Checks and Balances

For this data challenge, your task is to extend the analysis we did in the lecture notes by conducting a directed-dyad analysis instead of a monadic (country) analysis. Jessica Weeks in her 2012 paper makes the case that we need to factor in certain characteristics about the strategic environment that countries inhabit to appropriately estimate the impact of different regime types of MID onset initiation. The country-level analysis we did in the lecture notes can't fully capture the broader set of interactions between countries as possible confounding factors, but a dyadic analysis possibly can. I've provided some code on the next page to help you construct your dataset for the analysis. Basically, this code starts by creating a directed leader-dyad dataset and populates it with MID initiation data, peace spells, and data on leader willingness to use force. It then collapses it to the dyad-year level. This is done so the data can be merged with the replication dataset from Weeks which has the other variables we'll use for our analysis.

With this dataset, estimate a logit model with the variable `mid_init` as the outcome and an interaction between `regime1` and `lwuf1`, which denote the set of regime types and leader willingness to use force for side 1 in the data. You'll also include a cubic peace spells trend. From there, it's up to you to decide what factors you think you should try to control for in your analysis. At minimum, I'd consider including the following in my model, but you can look at including other factors as well.

- `cap_1`: Military capabilities (CINC score) for side 1
- `cap_2`: Military capabilities (CINC score) for side 2
- `I(cap_1 / (cap_1 + cap_2))`: Side 1's share of capabilities relative to side 2
- `logdist`: Natural log of the distance between side 1 and side 2
- `newregime_1`: Indicator of whether side 1 is a new country.
- `dependlow`: The minim level of trade dependence between sides.

For your submission, answer the following question: how robust are the results we obtained in the lecture notes to controlling for dyadic factors?

```r
## packages
library(tidyverse)
library(peacesciencer)
library(sandwich)
library(lmtest)
library(sjPlot)
library(coolorrr)
set_palette(
  qualitative = c("steelblue", "grey20", "red")
)

## build dayd-year data
create_leaderdyadyears(
  system = "cow"
) |>
  filter(year %in% 1945:1999) |>
  add_gml_mids() |>
  add_spells() |>
  add_lwuf() -> Data

Data |>
  group_by(
    ccode1, ccode2, year
  ) |>
  summarize(
    mid_init = max(gmlmidonset_init, na.rm = T),
    peace_spell = max(gmlmidinitspell, na.rm = T),
    lwuf1 = unique(theta2_mean1) |> mean(na.rm = T),
    lwuf2 = unique(theta2_mean2) |> mean(na.rm = T)
  ) -> Data

## read in data from Weeks (2012)
haven::read_dta(
  here::here( ## adjust location based on your files
    "DPR 390",
    "Quarto Notes",
    "studying-conflict-with-r",
    "Data",
    "WeeksAPSR2012",
    "WeeksAPSR2012.dta"
  )
```

```
) -> weeks_dt

## merge the data
weeks_dt |>
  left_join(
    Data
  ) -> Data

## one more thing...
## make a regime factor
library(socsci)
Data |>
  mutate(
    regime1 = frcode(
      democracy_1 == 1 ~ "Democracy",
      juntajlw_1 == 1 ~ "Junta",
      strongmanjlw_1 == 1 ~ "Strongman",
      machinejlw_1 == 1 ~ "Machine",
      bossjlw_1 == 1 ~ "Boss",
      TRUE ~ "Other"
    )
  ) |>
  drop_na(peace_spell) -> Data
```