# Replication file for Political Preferences, Policy Loss and Shifting Support for European Integration

Simon Hix and Bjørn Høyland

### Replication material

In this document, we describe the steps used to collect, prepare, analyse the data and visualize the results.

## Data

We rely on two data-sources for our analysis:

1. Eurobarometer surveys (1976 - 2023)
2. EU legislation (regulations and directives) from EurLex (1953 - 2024)

### Eurobarometer

For the Eurobarometer, we merged all surveys that contained data on support for the EU and left-right self-placement. All of these surveys also contained information on age and education. More information about Eurobarometer can be found at the Eurobarometer webpage. The code below reproduces Figure 1.

```
source("eb_support.R")
```

### EurLex data

We interacted with the EurLex API via `library(EurLex)` to collect all EU directives and regulations. We stored the directives and regulations in folders, clean them, including dropping all Commission legislation and stored them as a data-frame called `legisdata.RData`. In the interest of space, we only attach the resulting data-frame.

```
source("collect_legislation.R")
collect_legislation("directive")
collect_legislation("regulation")
source("legislation.R")
```

## Measuring EU policy-output

Next, we calculated our 7 left-right measures:

1. Keywords
2. CMPgen
3. CMPecon
4. CMPsocial
5. LLM
6. Factor

The next set of scripts adds each of these measures as columns to the `legisdata.Rdata` dataset.

```
library(data.table)
load("legisdata.RData")
source("keywords.R")
source("manifesto.R")
source("llama.R")
source("factor.R")
save(legis, file ="all_years_legis.RData")
```

### Annualized EU-policy-output

Having build the legislation-level dataset, we now aggregate the scores by year. As the LLM measure is categorical, we use the logit transformation at the aggregated level. We then merge the legislative measures with the Eurobarometer-data, and store this as `hh_data.RData`.

```
source("prepare_analysis.R")
```

```
\begin{table}[!htbp] \centering \renewcommand*{\arraystretch}{1.1}\caption{Descriptive statis
\begin{tabular}{lrrrrrrr}
\hline
\hline
Variable & N & Mean & Std. Dev. & Min & Pctl. 25 & Pctl. 75 & Max \\
\hline
EU Support & 1196114 & 0.61 & 0.49 & 0 & 0 & 1 & 1 \\
Policy-loss (Keywords) & 1196114 & 1 & 1.4 & 0.0000043 & 0.065 & 1.4 & 6.3 \\
Policy-loss (CMPgen) & 1196114 & 1 & 1.4 & 0.000092 & 0.057 & 1.4 & 7.3 \\
Policy-loss (CMPecon) & 1196114 & 1.1 & 1.5 & 0.00026 & 0.061 & 1.4 & 9.2 \\
Policy-loss (CMPsocial) & 1196114 & 1 & 1.4 & 0.000013 & 0.056 & 1.4 & 8.3 \\
Policy-loss (LLM) & 1196114 & 1.2 & 1.7 & 0.0000031 & 0.11 & 1.7 & 10 \\
Policy-loss (Factor) & 1196114 & 1 & 1.4 & 0.000000000097 & 0.059 & 1.4 & 7.5 \\
Left-right & 1196114 & 0.000024 & 1 & -2 & -0.62 & 0.78 & 2.2 \\
Age & 1196114 & 47 & 18 & 15 & 32 & 61 & 99 \\
Age squared & 1196114 & 25 & 18 & 2.2 & 10 & 37 & 98 \\
Education & 1196114 &  &  &  &  &  &  \\
... High School & 872990 & 73\% &  &  &  &  &  \\
... University & 323124 & 27\% &  &  &  &  &  \\
Sex & 1196114 &  &  &  &  &  &  \\
...  Male & 585769 & 49\% &  &  &  &  &  \\
... Female & 610345 & 51\% &  &  &  &  &  \\
Identity & 251381 & 0.57 & 0.49 & 0 & 0 & 1 & 1\\
\hline
\hline
\end{tabular}
}
\end{table}
```

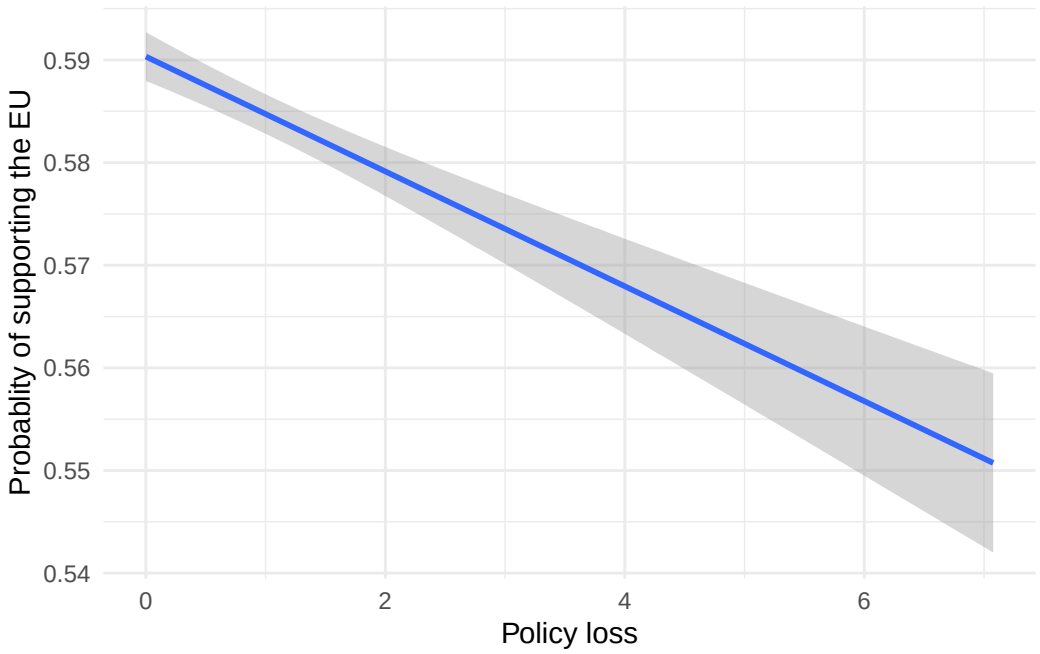
### Examples

```
examples <- legis[celex %in% c("32008L0115", "32003L0088", "32004L0025", "32006L0123"),
                  list(title, keywords,keywords_score, CMPgen, llm_class, CMPecon,
```

```
                        CMPsocial, Factor, keywords_mean, gen_mean,econ_mean,
                        social_mean, Factor_mean),]
# mean-centering examples
examples[,Keywords := round(keywords_score - keywords_mean, 2)]
examples[, CMPgen := round(CMPgen - gen_mean,2)]
examples[, CMPecon := round(CMPecon - econ_mean,2)]
examples[, CMPsocial := round(CMPsocial - social_mean,2)]
examples[, Factor := round(Factor - Factor_mean,2)]
examples <- examples |>
  rename(LLM = llm_class) |>
  rename(Title = title) |>
  select(Title, Keywords, CMPgen, CMPecon, CMPsocial, LLM, Factor)
x <-xtable(examples, align = c("R{.5cm}","p{5cm}", "p{1.3cm}",
                               "p{1.2cm}",  "p{1.3cm}",
                               "p{1.4cm}", " p{1cm}","p{1cm}"),
           caption = "Coding examples",label = "tab:examples")
print(x, floating = TRUE, include.rownames = FALSE, size = "footnotesize",
      file = "../tables/example.tex")
```

**Descriptive data**

```
hh_data |>
  na.omit() |>
  ggplot(aes(x = policy_loss_Factor, y = proeu)) +
  geom_smooth(method = "lm")+
  theme_minimal()+
  xlab("Policy loss")+
  ylab("Probablity of supporting the EU") +
  theme(legend.title=element_blank())
```




```
ggsave("figures/policy_loss_eu.png",
       width = 6, height = 2.65, units = "in", dpi = 450)
```

```
####
labs <- c("EU Support", "Policy-loss (Keywords)", "Policy-loss (CMPgen)",
          "Policy-loss (CMPecon)", "Policy-loss (CMPsocial)", "Policy-loss (LLM)", "Policy-lo
          "Left-right", "Age", "Age squared", "Education", "Sex", "Identity")
hh_data |>
  select(proeu,policy_loss_Keywords, policy_loss_CMPgen, policy_loss_CMPecon,
         policy_loss_CMPsocial, policy_loss_LLM, policy_loss_Factor,
         left_right,  age, age_sq, education,sex, identity) |>
vtable::st(out = "latex", labels = labs, title = "Descriptive statistics",
           anchor = "tab:desc_summary")
```

```
\begin{table}[!htbp] \centering \renewcommand*{\arraystretch}{1.1}\caption{Descriptive statis
\begin{tabular}{lrrrrrrr}
\hline
\hline
Variable & N & Mean & Std. Dev. & Min & Pctl. 25 & Pctl. 75 & Max \\
\hline
EU Support & 1196114 & 0.61 & 0.49 & 0 & 0 & 1 & 1 \\
Policy-loss (Keywords) & 1196114 & 1 & 1.4 & 0.0000043 & 0.065 & 1.4 & 6.3 \\
Policy-loss (CMPgen) & 1196114 & 1 & 1.4 & 0.000092 & 0.057 & 1.4 & 7.3 \\
```

```
Policy-loss (CMPecon) & 1196114 & 1.1 & 1.5 & 0.00026 & 0.061 & 1.4 & 9.2 \\
Policy-loss (CMPsocial) & 1196114 & 1 & 1.4 & 0.000013 & 0.056 & 1.4 & 8.3 \\
Policy-loss (LLM) & 1196114 & 1.2 & 1.7 & 0.0000031 & 0.11 & 1.7 & 10 \\
Policy-loss (Factor) & 1196114 & 1 & 1.4 & 0.000000000097 & 0.059 & 1.4 & 7.5 \\
Left-right & 1196114 & 0.000024 & 1 & -2 & -0.62 & 0.78 & 2.2 \\
Age & 1196114 & 47 & 18 & 15 & 32 & 61 & 99 \\
Age squared & 1196114 & 25 & 18 & 2.2 & 10 & 37 & 98 \\
Education & 1196114 &  &  &  &  &  &  \\
... High School & 872990 & 73\% &  &  &  &  &  \\
... University & 323124 & 27\% &  &  &  &  &  \\
Sex & 1196114 &  &  &  &  &  &  \\
...  Male & 585769 & 49\% &  &  &  &  &  \\
... Female & 610345 & 51\% &  &  &  &  &  \\
Identity & 251381 & 0.57 & 0.49 & 0 & 0 & 1 & 1\\
\hline
\hline
\end{tabular}
}
\end{table}
```

### Analysis

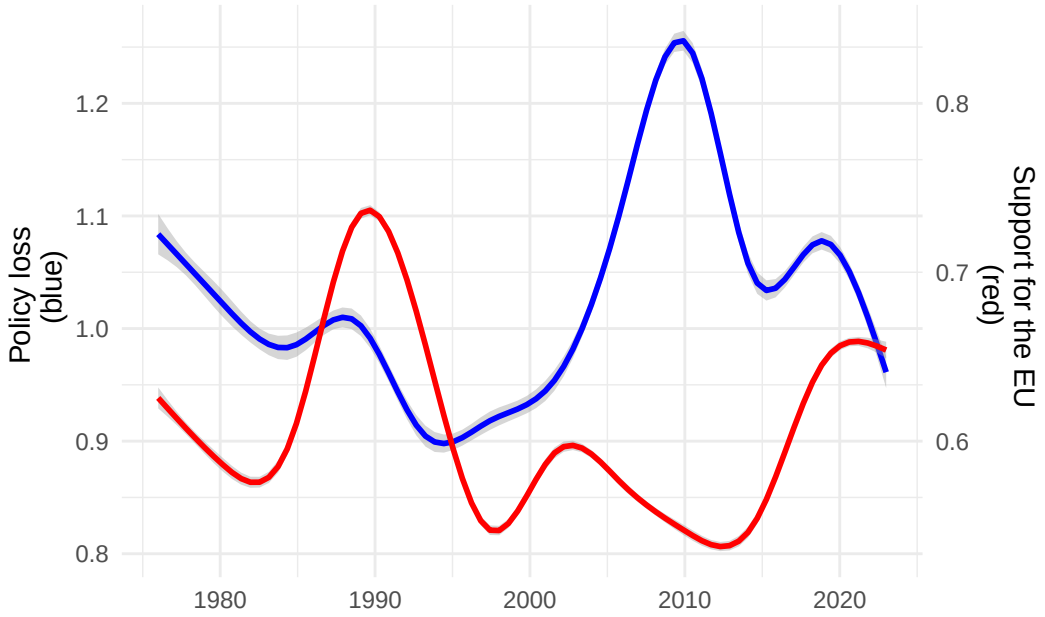
We are now ready to analyse the data.

```
library(modelsummary)
library(marginaleffects)
library(tidyverse)
library(data.table)
library(ggplot2)
library(ggthemes); library(ggExtra); library(broom)
library(fixest);library(ggeffects)
load("hh_data.RData")
source("helpers.R")
dir.create("figures")
dir.create("tables")
## average policy loss over time

factor <- 1.5


### Figure 3
hh_data |>
```

```r
ggplot(aes(x=year)) +
  geom_smooth(aes(y=policy_loss_Factor), method = "gam",
              formula = y ~ s(x, k = 12, bs = "cs"), col="blue") +
  geom_smooth(aes(y=proeu*factor), method = "gam",
              formula = y ~ s(x, k = 12, bs = "cs"), col="red") +
  scale_y_continuous(name = "Policy loss \n (blue)",
                     sec.axis = sec_axis(~ . / factor,
                                         name = "Support for the EU \n (red)")) +
  theme_minimal() +
  xlab("")
```




```r
ggsave("figures/policy_and_eu_support.png",
       width = 6, height = 4, units = "in", dpi = 450)

source("analysis.R")
```

All resulting tables and figures can be found in the tables and figures sub-folders generated by running this script.

```r
sessionInfo()
```

```
R version 4.5.2 (2025-10-31)
Platform: aarch64-apple-darwin20
```

```
Running under: macOS Tahoe 26.1

Matrix products: default
BLAS:   /System/Library/Frameworks/Accelerate.framework/Versions/A/Frameworks/vecLib.framewo
LAPACK: /Library/Frameworks/R.framework/Versions/4.5-arm64/Resources/lib/libRlapack.dylib;  

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

time zone: Europe/Oslo
tzcode source: internal

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
 [1] ggeffects_2.2.1        fixest_0.12.1          broom_1.0.8
 [4] marginaleffects_0.25.1 modelsummary_2.3.0     ggExtra_0.10.1
 [7] ggthemes_5.1.0         data.table_1.17.8      lubridate_1.9.4
[10] forcats_1.0.0          stringr_1.5.2          purrr_1.1.0
[13] readr_2.1.5            tidyr_1.3.1            tibble_3.3.0
[16] tidyverse_2.0.0        xtable_1.8-4           directlabels_2025.6.24
[19] dplyr_1.1.4            ggplot2_4.0.0          haven_2.5.4

loaded via a namespace (and not attached):
 [1] sjlabelled_1.2.0    tidyselect_1.2.1    viridisLite_0.4.2
 [4] farver_2.1.2        S7_0.2.0            fastmap_1.2.0
 [7] promises_1.3.2      digest_0.6.37       timechange_0.3.0
[10] estimability_1.5.1  mime_0.13           lifecycle_1.0.4
[13] dreamerr_1.5.0      magrittr_2.0.4      compiler_4.5.2
[16] rlang_1.1.6         tools_4.5.2         yaml_2.3.10
[19] knitr_1.50          texreg_1.39.4       labeling_0.4.3
[22] vtable_1.4.8        xml2_1.4.0          RColorBrewer_1.1-3
[25] miniUI_0.1.1.1      withr_3.0.2         numDeriv_2016.8-1.1
[28] datawizard_1.0.2    grid_4.5.2          emmeans_1.11.2-8
[31] scales_1.4.0        tinytex_0.57        insight_1.2.0
[34] cli_3.6.5           mvtnorm_1.3-3       rmarkdown_2.30
[37] ragg_1.4.0          generics_0.1.4      rstudioapi_0.17.1
[40] httr_1.4.7          tzdb_0.5.0          splines_4.5.2
[43] stringmagic_1.2.0   vctrs_0.6.5         Matrix_1.7-4
[46] sandwich_3.1-1      jsonlite_2.0.0      hms_1.1.3
[49] Formula_1.2-5       systemfonts_1.2.3   glue_1.8.0
[52] stringi_1.8.7       gtable_0.3.6        later_1.4.2
```

```
[55] tables_0.9.31       quadprog_1.5-8      pillar_1.11.1
[58] htmltools_0.5.8.1   R6_2.6.1            textshaping_1.0.0
[61] evaluate_1.0.5      shiny_1.10.0        kableExtra_1.4.0
[64] lattice_0.22-7      backports_1.5.0     httpuv_1.6.16
[67] Rcpp_1.1.0          svglite_2.2.1       coda_0.19-4.1
[70] nlme_3.1-168        mgcv_1.9-3          xfun_0.53
[73] zoo_1.8-14          pkgconfig_2.0.3
```