

EMI and polarization

David Garcia and Segun Aroyehun

pre: observations on EMI

```
library(dplyr)
library(lmtest)
library(sandwich)
library(tseries)
library(ggplot2)
library(ggrepel)
library(stargazer)

source("Functions.R")
output_path <- "output"

df <- read.csv("data/emi_congressw2v_prod_variables_public_laws.csv")

df_avg <- df %>% rowwise() %>% mutate(Avg_pol=mean(c(House_party.mean.diff.d1, Senate_party.mean.diff.d1)))
EMI <- df_avg$evidence_minus_intuition_score

peakyear <- df_avg$starting_year[EMI==max(EMI)]
peakyear

## [1] 1975
max(EMI)

## [1] 0.3582505
EMI[df_avg$starting_year==peakyear-2]

## [1] 0.3551387
mean(EMI)

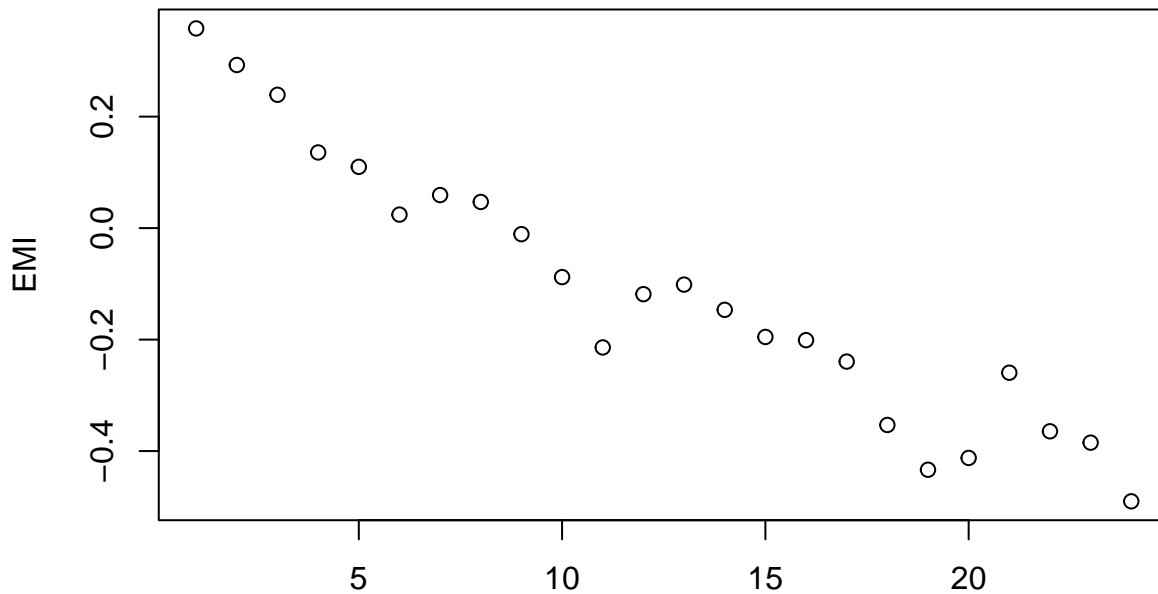
## [1] -0.01712737
sd(EMI)

## [1] 0.1736839
EMIits <- EMI[df_avg$starting_year>=peakyear]
t <- seq(0, length(EMIits)-1)
lin <- lm(EMIits~t)
summary(lin)

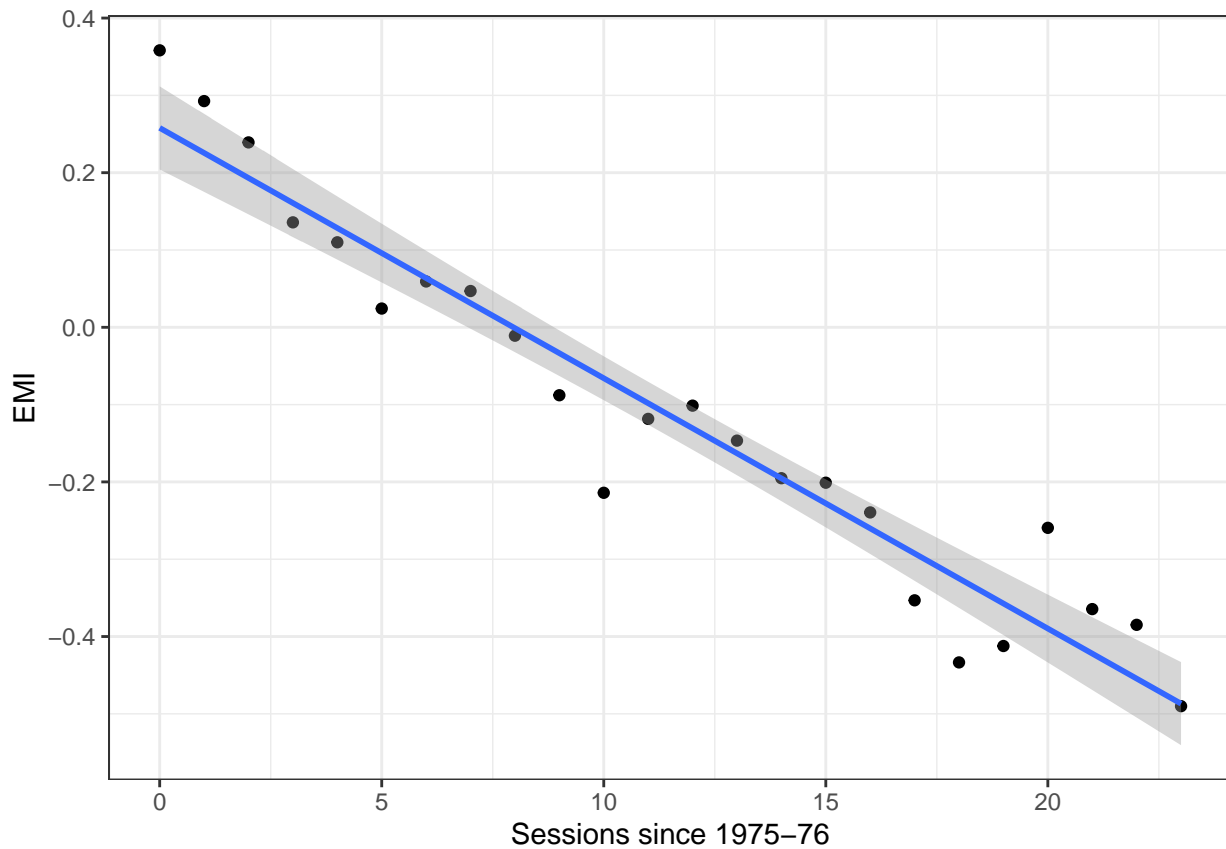
##
## Call:
## lm(formula = EMIits ~ t)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.148219 -0.032384 -0.001557  0.033471  0.130269
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.257925   0.025926   9.949 1.33e-09 ***
## t            -0.032380   0.001932 -16.764 5.14e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0655 on 22 degrees of freedom
## Multiple R-squared:  0.9274, Adjusted R-squared:  0.9241
## F-statistic: 281 on 1 and 22 DF, p-value: 5.141e-14
coeftest(lin, vcov. = vcovHAC(lin))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.2579248  0.0352554   7.3159 2.519e-07 ***
## t            -0.0323797  0.0026794 -12.0847 3.473e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
plot(EMIts, xlab = "", ylab="EMI")
```



```
ggplot(data.frame(EMIts, t), aes(x=t, y=EMIts)) +
  geom_point() +
  geom_smooth(method = "lm") +
  xlab("Sessions since 1975-76") +
  ylab("EMI") + theme_bw()
```



```
ggsave(filename = file.path(output_path, "EMI-trend.pdf"), width = 6.5, height = 3, dpi = 300, device="pdf")
```

1: Cross-correlations between EMI and polarization

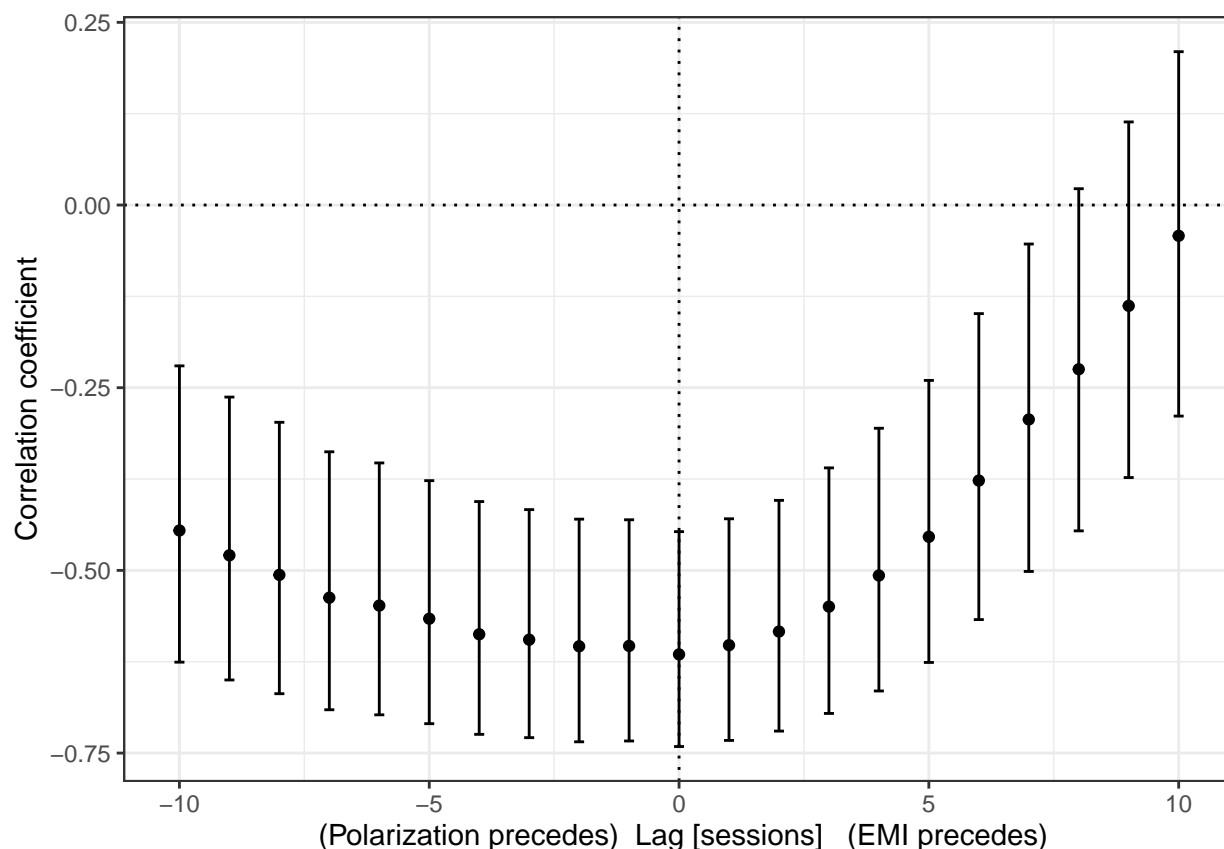
```
Pol <- df_avg$Avg_pol

lagmax <- 10
lags <- seq(-1*lagmax, lagmax)

corrs <- crosscor(Pol, EMI, lagmax)

plotdf <- data.frame(x=lags, y=corrs$corr, yminus = corrs$low, yplus= corrs$high)

ggplot(plotdf, aes(x=x, y=y)) +
  geom_point() + geom_vline(xintercept = 0, linetype="dotted") + geom_hline(yintercept = 0, linetype="dotted") +
  geom_errorbar(aes(ymin=yminus, ymax=yplus), width=.2) +
  xlab("(Polarization precedes) Lag [sessions] (EMI precedes)") +
  ylab("Correlation coefficient") + theme_bw()
```



```
ggsave(filename = file.path(output_path, "Polemi_corr.png"), width = 6.5, height = 3, dpi = 300)
ggsave(filename = file.path(output_path, "Polemi_corr.svg"), width = 6.5, height = 3, dpi = 300, device = "svg")
ggsave(filename = file.path(output_path, "Polemi_corr.pdf"), width = 6.5, height = 3, dpi = 300, device = "pdf")
```

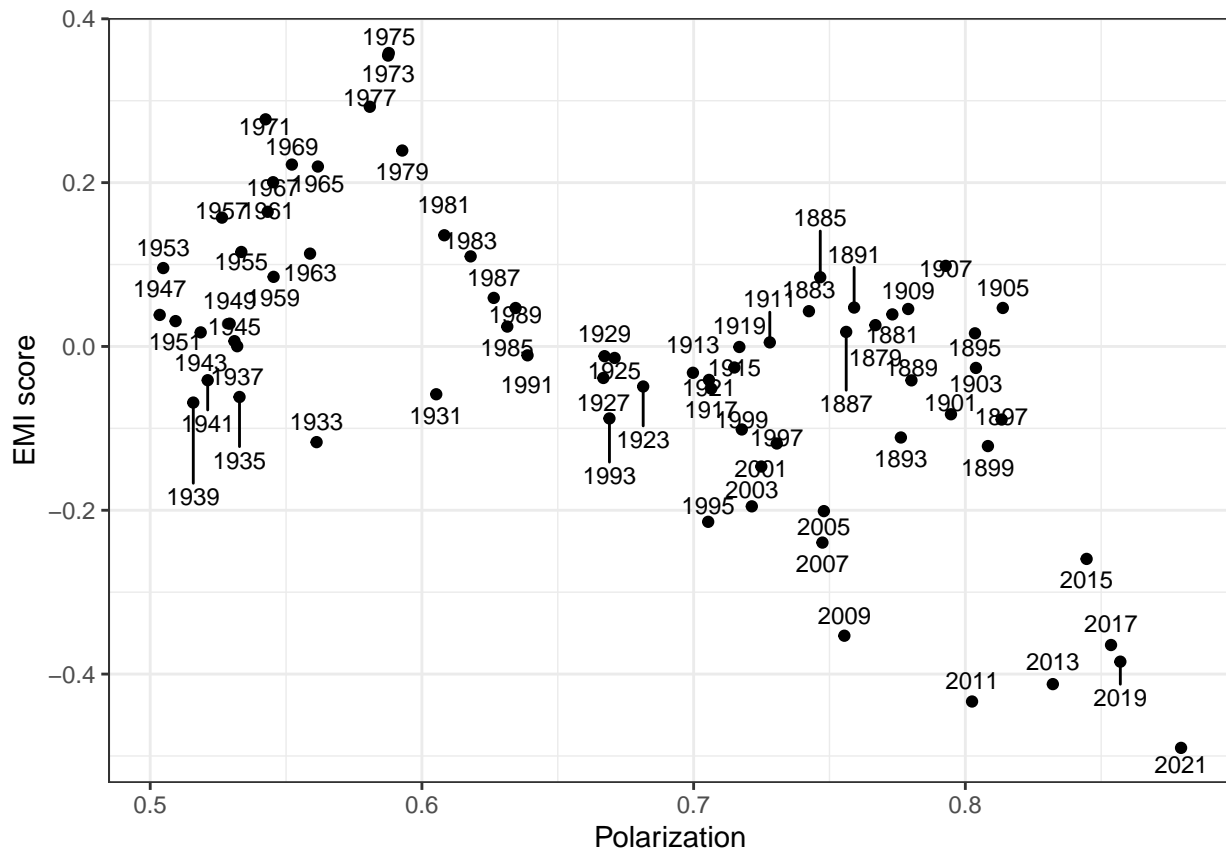
```
corrs$lag[abs(corrs$corr)==max(abs(corrs$corr))]
```

```
## [1] 0
```

```
cor.test(Pol, EMI)
```

```
##
## Pearson's product-moment correlation
##
## data: Pol and EMI
## t = -6.5246, df = 70, p-value = 9.083e-09
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.7410466 -0.4469583
## sample estimates:
## cor
## -0.6149511
```

```
plotdf <- data.frame(x=Pol, y=EMI, year=df$starting_year)
ggplot(plotdf, aes(x,y, label=year)) + geom_point() + geom_text_repel(direction="y",size=3) + xlab("Po")
```



```
ggsave(filename = file.path(output_path, "Polemi_scatter.png"), width = 6.5, height = 5, dpi = 300)
ggsave(filename = file.path(output_path, "Polemi_scatter.svg"), width = 6.5, height = 5, dpi = 300, dev = "svg")
ggsave(filename = file.path(output_path, "Polemi_scatter.pdf"), width = 6.5, height = 5, dpi = 300, dev = "pdf")
```

2: Regression models between EMI and polarization

```
attach(df_avg)
model_emipol0 <- lm(EMI ~ lag(EMI, 1))
model_emipol0_coefs <- coeftest(model_emipol0, vcov=vcovHAC(model_emipol0))
model_emipol0_coefs
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0074507  0.0070113  -1.0627  0.2916
## lag(EMI, 1)  0.9826538  0.0593096  16.5682  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
adf.test(residuals(model_emipol0))
```

```
##
## Augmented Dickey-Fuller Test
##
## data: residuals(model_emipol0)
## Dickey-Fuller = -3.3611, Lag order = 4, p-value = 0.06906
```

```

## alternative hypothesis: stationary
kpss.test(residuals(model_emipol0))

##
## KPSS Test for Level Stationarity
##
## data: residuals(model_emipol0)
## KPSS Level = 0.38428, Truncation lag parameter = 3, p-value = 0.08393
jarque.bera.test(residuals(model_emipol0))

##
## Jarque Bera Test
##
## data: residuals(model_emipol0)
## X-squared = 0.68717, df = 2, p-value = 0.7092
model_emipol <- lm(EMI ~ lag(EMI, 1) + Pol)
model_emipol_coefs <- coeftest(model_emipol, vcov=vcovHAC(model_emipol))
model_emipol_coefs

##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.094745   0.045290  2.0920  0.04018 *
## lag(EMI, 1)  0.921103   0.065218 14.1235 < 2e-16 ***
## Pol         -0.153168   0.070318 -2.1782  0.03286 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
adf.test(residuals(model_emipol))

##
## Augmented Dickey-Fuller Test
##
## data: residuals(model_emipol)
## Dickey-Fuller = -3.7934, Lag order = 4, p-value = 0.02404
## alternative hypothesis: stationary
kpss.test(residuals(model_emipol))

##
## KPSS Test for Level Stationarity
##
## data: residuals(model_emipol)
## KPSS Level = 0.52164, Truncation lag parameter = 3, p-value = 0.03679
jarque.bera.test(residuals(model_emipol))

##
## Jarque Bera Test
##
## data: residuals(model_emipol)
## X-squared = 0.8777, df = 2, p-value = 0.6448
waldtest(model_emipol0, model_emipol)

```

```

## Wald test
##
## Model 1: EMI ~ lag(EMI, 1)
## Model 2: EMI ~ lag(EMI, 1) + Pol
##   Res.Df Df       F   Pr(>F)
## 1      69
## 2      68   1 3.1156 0.08203 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

model_polemi0 <- lm(Pol ~ lag(Pol, 1))
model_polemi0_coefs <- coeftest(model_polemi0, vcov=vcovHAC(model_polemi0))
model_polemi0_coefs

##
## t test of coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.00015234  0.01505512 -0.0101    0.992
## lag(Pol, 1)  1.00259258  0.02305730 43.4827   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

adf.test(residuals(model_polemi0))

##
## Augmented Dickey-Fuller Test
##
## data:  residuals(model_polemi0)
## Dickey-Fuller = -2.886, Lag order = 4, p-value = 0.2147
## alternative hypothesis: stationary

kpss.test(residuals(model_polemi0))

## Warning in kpss.test(residuals(model_polemi0)): p-value smaller than printed
## p-value

##
## KPSS Test for Level Stationarity
##
## data:  residuals(model_polemi0)
## KPSS Level = 0.76365, Truncation lag parameter = 3, p-value = 0.01

jarque.bera.test(residuals(model_polemi0))

##
## Jarque Bera Test
##
## data:  residuals(model_polemi0)
## X-squared = 4.1157, df = 2, p-value = 0.1277

model_polemi <- lm(Pol ~ lag(Pol, 1) + EMI)
model_polemi_coefs <- coeftest(model_polemi, vcov=vcovHAC(model_polemi))
model_polemi_coefs

##
## t test of coefficients:
##

```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.018070   0.023192  0.7792   0.4386
## lag(Pol, 1)  0.974629   0.033279 29.2868   <2e-16 ***
## EMI          -0.028654   0.023882 -1.1998   0.2344
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

adf.test(residuals(model_polemi))

##
## Augmented Dickey-Fuller Test
##
## data: residuals(model_polemi)
## Dickey-Fuller = -2.5452, Lag order = 4, p-value = 0.3534
## alternative hypothesis: stationary

kpss.test(residuals(model_polemi))

##
## KPSS Test for Level Stationarity
##
## data: residuals(model_polemi)
## KPSS Level = 0.53749, Truncation lag parameter = 3, p-value = 0.03322

jarque.bera.test(residuals(model_polemi))

##
## Jarque Bera Test
##
## data: residuals(model_polemi)
## X-squared = 7.059, df = 2, p-value = 0.02932

waldtest(model_polemi0, model_polemi)

## Wald test
##
## Model 1: Pol ~ lag(Pol, 1)
## Model 2: Pol ~ lag(Pol, 1) + EMI
##   Res.Df Df      F Pr(>F)
## 1      69
## 2      68  1 2.6245 0.1099

stargazer(model_emipol0, model_emipol, model_polemi0, model_polemi, type = "latex", digits = 2, df = F,
           se=list(model_emipol0_coefs[,2], model_emipol_coefs[,2], model_polemi0_coefs[,2], model_polemi_coefs[,2]),
           p=list(model_emipol0_coefs[,4], model_emipol_coefs[,4], model_polemi0_coefs[,4], model_polemi_coefs[,4]),
           out="output/EMI-Pol-LM.tex")

##
## % Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac@sp.i.cas.cz
## % Date and time: Tue, Mar 26, 2024 - 17:23:02
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
##   \begin{tabular}{@{\extracolsep{5pt}}lcccc}
##     \\\[-1.8ex]\hline
##     \\\[-1.8ex]\hline
##     & \multicolumn{4}{c}{\textit{Dependent variable:}} & \\
##     \\\[-1.8ex]\hline

```



```
## \[-1.8ex] & \multicolumn{2}{c}{EMI} & \multicolumn{2}{c}{Pol} \\
## \[-1.8ex] & (1) & (2) & (3) & (4)\\
## \hline \[-1.8ex]
## lag(EMI, 1) & 0.98$^{***}$ & 0.92$^{***}$ & & \\
## & (0.06) & (0.07) & & \\
## & & & & \\
## Pol & & $-0.15$^{**}$ & & \\
## & & (0.07) & & \\
## & & & & \\
## lag(Pol, 1) & & & 1.00$^{***}$ & 0.97$^{***}$ \\
## & & & (0.02) & (0.03) \\
## & & & & \\
## EMI & & & $-0.03$ \\
## & & & (0.02) \\
## & & & \\
## Constant & $-0.01$ & 0.09$^{**}$ & $-0.0002$ & 0.02 \\
## & (0.01) & (0.05) & (0.02) & (0.02) \\
## & & & & \\
## \hline \[-1.8ex]
## Observations & 71 & 71 & 71 & 71 \\
## R$^{2}$ & 0.86 & 0.87 & 0.96 & 0.97 \\
## Adjusted R$^{2}$ & 0.86 & 0.87 & 0.96 & 0.97 \\
## Residual Std. Error & 0.06 & 0.06 & 0.02 & 0.02 \\
## F Statistic & 438.71$^{***}$ & 227.64$^{***}$ & 1,887.01$^{***}$ & 967.03$^{***}$ \\
## \hline
## \hline \[-1.8ex]
## \textit{Note:} & \multicolumn{4}{r}{\textit{$^{*}$p} < $0.1$; \textit{$^{**}$p} < $0.05$; \textit{$^{***}$p} < $0.01}} \\
## \end{tabular}
## \end{table}
```

```
stargazer(model_emipol0, model_emipol, model_polemi0, model_polemi, type = "text", digits = 2, df = F,
          se=list(model_emipol0_coefs[,2], model_emipol_coefs[,2], model_polemi0_coefs[,2], model_polemi_coefs[,2]),
          p=list(model_emipol0_coefs[,4], model_emipol_coefs[,4], model_polemi0_coefs[,4], model_polemi_coefs[,4]),
          out="output/EMI-Pol-LM.txt")
```

```
##
## =====
##
## Dependent variable:
## -----
##
## EMI
## (1) (2) (3) Pol (4)
## -----
## lag(EMI, 1) 0.98*** 0.92***
## (0.06) (0.07)
##
## Pol -0.15**
## (0.07)
##
## lag(Pol, 1) 1.00*** 0.97***
## (0.02) (0.03)
##
## EMI -0.03
## (0.02)
##
## Constant -0.01 0.09** -0.0002 0.02
```

```

##          (0.01)    (0.05)    (0.02)    (0.02)
##
## -----
## Observations      71      71      71      71
## R2                0.86    0.87    0.96    0.97
## Adjusted R2       0.86    0.87    0.96    0.97
## Residual Std. Error 0.06    0.06    0.02    0.02
## F Statistic      438.71*** 227.64*** 1,887.01*** 967.03***
## =====
## Note:                *p<0.1; **p<0.05; ***p<0.01

```