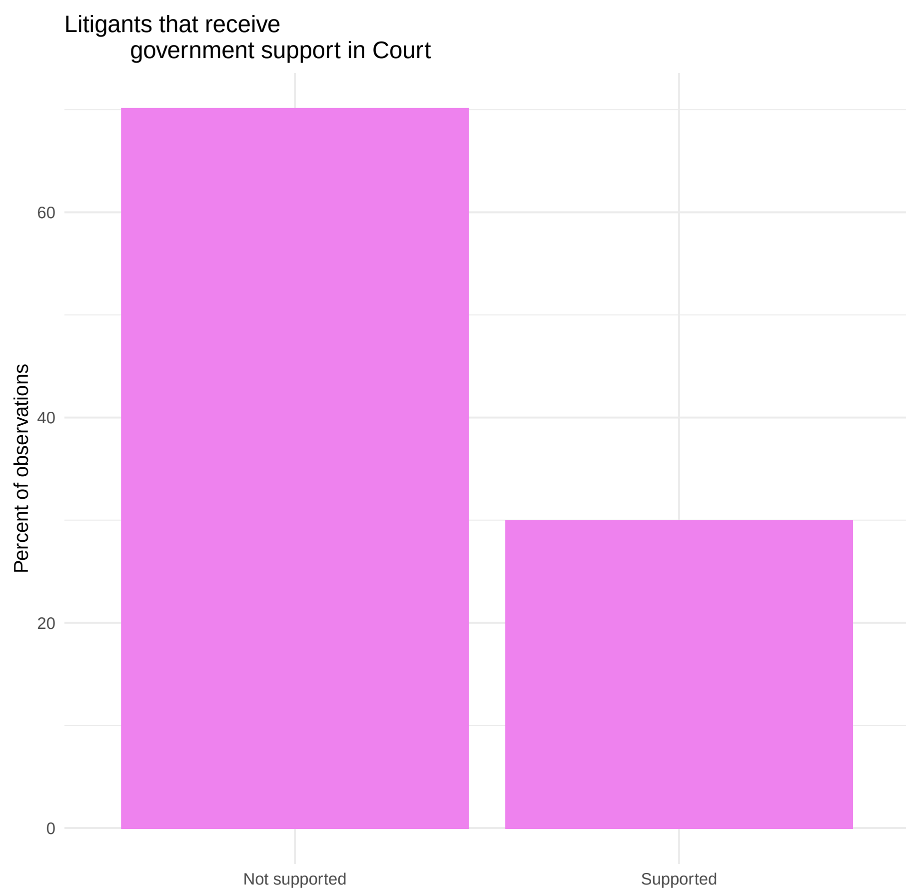



```
data %>%
  ggplot(aes(as.factor(government_support), fill = as.factor(type))) +
  geom_bar(position = "dodge", aes(y = (..count..)/sum(..count..)*100)) +
  ylab("Percent of cases") +
  labs(fill = "Government support")  +
  ggtitle("Who receives government support?")+
  xlab("")+
  theme_minimal()
```




```
data$government_support_binary <-
  ifelse(data$government_support > 0, 1, 0)

table(data$government_support)

##
##    0    1    2    3    4    5    6    7
## 1941  620  145   43   12    3    2    4

table(data$government_support_binary)

##
##    0    1
## 1941  829

data %>%
  ggplot(aes(as.factor(government_support_binary))) +
```

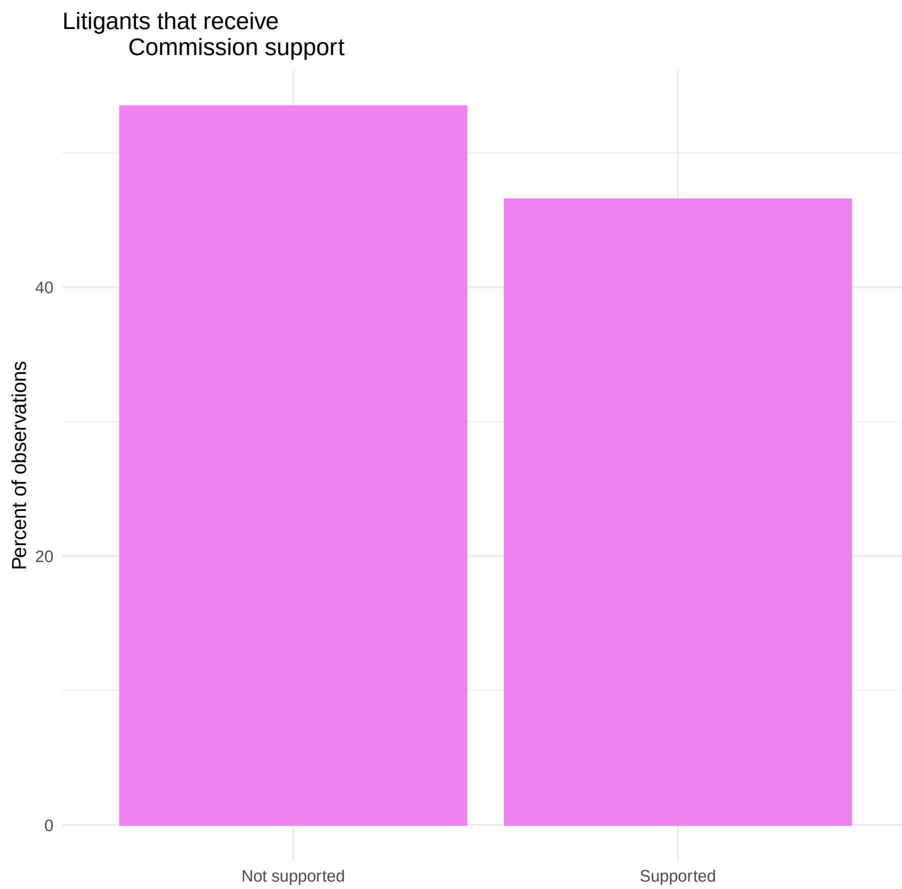
```
  geom_bar(position = "dodge", aes(y = (..count..)/sum(..count..)*100), colour = "violet", f
  ylab("Percent of observations") +
  ggtitle("Litigants that receive
          government support in Court")+
  theme_minimal() +
  xlab("")+
  scale_x_discrete(labels=c("Not supported", "Supported"))
```



5.4 Commission support

This variable indicates whether or not the applicant or defendant in the legal dispute has support from the Commission. The variable is binary. Support from the Commission is flagged as 1.

```
table(data$commission_support)

##
##    0    1
## 1481 1289

prop.table(table(data$commission_support))*100

##
##       0       1
## 53.4657 46.5343

# Who has commission support in Court?


data %>%
  ggplot(aes(as.factor(commission_support))) +
  geom_bar(position = "dodge", aes(y = (..count..)/sum(..count..)*100), colour= "violet", fi
  ylab("Percent of observations") +
  ggtitle("Litigants that receive
          Commission support")+
  xlab("")+
  scale_x_discrete(labels = c("Not supported", "Supported"))+
  theme_minimal()
```



5.5 Number of member states

```
table(data$member_state)

##
##   10   12    6    9
##  658 1146  240  726
```

6 Descriptive statistics

This table gives an overview of the numeric variables in the data.

```r
stats_data <- data.frame(data)
names(stats_data)
```

```
##  [1] "celex"                   "role"
##  [3] "type"                    "experience"
##  [5] "win"                     "type_opponent"
##  [7] "salience"                "lawyer"
##  [9] "binary_salience"         "government_support"
## [11] "commission_support"      "member_state"
## [13] "ecjplaintiffagree"       "year"
## [15] "time_period_salience"    "government_support_binary"
```

```r
stats_data <- stats_data %>%
  select(4:5, 7:11, 15)
stargazer(stats_data, type = "text")
```

```
##
## ============================================================================
## Statistic                   N   Mean  St. Dev.  Min  Pctl(25) Pctl(75)  Max
## ----------------------------------------------------------------------------
## experience                2,770 0.228  0.420     0      0        0       1
## win                       2,770 0.500  0.500     0      0        1       1
## salience                  2,770 0.825  1.035     0      0        1       7
## lawyer                    2,770 0.545  0.498     0      0        1       1
## binary_salience           2,770 0.198  0.398     0      0        0       1
## government_support        2,770 0.412  0.772     0      0        1       7
## commission_support        2,770 0.465  0.499     0      0        1       1
## ecjplaintiffagree         2,770 0.471  0.499     0      0        1       1
## time_period_salience      2,770 1.566  0.108   1.255  1.513    1.636   1.636
## government_support_binary 2,770 0.299  0.458     0      0        1       1
## ----------------------------------------------------------------------------
```

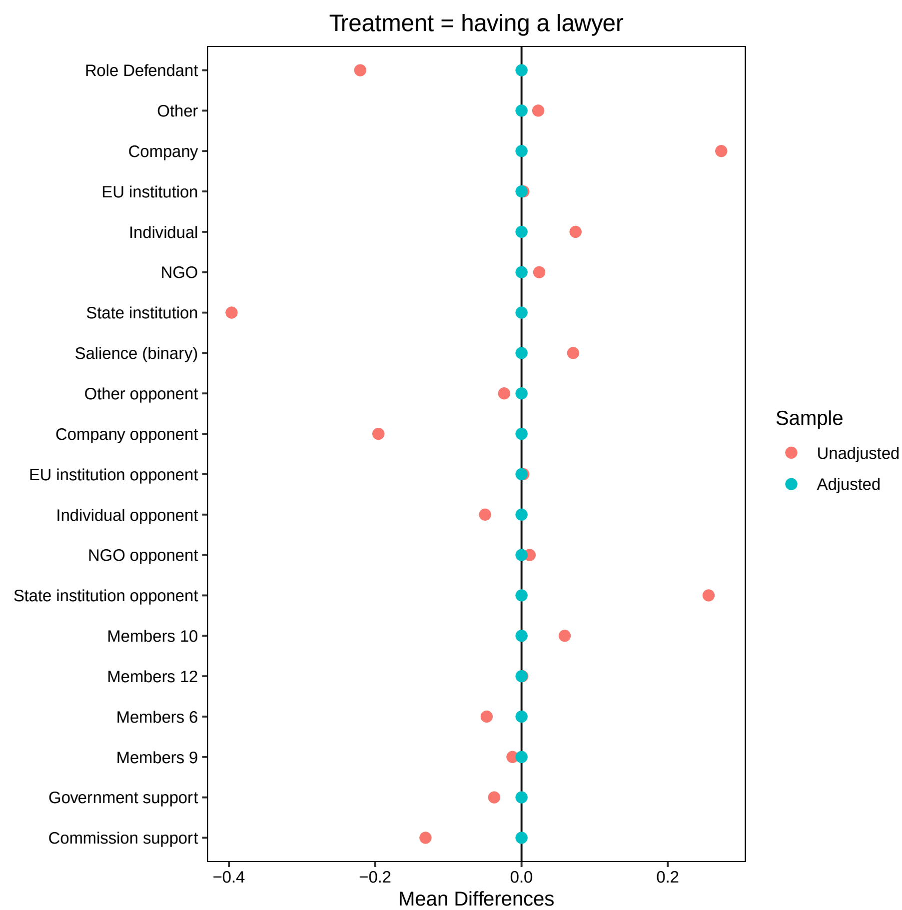
```r
rm(stats_data)
```

7 Treatment-is-lawyer sample

7.1 Matching

```r
# Treatment = having a lawyer
sample_1 <- matchit(lawyer ~ role +
                      type + binary_salience +
                         type_opponent
                    + member_state
                    + government_support
```

```
                          + commission_support,
                            method = "cem",
                            estimand = "ATT",
                            data = data)


sample1 <- match.data(sample_1, data=data)

love.plot(sample_1,
          title = "Treatment = having a lawyer", var.names = var_names)
```




```
### Add more plots

### Review the sample

length(unique(sample1$celex))
```

```
## [1] 1118

prop.table(table(sample1$ecjplaintiffagree))

##
##         0         1
## 0.5436987 0.4563013

table(sample1$type)

##
##             Other           Company    EU institution        Individual
##               201               619                 0               398
##               NGO State institution
##                22               831

sample1$type <- relevel(sample1$type, ref = "Individual")

sample1$type_opponent <- relevel(sample1$type_opponent, ref = "Individual")
```

7.2 Analysis

```
## Estimating a model without interaction
law1 <- glm(win~ lawyer + role + type +
            type_opponent + member_state +
            binary_salience
           + government_support
           + commission_support,
          family = binomial(link = "logit"),
          data = sample1)

beta_law1 <- law1$coefficients

tab_law1 <- (exp(beta_law1)-1)*100
tab_law1

##                   (Intercept)                        lawyer
##                    -73.682257                      4.199343
##                 roledefendant                     typeOther
##                      4.632119                    -18.077924
##                   typeCompany                       typeNGO
##                    -15.755639                    -20.661474
##          typeState institution            type_opponentOther
##                    -36.170960                    -11.711833
```

```
##          type_opponentCompany                   type_opponentNGO
##                     21.721787                           3.959007
## type_opponentState institution                    member_state12
##                     16.843448                          15.103636
##                  member_state6                      member_state9
##                    -18.987031                          -4.043076
##                binary_salience                government_support
##                    -42.906321                          63.778545
##             commission_support
##                   1657.063247
```

```
vcov_law1 <- vcovHC(law1, "HC1")

## Estimating a model with controls + interaction

law <- glm(win~ lawyer + role + type +
            type_opponent + member_state +
            binary_salience
           + lawyer*binary_salience
           + government_support
           + commission_support,
          family = binomial(link = "logit"),
          data = sample1)



beta_law <- law$coefficients

tab_law <- (exp(beta_law)-1)*100
tab_law
```

```
##                    (Intercept)                            lawyer
##                    -74.977101                          13.151103
##                 roledefendant                          typeOther
##                      4.031611                         -19.286908
##                    typeCompany                           typeNGO
##                    -17.093649                         -20.063564
##          typeState institution                type_opponentOther
##                    -36.144294                          -9.980476
##          type_opponentCompany                   type_opponentNGO
##                     24.290846                           4.532891
## type_opponentState institution                    member_state12
##                     16.916469                          15.763789
##                  member_state6                      member_state9
##                    -18.474206                          -3.563876
##                binary_salience                government_support
```

```
##                   -16.644408                        62.304878
##            commission_support        lawyer:binary_salience
##                  1677.572955                       -53.252588

vcov_law <- vcovHC(law, "HC1")

stargazer(law1, law, type = "text", se = list(vcov_law1, vcov_law))

##
## ===========================================================
##                                    Dependent variable:
##                                ----------------------------
##                                            win
##                                     (1)            (2)
## -----------------------------------------------------------
## lawyer                            0.041***      0.124***
##                                  (-0.010)      (-0.011)
##
## roledefendant                     0.045***      0.040***
##                                  (-0.011)      (-0.011)
##
## typeOther                        -0.199***     -0.214***
##                                  (-0.024)      (-0.024)
##
## typeCompany                      -0.171***     -0.187***
##                                  (-0.022)      (-0.021)
##
## typeNGO                          -0.231***     -0.224***
##                                  (-0.063)      (-0.063)
##
## typeState institution            -0.449***     -0.449***
##                                  (-0.041)      (-0.041)
##
## type_opponentOther               -0.125***     -0.105***
##                                  (-0.027)      (-0.028)
##
## type_opponentCompany              0.197***      0.217***
##                                  (-0.022)      (-0.023)
##
## type_opponentNGO                   0.039         0.044
##                                  (-0.048)      (-0.048)
##
## type_opponentState institution    0.156***      0.156***
##                                  (-0.046)      (-0.046)
##
## member_state12                    0.141***      0.146***
```

```
##                                     (-0.015)      (-0.014)
##
## member_state6                       -0.211***     -0.204***
##                                     (-0.011)      (-0.011)
##
## member_state9                       -0.041***     -0.036***
##                                     (-0.014)      (-0.014)
##
## binary_salience                     -0.560***     -0.182***
##                                     (-0.009)      (-0.016)
##
## government_support                   0.493***      0.484***
##                                     (-0.002)      (-0.002)
##
## commission_support                   2.866***      2.878***
##                                     (-0.005)      (-0.005)
##
## lawyer:binary_salience                            -0.760***
##                                                    (0.012)
##
## Constant                            -1.335***     -1.385***
##                                     (0.084)        (0.084)
##
## -----------------------------------------------------------
## Observations                          2,071         2,071
## Log Likelihood                      -1,004.453    -1,002.523
## Akaike Inf. Crit.                   2,042.905      2,041.047
## ===========================================================
## Note:                             *p<0.1; **p<0.05; ***p<0.01
```

```
stargazer(law1, law, type = "latex", se = list(vcov_law1, vcov_law), style = "all2",
          single.row = TRUE, no.space = TRUE, font.size = "scriptsize", align = TRUE,
          dep.var.caption = "Binomial logistic regression",
          dep.var.labels = "Favorable ruling",
         keep = c("lawyer", "role",
                  "type", "type_opponent",
                   "binary_salience",
                  "government_support",
                  "commission_support", "Constant"),
         covariate.labels = c("Lawyer",
                              "Defendant",
"State institution",
"Company",
"Other",
"NGO",
  "Opponent Company",
```

```
"Opponent Other",

 "Opponent NGO",

 "Opponent State Institution",
                              "Salience",
                              "Government support",
                              "Commission support",
                              "Interaction lawyer * salience"), flip = TRUE)
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac
## % Date and time: Wed, May 26, 2021 - 13:33:22
## % Requires LaTeX packages: dcolumn
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \scriptsize
## \begin{tabular}{@{\extracolsep{5pt}}lD{.}{.}{-3} D{.}{.}{-3} }
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{2}{c}{Binomial logistic regression} \\
## \cline{2-3}
## \\[-1.8ex] & \multicolumn{2}{c}{Favorable ruling} \\
## \\[-1.8ex] & \multicolumn{1}{c}{(1)} & \multicolumn{1}{c}{(2)}\\
## \hline \\[-1.8ex]
##  Lawyer & 0.041^{***}$ $(-0.010) & 0.124^{***}$ $(-0.011) \\
##   Defendant & 0.045^{***}$ $(-0.011) & 0.040^{***}$ $(-0.011) \\
##   State institution & -0.199^{***}$ $(-0.024) & -0.214^{***}$ $(-0.024) \\
##   Company & -0.171^{***}$ $(-0.022) & -0.187^{***}$ $(-0.021) \\
##   Other & -0.231^{***}$ $(-0.063) & -0.224^{***}$ $(-0.063) \\
##   NGO & -0.449^{***}$ $(-0.041) & -0.449^{***}$ $(-0.041) \\
##   Opponent Company & -0.125^{***}$ $(-0.027) & -0.105^{***}$ $(-0.028) \\
##   Opponent Other & 0.197^{***}$ $(-0.022) & 0.217^{***}$ $(-0.023) \\
##   Opponent NGO & 0.039$ $(-0.048) & 0.044$ $(-0.048) \\
##   Opponent State Institution & 0.156^{***}$ $(-0.046) & 0.156^{***}$ $(-0.046) \\
##   Salience & -0.560^{***}$ $(-0.009) & -0.182^{***}$ $(-0.016) \\
##   Government support & 0.493^{***}$ $(-0.002) & 0.484^{***}$ $(-0.002) \\
##   Commission support & 2.866^{***}$ $(-0.005) & 2.878^{***}$ $(-0.005) \\
##   Interaction lawyer * salience &  & -0.760^{***}$ $(0.012) \\
##   Constant & -1.335^{***}$ $(0.084) & -1.385^{***}$ $(0.084) \\
##  \hline \\[-1.8ex]
## Observations & \multicolumn{1}{c}{2,071} & \multicolumn{1}{c}{2,071} \\
## Log Likelihood & \multicolumn{1}{c}{-1,004.453} & \multicolumn{1}{c}{-1,002.523} \\
## Akaike Inf. Crit. & \multicolumn{1}{c}{2,042.905} & \multicolumn{1}{c}{2,041.047} \\
## Residual Deviance & \multicolumn{1}{c}{2,008.905 (df = 2054)} & \multicolumn{1}{c}{2,005.
```

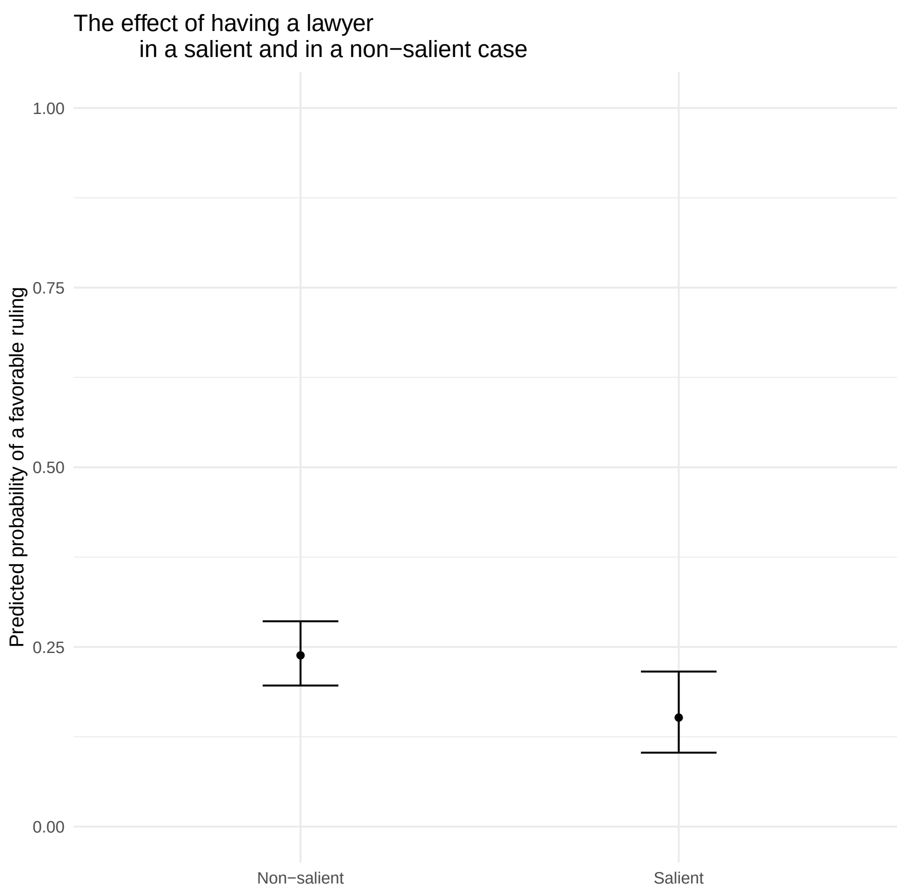
```
## Null Deviance (df = 2070) & \multicolumn{1}{c}{2,869.760} & \multicolumn{1}{c}{2,869.760}
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{2}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \\
## \end{tabular}
## \end{table}
```

7.3 Plotting effects

```
law$vcov_law <- vcov_law

law1$vcov_law <- vcov_law1
# Setting seed
set.seed(24)

simBetas <- mvrnorm(n = 1000,
                    mu = coefficients(law1),
                    Sigma = law1$vcov_law)


names(coefficients(law1))

##  [1] "(Intercept)"                     "lawyer"
##  [3] "roledefendant"                   "typeOther"
##  [5] "typeCompany"                     "typeNGO"
##  [7] "typeState institution"           "type_opponentOther"
##  [9] "type_opponentCompany"            "type_opponentNGO"
## [11] "type_opponentState institution" "member_state12"
## [13] "member_state6"                   "member_state9"
## [15] "binary_salience"                 "government_support"
## [17] "commission_support"

xMatrix <- cbind(1, #the intercept
                 1, # Party has a lawyer
                 0, # defendant
                 0, # other
                 1, # Company
                 0, # NGO
                 0, # SI
                 0, # opponent other
                 0, # opponent compnay
                 0, # opponent NGO
                 1, # opponent state institution
                 1, # M 12
```

```r
                 0, # M 6
                 0, # M 9,
                 c(0, 1), # Salience
                 0, # Government support,
                 0 # Commission support
                  # interaction
                 )




ncol(simBetas) == ncol(xMatrix) #yay!!

## [1] TRUE

### Calculating predicted probabilities: Her multipliserer du simuleringen med xmatrisen
xBetaMatrix <- xMatrix %*% t(simBetas ) ## this just means x times the betas
predProbs <- 1/(1+exp(-xBetaMatrix)) #This is the predicted probability, for another type of


### Getting point estimates and confidence intervals:
quantileValues <- apply(X = predProbs,  ## read up on the apply() family of functions!
                        MARGIN = 1, ## this means we are applying a function to all the rows
                        FUN =  ## The fun argument defines what I want to do with all the ro
                          ## What we want to do here is to use quantile to get the quantiles
                          ### bounds of the confidence intervals and our point estimates:
                          quantile, probs = c(.05,.5,.95))
quantileValues <- as.data.frame(t(quantileValues))

plotPoints <- cbind(c("Non-salient", "Salient"),quantileValues)
plotPoints

##   c("Non-salient", "Salient")        5%       50%       95%
## 1                 Non-salient 0.1962999 0.2382884 0.2857224
## 2                     Salient 0.1028691 0.1516874 0.2157412

colnames(plotPoints) <- c("Salient", "lower", "estimate", "upper")




ggplot(plotPoints,
       aes(x = Salient,
           y = estimate,
           ymin = lower,
           ymax = upper)) +
  geom_errorbar(width =.2)+
```

```r
  geom_point()+
  ylim(0,1)+
  ylab("Predicted probability of a favorable ruling")+
  xlab("")+
  theme_minimal() +
  ggtitle("The effect of having a lawyer
          in a salient and in a non-salient case")
```

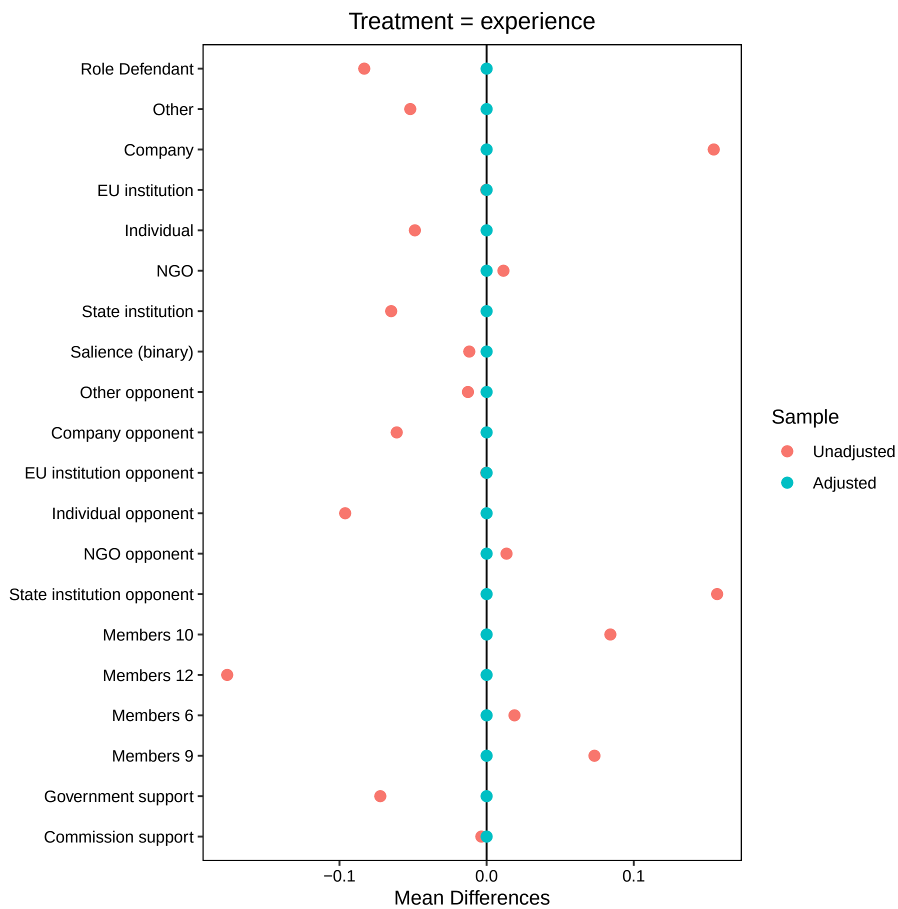


The figure shows the different effect lawyers on the decision-making of the CJEU. The two different estimates shows the difference in the effect of lawyers in cases that are salient and non-salient. The error bars illustrate the predicted probability of a favorable ruling for a company (applicant) with legal representation litigating against a state institution (defendant), without government or Commission support. Having legal representation in preliminary reference cases before the CJEU increases the likelihood of a favorable ruling. The effect of the legal representation on the decision-outcome is higher in non-salient cases, than in salient cases.

8 Treatment-is-experience sample

8.1 Matching

```
# Filtering out a sample consisting of parties with only lawyers
df <- data %>%
  filter(lawyer == 1) %>%
  ungroup()


sample_2 <- matchit(experience ~ role + type
                    + binary_salience +
                         type_opponent
                    + member_state +
                    government_support
                    + commission_support,
                       method = "cem",
                       estimand = "ATT",
                       data = df)


sample2 <- match.data(sample_2, data=df)


# Checking balance

love.plot(sample_2, var.names = var_names, title = "Treatment = experience")
```




```
# Add more balance plots

length(unique(sample2$celex))

## [1] 803

table(sample2$role)

##
## applicant defendant
##       646       344

prop.table(table(sample2$ecjplaintiffagree))

##
##         0         1
## 0.5181818 0.4818182
```

```r
prop.table(table(sample2$experience))*100

##
##        0        1
## 51.81818 48.18182

sample2$type <- relevel(sample2$type, ref = "Individual")

sample2$type_opponent <- relevel(sample2$type_opponent, ref = "Individual")

prop.table(table(sample2$type))*100

##
##        Individual             Other           Company    EU institution
##         21.919192         10.303030         50.404040          0.000000
##               NGO State institution
##          1.111111         16.262626

table(sample2$type_opponent)

##
##        Individual             Other           Company    EU institution
##               118                92               165                 0
##               NGO State institution
##                12               603

prop.table(table(sample2$ecjplaintiffagree))*100

##
##        0        1
## 51.81818 48.18182
```

8.2 Analysis

```r
## Estimating a model with controls
m1 <- glm(win~ experience + role + type +
            type_opponent + member_state +
            binary_salience +
          + government_support
          + commission_support,
          family = binomial(link = "logit"),
          data = sample2)

beta1 <- m1$coefficients
```

```r
tab1 <- (exp(beta1)-1)*100
tab1

##                  (Intercept)                   experience
##                   -79.095605                   -18.603892
##                roledefendant                    typeOther
##                    -2.347624                     2.907999
##                  typeCompany                      typeNGO
##                     3.280564                   134.537811
##          typeState institution            type_opponentOther
##                   -17.919871                   -12.922009
##          type_opponentCompany              type_opponentNGO
##                   -10.986832                    12.017837
## type_opponentState institution               member_state12
##                    41.225159                    21.563426
##                member_state6                 member_state9
##                    -3.700315                    -3.858882
##              binary_salience            government_support
##                   -63.875567                   115.759881
##           commission_support
##                  2603.070919

vcov_m1 <- vcovHC(m1, "HC1")

# Model with interaction term
m2 <- glm(win~ experience + role + type +
            type_opponent + member_state +
            binary_salience +
            experience*binary_salience
          + government_support
          + commission_support,
          family = binomial(link = "logit"),
          data = sample2)

beta2 <- m2$coefficients

tab2 <- (exp(beta2)-1)*100
tab2

##                  (Intercept)                   experience
##                   -78.741692                   -23.597476
##                roledefendant                    typeOther
##                    -2.611224                     3.177026
##                  typeCompany                      typeNGO
##                     4.706000                   139.065866
##          typeState institution            type_opponentOther
```

```
##                     -17.392524                        -12.795361
##            type_opponentCompany                 type_opponentNGO
##                     -11.296000                         11.976197
## type_opponentState institution                   member_state12
##                      41.381746                         21.870521
##                   member_state6                     member_state9
##                      -2.469659                         -3.967865
##                 binary_salience               government_support
##                     -73.314642                        114.092940
##              commission_support       experience:binary_salience
##                    2638.272266                         85.551667
```

```r
vcov_m2 <- vcovHC(m2, "HC1")

stargazer(m1, m2, type = "text", se = list(vcov_m1, vcov_m2))
```

```
##
## ===========================================================
##                                    Dependent variable:
##                                ----------------------------
##                                            win
##                                     (1)            (2)
## -----------------------------------------------------------
## experience                       -0.206***      -0.269***
##                                   (-0.013)       (-0.014)
##
## roledefendant                      -0.024         -0.026
##                                   (-0.034)       (-0.035)
##
## typeOther                           0.029          0.031
##                                   (-0.081)       (-0.081)
##
## typeCompany                         0.032          0.046
##                                   (-0.061)       (-0.062)
##
## typeNGO                           0.852***       0.872***
##                                   (-0.153)       (-0.154)
##
## typeState institution              -0.197         -0.191
##                                   (-0.133)       (-0.133)
##
## type_opponentOther                 -0.138         -0.137
##                                   (-0.133)       (-0.133)
##
## type_opponentCompany               -0.116         -0.120
##                                   (-0.102)       (-0.102)
```

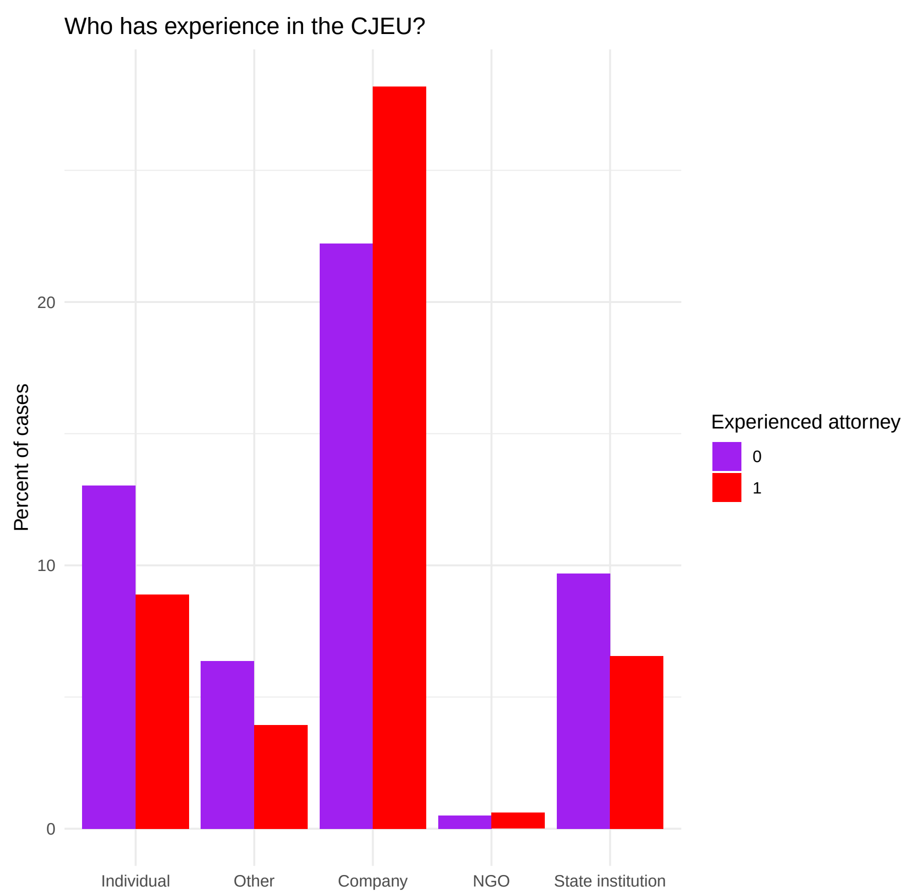
```
##
## type_opponentNGO                       0.113          0.113
##                                       (-0.150)       (-0.151)
##
## type_opponentState institution        0.345**        0.346**
##                                       (-0.163)       (-0.163)
##
## member_state12                        0.195***       0.198***
##                                       (-0.026)       (-0.027)
##
## member_state6                        -0.038***      -0.025***
##                                       (-0.009)       (-0.010)
##
## member_state9                         -0.039         -0.040
##                                       (-0.028)       (-0.029)
##
## binary_salience                      -1.018***      -1.321***
##                                       (-0.017)       (-0.024)
##
## government_support                    0.769***       0.761***
##                                       (-0.018)       (-0.020)
##
## commission_support                    3.297***       3.310***
##                                       (-0.006)       (-0.006)
##
## experience:binary_salience                           0.618***
##                                                      (0.015)
##
## Constant                             -1.565***      -1.548***
##                                       (0.249)        (0.252)
##
## -----------------------------------------------------------
## Observations                            990            990
## Log Likelihood                        -445.946       -445.365
## Akaike Inf. Crit.                     925.892        926.730
## ===========================================================
## Note:                               *p<0.1; **p<0.05; ***p<0.01
```

```
stargazer(m1, m2, type = "latex",
          se = list(vcov_m1, vcov_m2),
          style = "all2",
          single.row = TRUE,
          no.space = TRUE,
          font.size = "scriptsize",
          align = TRUE,
          dep.var.caption =
```

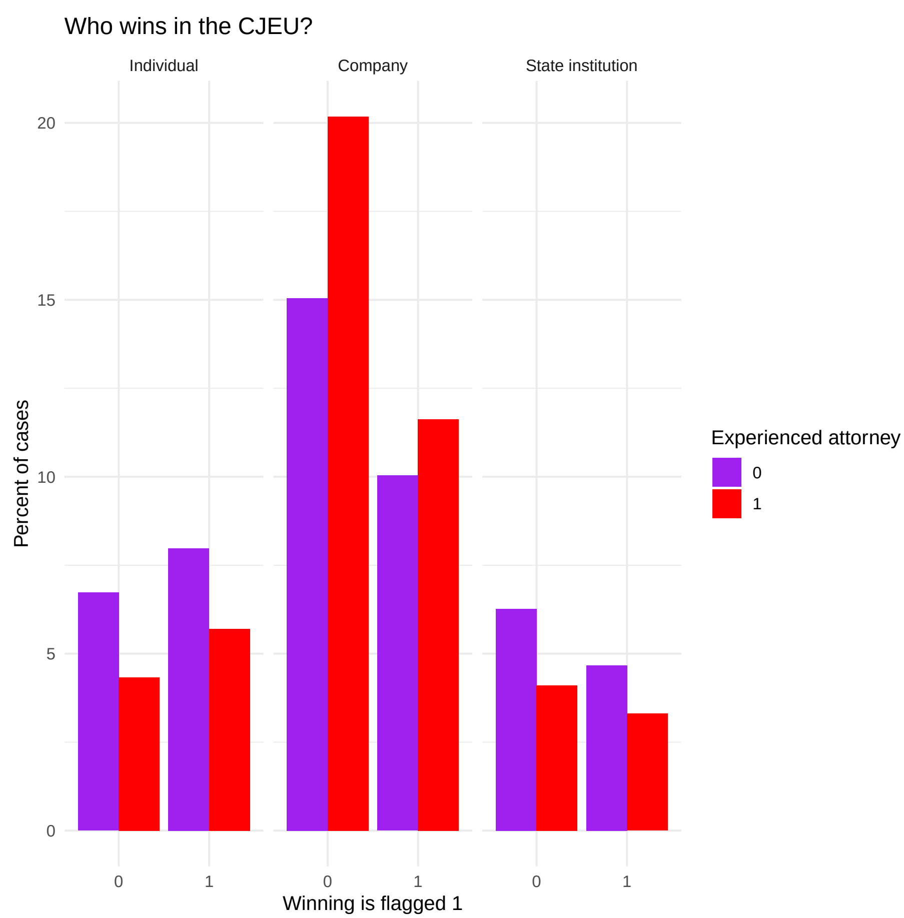
```
            "Binomial logistic regression",
          dep.var.labels = "Favorable ruling",
         keep = c("Constant","experience", "role",
                  "type",
                  "type_opponent",
                   "binary_salience",
                  "government_support",
                  "commission_support"),
         covariate.labels = c("Experience",
                              "Defendant",
                              "Other",
                              "Company", "NGO",
                              "State institution",
                              "Opponent Other",
                              "Opponent Company",
                              "Opponent NGO",
                              "Opponent State institution",
                              "Salience",
                              "Government support",
                              "Commission support",
                              "Interaction experience * salience"
                  ))
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac
## % Date and time: Wed, May 26, 2021 - 13:33:24
## % Requires LaTeX packages: dcolumn
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \scriptsize
## \begin{tabular}{@{\extracolsep{5pt}}lD{.}{.}{-3} D{.}{.}{-3} }
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{2}{c}{Binomial logistic regression} \\
## \cline{2-3}
## \\[-1.8ex] & \multicolumn{2}{c}{Favorable ruling} \\
## \\[-1.8ex] & \multicolumn{1}{c}{(1)} & \multicolumn{1}{c}{(2)}\\
## \hline \\[-1.8ex]
##  Experience & -0.206^{***}$ $(-0.013) & -0.269^{***}$ $(-0.014) \\
##   Defendant & -0.024$ $(-0.034) & -0.026$ $(-0.035) \\
##   Other & 0.029$ $(-0.081) & 0.031$ $(-0.081) \\
##   Company & 0.032$ $(-0.061) & 0.046$ $(-0.062) \\
##   NGO & 0.852^{***}$ $(-0.153) & 0.872^{***}$ $(-0.154) \\
##   State institution & -0.197$ $(-0.133) & -0.191$ $(-0.133) \\
##   Opponent Other & -0.138$ $(-0.133) & -0.137$ $(-0.133) \\
```

```
##   Opponent Company & -0.116$ $(-0.102) & -0.120$ $(-0.102) \\
##   Opponent NGO & 0.113$ $(-0.150) & 0.113$ $(-0.151) \\
##   Opponent State institution & 0.345^{**}$ $(-0.163) & 0.346^{**}$ $(-0.163) \\
##   Salience & -1.018^{***}$ $(-0.017) & -1.321^{***}$ $(-0.024) \\
##   Government support & 0.769^{***}$ $(-0.018) & 0.761^{***}$ $(-0.020) \\
##   Commission support & 3.297^{***}$ $(-0.006) & 3.310^{***}$ $(-0.006) \\
##   Interaction experience * salience &  & 0.618^{***}$ $(0.015) \\
##   Constant & -1.565^{***}$ $(0.249) & -1.548^{***}$ $(0.252) \\
##  \hline \\[-1.8ex]
## Observations & \multicolumn{1}{c}{990} & \multicolumn{1}{c}{990} \\
## Log Likelihood & \multicolumn{1}{c}{-445.946} & \multicolumn{1}{c}{-445.365} \\
## Akaike Inf. Crit. & \multicolumn{1}{c}{925.892} & \multicolumn{1}{c}{926.730} \\
## Residual Deviance & \multicolumn{1}{c}{891.892 (df = 973)} & \multicolumn{1}{c}{890.730
## Null Deviance (df = 989) & \multicolumn{1}{c}{1,358.808} & \multicolumn{1}{c}{1,358.808}
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{2}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \
## \end{tabular}
## \end{table}
```

```
sample2 %>%
  group_by(celex) %>%
  ggplot(aes(type, fill = as.factor(experience))) +
  geom_bar(aes(y = (..count..)/sum(..count..)*100), position = "dodge") +
  ylab("Percent of cases") +
  labs(fill = "Experienced attorney") +
  scale_fill_manual(values = c( "purple", "red")) +
  ggtitle("Who has experience in the CJEU?")+
  xlab("")+
  theme_minimal()
```




```
sample2 %>%
  group_by(celex) %>%
  filter(type != "NGO" & type != "Other") %>%
  ggplot(aes(as.factor(win), fill = as.factor(experience))) +
  labs(fill = "Experienced attorney")+
  geom_bar(aes(y = (..count..)/sum(..count..)*100), position = "dodge") +
  ylab("Percent of cases") +
  ggtitle("Who wins in the CJEU?")+
  xlab("")+
  theme_minimal() +
  facet_wrap(~ type) +
  xlab("Winning is flagged 1")+
  scale_fill_manual(values = c( "purple", "red"))
```



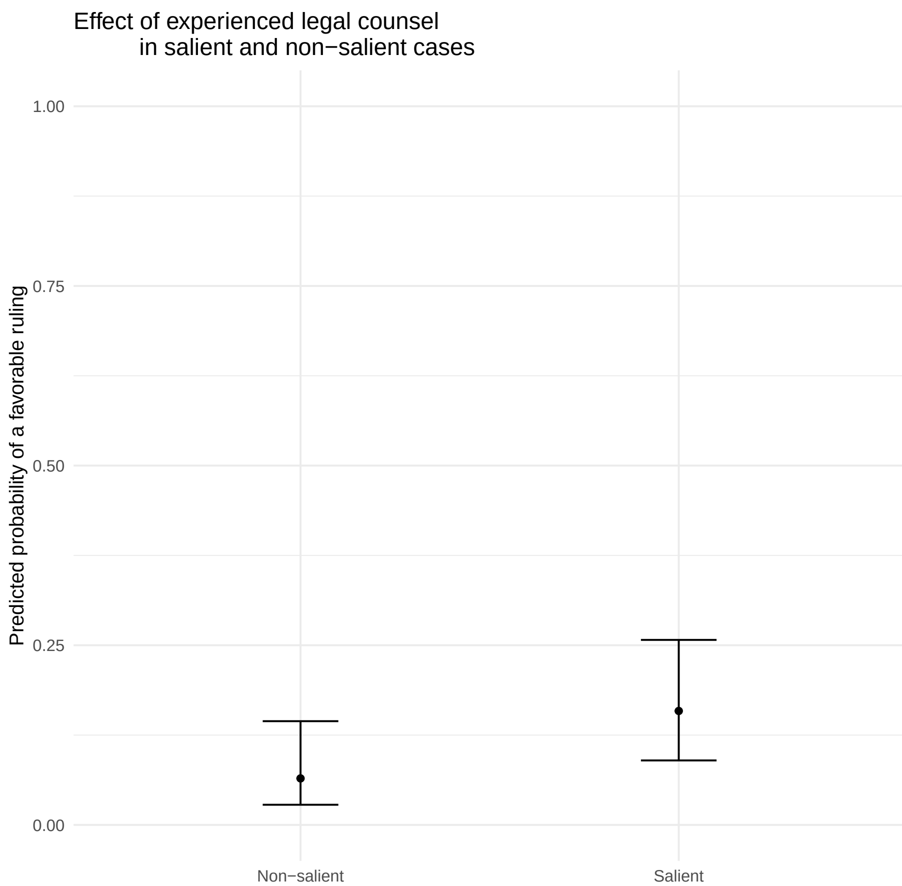

```
table(sample2$type)

##
##        Individual             Other           Company    EU institution
##               217               102               499                 0
##               NGO State institution
##                11               161
```

8.3 Plotting effects – creating scenarios

```
# Saving my se's in the model object


m2$cluster_m2 <- vcov_m2
```

```r
m1$cluster <- vcov_m1
# Setting seed
set.seed(24)

simBetas <- mvrnorm(n = 1000,
                    mu = coefficients(m1),
                    Sigma = m1$cluster)


names(coefficients(m1))

##  [1] "(Intercept)"                       "experience"
##  [3] "roledefendant"                     "typeOther"
##  [5] "typeCompany"                       "typeNGO"
##  [7] "typeState institution"             "type_opponentOther"
##  [9] "type_opponentCompany"              "type_opponentNGO"
## [11] "type_opponentState institution" "member_state12"
## [13] "member_state6"                     "member_state9"
## [15] "binary_salience"                   "government_support"
## [17] "commission_support"

table(sample2$member_state)

##
##  10  12   6   9
## 256 390  65 279

xMatrix <- cbind(1, #the intercept
                 1, # experience
                 0, # defendant
                 0, # other
                 1, # Company
                 0, # NGO
                 0, # State institution
                 0, # opponent other
                 1, # opponent company
                 0, # opponent NGO
                 0, # opponent state institution
                 1, # M 12
                 0, # M 6
                 0, # M 9,

                 c(0, 1), # binary salience,
                 0, # Government support,
                 0 # Commission support
```

```
                 )



ncol(simBetas) == ncol(xMatrix) #yay!!

## [1] TRUE

### Calculating predicted probabilities: Her multipliserer du simuleringen med xmatrisen
xBetaMatrix <- xMatrix %*% t(simBetas ) ## this just means x times the betas
predProbs <- 1/(1+exp(-xBetaMatrix)) #This is the predicted probability, for another type of


### Getting point estimates and confidence intervals:
quantileValues <- apply(X = predProbs,  ## read up on the apply() family of functions!
                        MARGIN = 1, ## this means we are applying a function to all the rows
                        FUN =  ## The fun argument defines what I want to do with all the ro
                          ## What we want to do here is to use quantile to get the quantiles
                          ### bounds of the confidence intervals and our point estimates:
                          quantile, probs = c(.05,.5,.95))
quantileValues <- as.data.frame(t(quantileValues))

plotPoints <- cbind(c("Salient", "Non-salient"),quantileValues)
plotPoints

##   c("Salient", "Non-salient")         5%        50%       95%
## 1                     Salient 0.08974371 0.15857011 0.2574454
## 2                 Non-salient 0.02804750 0.06471242 0.1443523

colnames(plotPoints) <- c("Salient", "lower", "estimate", "upper")



ggplot(plotPoints,
       aes(x = Salient,
           y = estimate,
           ymin = lower,
           ymax = upper)) +
  geom_errorbar(width =.2)+
  geom_point()+
  ylim(0,1)+
  ylab("Predicted probability of a favorable ruling")+
  xlab("")+
  theme_minimal() +
  ggtitle("Effect of experienced legal counsel
          in salient and non-salient cases")
```



9 Digasnostics

The underlying assumption of the logistical regression model are that (1) the dependent variable is binary; (2) the probability curve is S-shaped and the logit curve is linear; (3) there are no influential observations; (4) there is no multicollinearity among the predictors; (5) there are no empty cells; (6) there is no "complete separation"; (7) no omitted variable bias and (8) the observations are independent and identically distributed.

In this section I go through each of the assumptions and evaluate the model where I control salience, role and type of actor.

The first assumption holds as the dependent variable is binary – 0 for lost case and 1 for won case. The assumption that the observations are independent and identically distributed is not relevant because I am dealing with observational data – and the data at hand is the sample of cases are more or less equal to universe of cases. No omitted variable bias is a theoretical assumption. There are many unobserved factors that may affect not just the assignment to treatment, but also the decision of the Court and the cases that end up in the CJEU in the first place. This is not controlled for because many of the mechanisms that may affect assignment to treatment, cases being referred and the decision of the CJEU are unobservable.

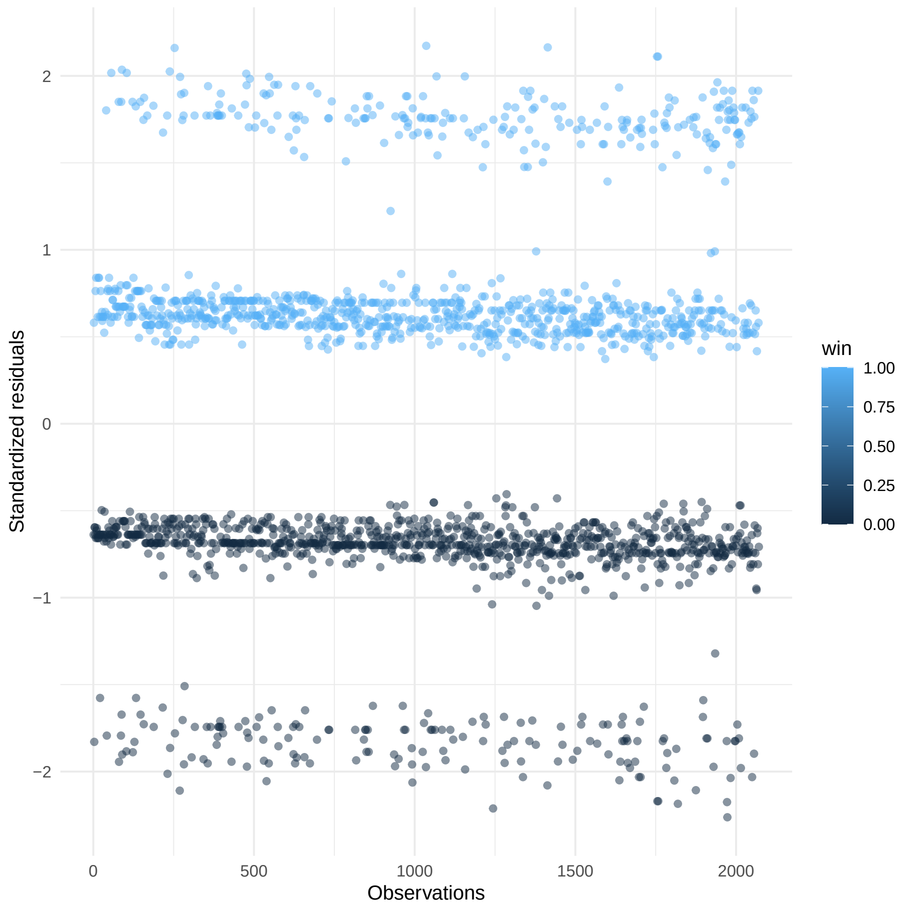
9.0.1 The regression has the shape of an S

In order to investigate if the second assumption holds I make sure that the relationship between the independent variables and the logit-outcome is linear. To show this graphically is difficult when the variables are categorical – like most of my variables.

9.0.2 Influential values

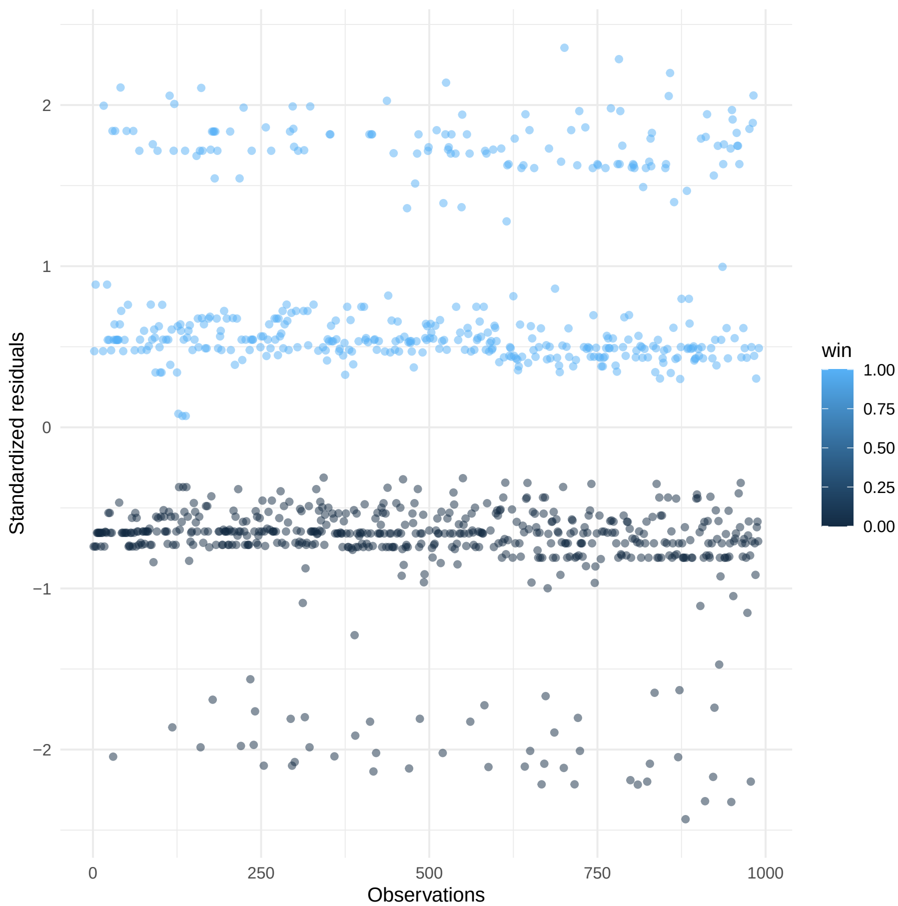
Checking for influential values. Plotting the standardized residuals. The standard normal distribution lies between -4 and 4. Values above three indicates outliers and should be further investigated as they might affect the results.

```
# Sample1 model
model.data <- augment(law) %>%
  mutate(index = 1:n())

model.data %>%
  top_n(3, .cooksd)

## # A tibble: 3 x 17
##   .rownames   win lawyer role  type  type_opponent member_state binary_salience
##   <chr>     <dbl>  <dbl> <chr> <fct> <fct>         <chr>                  <dbl>
## 1 329           1      0 appl~ Stat~ NGO           9                          0
## 2 331           0      0 defe~ NGO   State instit~ 9                          0
## 3 2403          1      0 appl~ NGO   Individual    12                         0
## # ... with 9 more variables: government_support <dbl>,
## #   commission_support <dbl>, .fitted <dbl>, .resid <dbl>, .std.resid <dbl>,
## #   .hat <dbl>, .sigma <dbl>, .cooksd <dbl>, index <int>

ggplot(model.data, aes(index, .std.resid))+
  geom_point(aes(color = win), alpha = .5) +
  theme_minimal() +
  ylab("Standardized residuals") +
  xlab("Observations")
```




```
# Manual count
model.data %>%
  filter(abs(.std.resid) > 3)

## # A tibble: 0 x 17
## # ... with 17 variables: .rownames <chr>, win <dbl>, lawyer <dbl>, role <chr>,
## #   type <fct>, type_opponent <fct>, member_state <chr>, binary_salience <dbl>,
## #   government_support <dbl>, commission_support <dbl>, .fitted <dbl>,
## #   .resid <dbl>, .std.resid <dbl>, .hat <dbl>, .sigma <dbl>, .cooksd <dbl>,
## #   index <int>

# Sample2 model
model.data <- augment(m2) %>%
  mutate(index = 1:n())

model.data %>%
  top_n(3, .cooksd)
```

```
## # A tibble: 3 x 17
##   .rownames   win experience role  type  type_opponent member_state
##   <chr>     <dbl>      <dbl> <chr> <fct> <fct>         <chr>
## 1 1246          1          1 appl~ NGO   NGO           12
## 2 1311          0          1 defe~ Indi~ NGO           12
## 3 1389          0          0 defe~ Other Individual    12
## # ... with 10 more variables: binary_salience <dbl>, government_support <dbl>,
## #   commission_support <dbl>, .fitted <dbl>, .resid <dbl>, .std.resid <dbl>,
## #   .hat <dbl>, .sigma <dbl>, .cooksd <dbl>, index <int>

ggplot(model.data, aes(index, .std.resid))+
  geom_point(aes(color = win), alpha = .5) +
  theme_minimal() +
  ylab("Standardized residuals") +
  xlab("Observations")
```



```r
# Manual count
model.data %>%
  filter(abs(.std.resid) > 3)

## # A tibble: 0 x 17
## # ... with 17 variables: .rownames <chr>, win <dbl>, experience <dbl>,
## #   role <chr>, type <fct>, type_opponent <fct>, member_state <chr>,
## #   binary_salience <dbl>, government_support <dbl>, commission_support <dbl>,
## #   .fitted <dbl>, .resid <dbl>, .std.resid <dbl>, .hat <dbl>, .sigma <dbl>,
## #   .cooksd <dbl>, index <int>
```

9.0.3 Multicollinearity

VIF-test measures how much of the variance in each independent variable can be explained by the other variables in the analysis. As a general rule of thumb a VIF-value under 5 indicates no multicolinearity, whilst values between 5 and 10 is considered no ideal, but yet not very problematic. Values above 10 indicates strong multicolinearity(ibid).

```r
vif(law)

##                             GVIF Df GVIF^(1/(2*Df))
## lawyer                  1.365042  1        1.168350
## role                    1.576987  1        1.255782
## type                    3.084576  4        1.151196
## type_opponent           3.003929  4        1.147390
## member_state            1.224151  3        1.034282
## binary_salience         2.432330  1        1.559593
## government_support      1.518394  1        1.232231
## commission_support      1.187288  1        1.089628
## lawyer:binary_salience 2.026112  1        1.423416

vif(m2)

##                                 GVIF Df GVIF^(1/(2*Df))
## experience                  1.175269  1        1.084098
## role                        1.516411  1        1.231427
## type                        3.313835  4        1.161559
## type_opponent               2.995752  4        1.147000
## member_state                1.367333  3        1.053527
## binary_salience             2.323467  1        1.524292
## government_support          1.386600  1        1.177540
## commission_support          1.197835  1        1.094456
## experience:binary_salience 2.119332  1        1.455793
```

Running the VIF-test I find that all my independent variables have VIF-values between 1 and 3 which indicates no multicolinearity.

9.0.4 Complete seperation

Checking for complete separation is easily done by plotting the data. The plot below indicates that also this assumption is met. Checking for empty cells is unecessary are observations with missing values are not included in the model.

9.1 Goodness of fit

9.1.1 McFadden's pseudo R2

McFadden's pseudo R2 is a measure that compares the log-likelihood value for my model and compares it to the log-likelihood value of a model without any variables – an intercept-only model. The value ranges from zero to one. Values closer to 1 indicates good predictive power. Values closer to zero indicates no predictive power. The results show that the model is better than an intercept-only model, however, the model does not explain much of the variation. The models with more variables have slightly higher McFadden scores.

```
PseudoR2(law, which = c("McFadden", "AIC", "BIC", "logLik"))

##      McFadden           AIC           BIC        logLik
##     0.3013189  2041.0467818  2142.4909453 -1002.5233909

PseudoR2(law)

##  McFadden
## 0.3013189

PseudoR2(m2, which = c("McFadden", "AIC", "BIC", "logLik"))

##     McFadden          AIC          BIC       logLik
##    0.3444769  926.7302187 1014.8889077 -445.3651093

PseudoR2(m2)

##  McFadden
## 0.3444769
```

9.1.2 Hosmler-Lemeshow-test

Tests ho good the model fits the data by comparing observed and predicted values – meaning that it compares the observed, real values of 1 and + to the models fitted values (ibid). The test does this by comparing subgroups of the population estimated. The Hosmer-Lemeshow-test is not supposed to give significant results, because this means that the model is not a good fit for the data. The results from running the test are not significant suggesting that the model is good at describing the data.

```
############# Sample1 model
hl <- hoslem.test(law$y,
                  fitted(law),
                  g = 10)
# G= 10 ten subgroups
hl

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  law$y, fitted(law)
## X-squared = 3.0483, df = 8, p-value = 0.9313

# Shows difference in
# observed and expected Y-values
# for ten subgroups
cbind(hl$expected, hl$observed)

##                     yhat0     yhat1  y0  y1
## [0.0782,0.166] 195.01370  32.98630 195  33
## (0.166,0.192]  153.84614  34.15386 159  29
## (0.192,0.213]  167.88394  43.11606 164  47
## (0.213,0.233]  157.59929  44.40071 157  45
## (0.233,0.277]  168.94298  56.05702 165  60
## (0.277,0.774]   82.25313 117.74687  89 111
## (0.774,0.806]   42.66412 159.33588  43 159
## (0.806,0.833]   36.91588 167.08412  37 167
## (0.833,0.857]   31.64523 173.35477  30 175
## (0.857,0.934]   24.23560 181.76440  22 184

############# Sample2 model
hl <- hoslem.test(law$y,
                  fitted(law),
                  g = 10)
# G= 10 ten subgroups
hl

##
##  Hosmer and Lemeshow goodness of fit (GOF) test
##
## data:  law$y, fitted(law)
## X-squared = 3.0483, df = 8, p-value = 0.9313
```

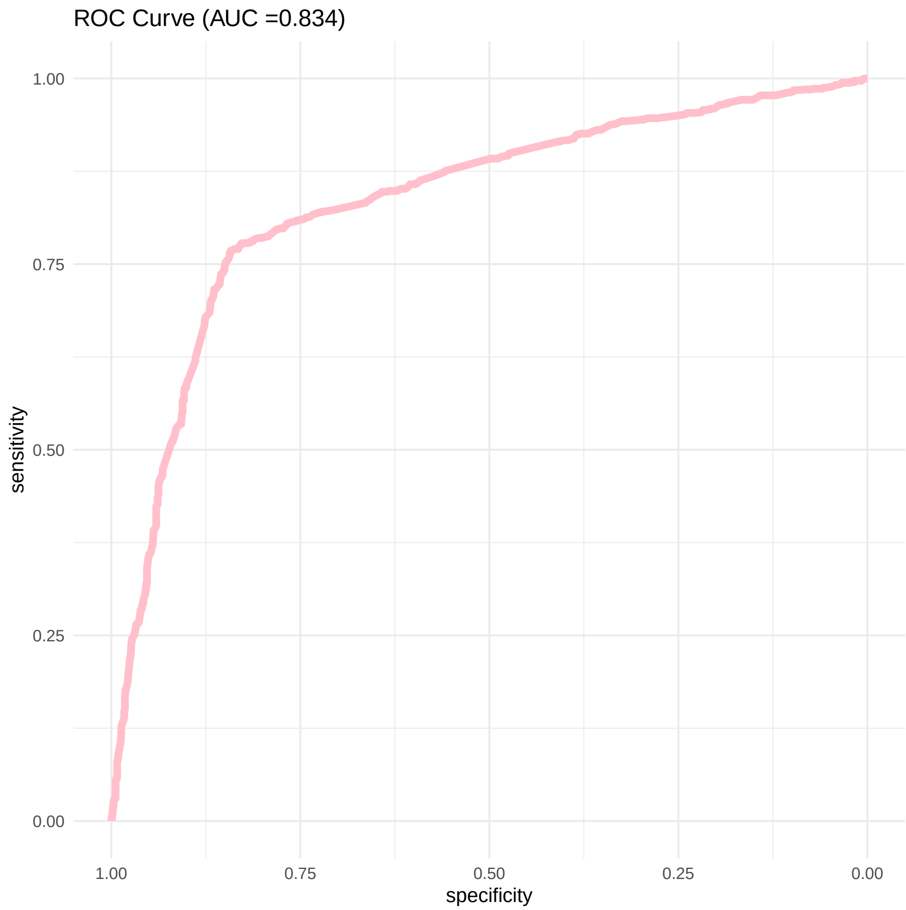
```
# Shows difference in
# observed and expected Y-values
# for ten subgroups
cbind(hl$expected, hl$observed)

##                     yhat0     yhat1  y0  y1
## [0.0782,0.166] 195.01370  32.98630 195  33
## (0.166,0.192]  153.84614  34.15386 159  29
## (0.192,0.213]  167.88394  43.11606 164  47
## (0.213,0.233]  157.59929  44.40071 157  45
## (0.233,0.277]  168.94298  56.05702 165  60
## (0.277,0.774]   82.25313 117.74687  89 111
## (0.774,0.806]   42.66412 159.33588  43 159
## (0.806,0.833]   36.91588 167.08412  37 167
## (0.833,0.857]   31.64523 173.35477  30 175
## (0.857,0.934]   24.23560 181.76440  22 184
```

10 How well does my model predict?

I am modelling predicted probabilitites. To estimate how well my models predict, a ROC-cruve can be helpful (Receiving Operating Characteristics). When using logistical regression, the goal is a model that predicts the outcome of the indepdent variable corectly at all times. The ROC-curve shows how well my model predics by determining the relationship between true positive values (the predictions my model predics as 1 that is observed to be 1) and false positive values (the prediction my model predicts as 1 but is actually 0) using various cut-off values. I create a ROC-curve to evaluate the overall performance of my model. The ROC-curve defines the optimal cut-off value for me – indicating at which point from 0 to 100 my model predicts correctly. Is my model correct in 70 percent of all the instances? Is my model correct in 60 percent of all instances?

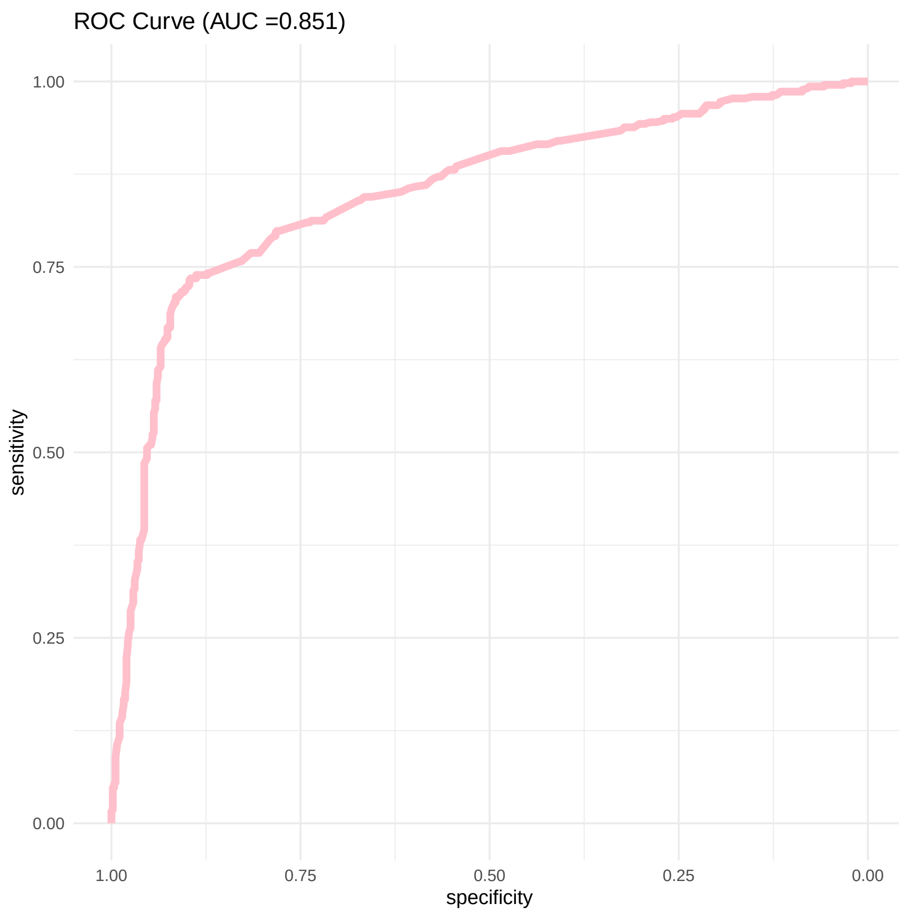
```
# Sample1 model
preds <- predict(law,
                 sample1,
                 type = "response")

roc_obj <- roc(sample1$win, preds)

## Setting levels:  control = 0, case = 1
## Setting direction:  controls < cases

auc_m2 <- auc(roc_obj)

auc_m2 <- round(auc_m2, digits = 3)
```

```r
law_roc <- ggroc(roc_obj, color="pink", size = 2) +
  ggtitle(paste0("ROC Curve ", "(AUC =", auc_m2, ")")) +
  theme_minimal()

law_roc
```




```r
test <- table(predicted = ifelse(preds > auc_m2, 1, 0),
              observed = sample1$win)

test

##          observed
## predicted    0    1
##         0 1011  681
##         1   50  329
```

```r
# Sample3 model
preds <- predict(m2,
                 sample2,
                 type = "response")

roc_obj <- roc(sample2$win, preds)

## Setting levels: control = 0, case = 1
## Setting direction: controls < cases

auc_m2 <- auc(roc_obj)

auc_m2 <- round(auc_m2, digits = 3)

m2_roc <- ggroc(roc_obj, color="pink", size = 2) +
  ggtitle(paste0("ROC Curve ", "(AUC =", auc_m2, ")")) +
  theme_minimal()

m2_roc
```



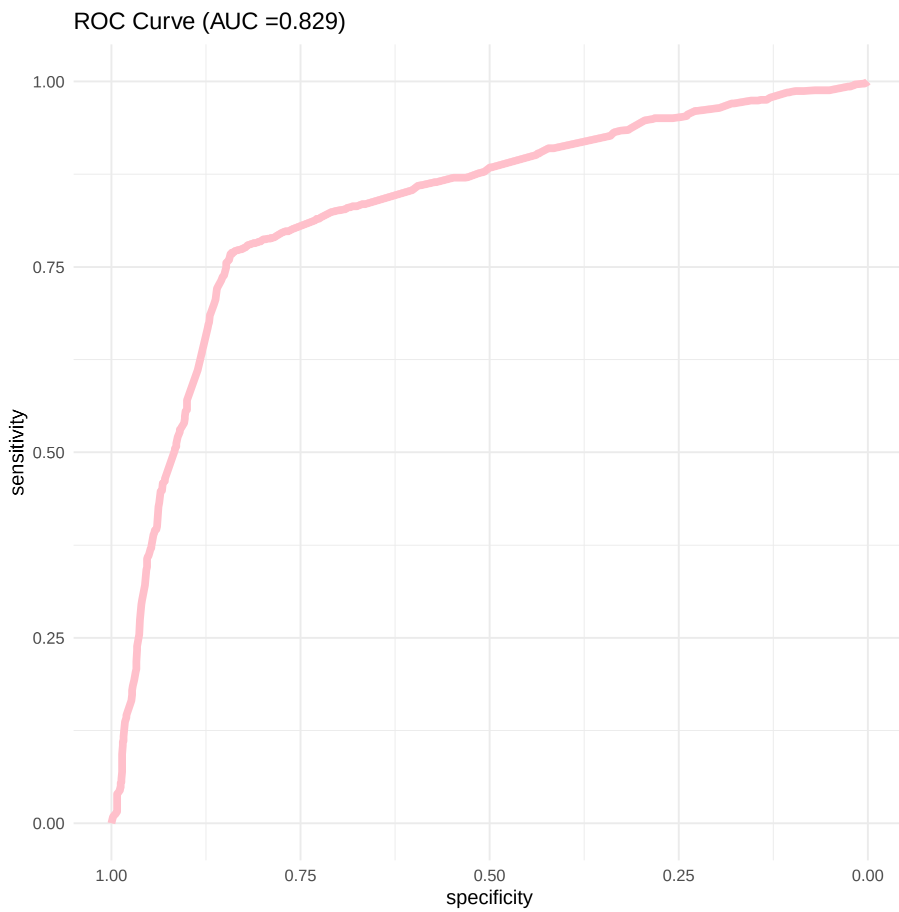

```
test <- table(predicted = ifelse(preds > auc_m2, 1, 0),
              observed = sample2$win)

test

##          observed
## predicted   0   1
##         0 523 208
##         1  30 229
```

Area under the curve (AUC) equal to 1 means that the model makes perfect predictions, meaning that the model predicts Y = 1 when Y=1 is observed in all incidents. The model predicts correctly in approximately 84 percent of all incidents. The model I have made is able to classify Y = 1 and Y = 0 correctly in approximately 84 percent of the time.

10.1 Predictions without main explanatory variables

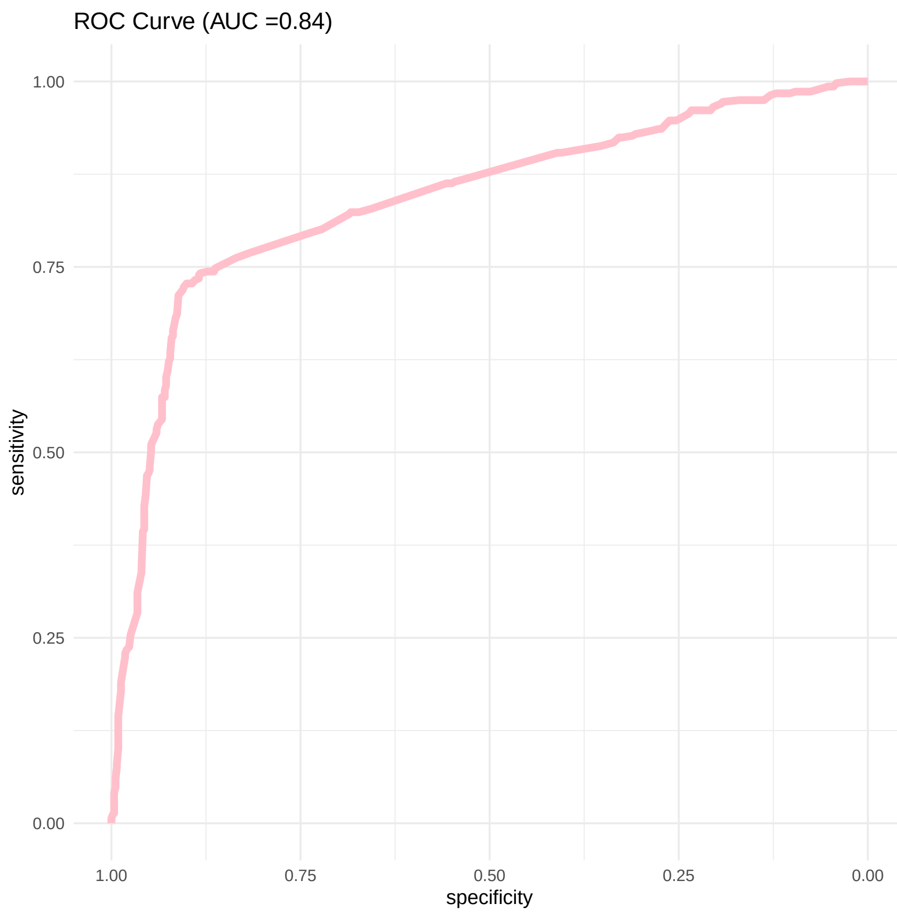
Here I will check to see if my models predics worse without legal representation and lawyer experience.

```
law2 <- glm(win~role + type +
            type_opponent + member_state
           + government_support
           + commission_support,
          family = binomial(link = "logit"),
          data = sample1)

preds <- predict(law2,
                 sample1,
                 type = "response")

roc_obj <- roc(sample1$win, preds)

## Setting levels:  control = 0, case = 1
## Setting direction:  controls < cases

auc_m2 <- auc(roc_obj)

auc_m2 <- round(auc_m2, digits = 3)

law_roc <- ggroc(roc_obj, color="pink", size = 2) +
  ggtitle(paste0("ROC Curve ", "(AUC =", auc_m2, ")")) +
  theme_minimal()

law_roc
```




```
test <- table(predicted = ifelse(preds > auc_m2, 1, 0),
              observed = sample1$win)

test

##           observed
## predicted    0    1
##          0 1006  636
##          1   55  374

# Experience model
m3 <- glm(win~ role + type +
            type_opponent + member_state
          + government_support
          + commission_support,
          family = binomial(link = "logit"),
```

```r
          data = sample2)

preds <- predict(m3,
                 sample2,
                 type = "response")

roc_obj <- roc(sample2$win, preds)

## Setting levels: control = 0, case = 1
## Setting direction: controls < cases

auc_m2 <- auc(roc_obj)

auc_m2 <- round(auc_m2, digits = 3)

law_roc <- ggroc(roc_obj, color="pink", size = 2) +
  ggtitle(paste0("ROC Curve ", "(AUC =", auc_m2, ")")) +
  theme_minimal()

law_roc
```

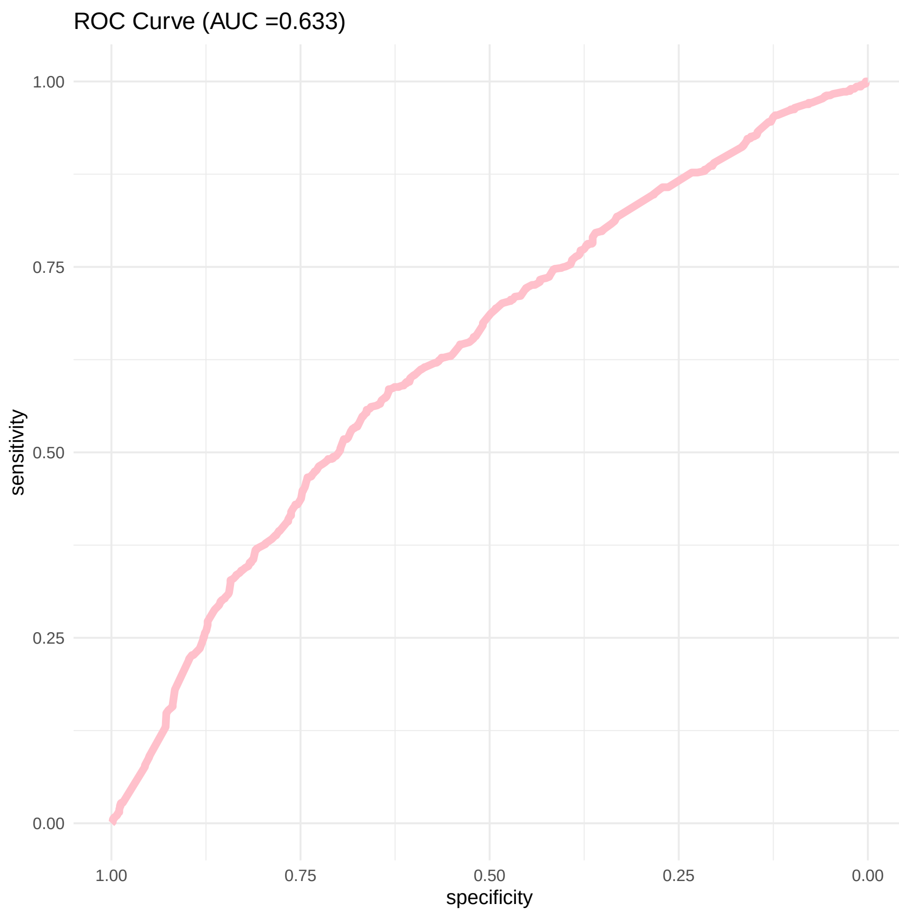



```
test <- table(predicted = ifelse(preds > auc_m2, 1, 0),
              observed = sample1$win)
test

##           observed
## predicted    0    1
##          0 1006  636
##          1   55  374
```

10.2 Predictions without government and Commission supports

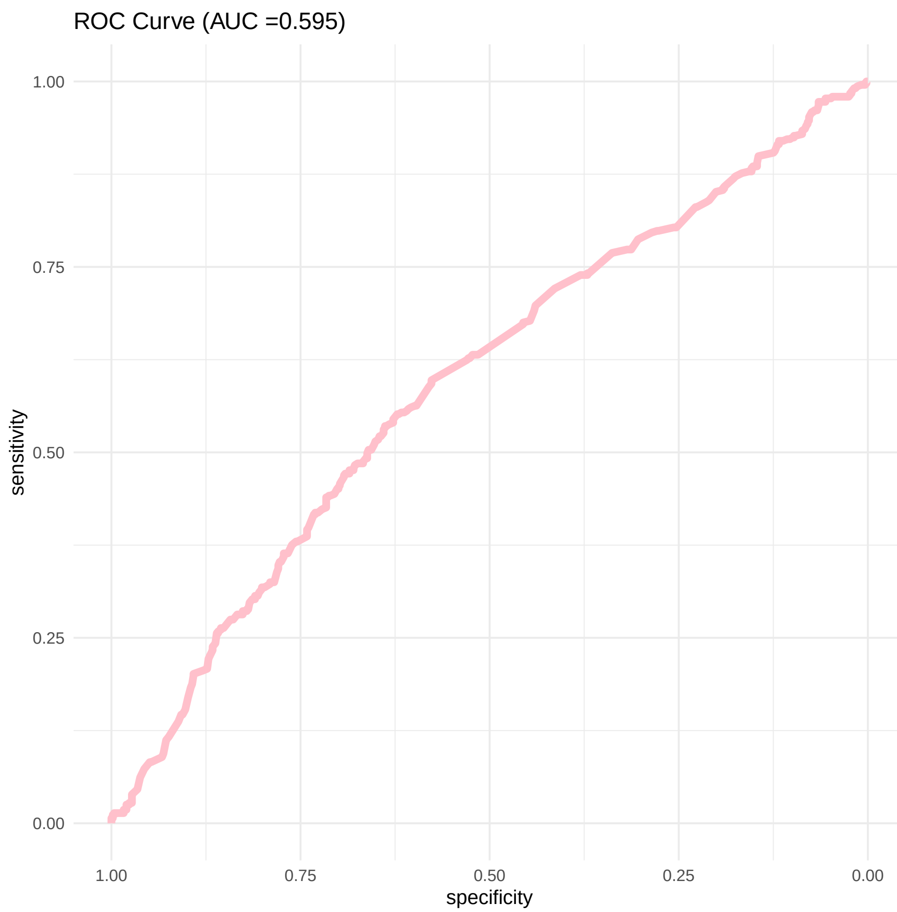
```
law2 <- glm(win~role + lawyer + type +
            type_opponent + member_state
```

```
            + salience + salience*lawyer,
           family = binomial(link = "logit"),
           data = sample1)

preds <- predict(law2,
                 sample1,
                 type = "response")

roc_obj <- roc(sample1$win, preds)

## Setting levels:  control = 0, case = 1
## Setting direction:  controls < cases

auc_m2 <- auc(roc_obj)

auc_m2 <- round(auc_m2, digits = 3)

law_roc <- ggroc(roc_obj, color="pink", size = 2) +
  ggtitle(paste0("ROC Curve ", "(AUC =", auc_m2, ")")) +
  theme_minimal()

law_roc
```




```
test <- table(predicted = ifelse(preds > auc_m2, 1, 0),
              observed = sample1$win)

test

##           observed
## predicted   0   1
##          0 952 787
##          1 109 223

# Experience model
m3 <- glm(win~ experience + role + type +
            type_opponent + member_state +
          salience + experience*salience ,
          family = binomial(link = "logit"),
          data = sample2)
```

```r
preds <- predict(m3,
                 sample2,
                 type = "response")

roc_obj <- roc(sample2$win, preds)

## Setting levels: control = 0, case = 1
## Setting direction: controls < cases

auc_m2 <- auc(roc_obj)

auc_m2 <- round(auc_m2, digits = 3)

law_roc <- ggroc(roc_obj, color="pink", size = 2) +
  ggtitle(paste0("ROC Curve ", "(AUC =", auc_m2, ")")) +
  theme_minimal()

law_roc
```




```
test <- table(predicted = ifelse(preds > auc_m2, 1, 0),
              observed = sample1$win)
test

##           observed
## predicted   0   1
##         0 952 787
##         1 109 223
```