Models

Joshua Ashkinaze

2023-03-28

Load packages

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

##
## Attaching package: 'plm'

## The following objects are masked from 'package:dplyr':
##
##     between, lag, lead

##
## Please cite as:

##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union

## Loading required package: Matrix
```

Load Data

```
## New names:
## Rows: 16314 Columns: 24
## -- Column specification
## --------------------------------------------------------- Delimiter: "," chr
## (6): search_type, event, kw, index, kwe, period dbl (13): ...1, value,
## rumor_delta, announce_delta, rumor_announce_gap, stu... date (5): date,
## rumor_day, announce_day, max_date, min_date
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
## * `` -> `...1`
```

Modeling

- Point estimates are nearly identical between fixed efffects model with Newey West errors and nested rfx, so that's good
- Crossed rfx don't converge so we won't use that

Random Effects models

```r
# Make mixed model
model.crossed <- lmer(value ~ start_delta + year + month + period*search_type + (1 | kw) + (1|event), da
```

```
## boundary (singular) fit: see help('isSingular')
```

```r
model.nested <- lmer(value ~ start_delta + year + month + period*search_type + (1 | event/kw), data = d
```

Panel Model

```r
# Fit the fixed effects model and then get newey west standard errors
fem <- plm(value ~ period * search_type, data = df, model = "within", index = c("kwe", "date", "search_
```

```
## Warning in pdata.frame(data, index): duplicate couples (id-time) in resulting pdata.frame
##   to find out which, use, e.g., table(index(your_pdataframe), useNA = "ifany")
```

```r
fixed_ses <- summary(fem, vcov = vcovNW)
fem_robust_se <- fixed_ses$coefficients[, 2]
fem_p_values <- fixed_ses$coefficients[, 4]
```

Look at contrasts and graph emmeans

Contrasts

```r
# Look at contrasts:
# For rumors, is attention higher for google news and YT vs web?
# For announcements, is attention higher for web vs google news and YT?
em <- emmeans(model.nested, ~ period*search_type)
```

```
## Note: D.f. calculations have been disabled because the number of observations exceeds 3000.
## To enable adjustments, add the argument 'pbkrtest.limit = 16314' (or larger)
## [or, globally, 'set emm_options(pbkrtest.limit = 16314)' or larger];
## but be warned that this may result in large computation time and memory use.
```

```
## Note: D.f. calculations have been disabled because the number of observations exceeds 3000.
## To enable adjustments, add the argument 'lmerTest.limit = 16314' (or larger)
## [or, globally, 'set emm_options(lmerTest.limit = 16314)' or larger];
## but be warned that this may result in large computation time and memory use.
```

```r
em_df <- as.data.frame(em)
pairs <- pairs(em, by = "period", type = "response", rev = TRUE)
print(pairs)
```

```
## period = control:
##  contrast            estimate    SE  df z.ratio p.value
##  google_news - web      2.620 0.345 Inf   7.603  <.0001
```

```
##  youtube - web              2.846 0.345 Inf   8.258  <.0001
##  youtube - google_news    0.226 0.345 Inf   0.656  0.7891
##
## period = announce_period:
##  contrast              estimate    SE  df z.ratio p.value
##  google_news - web      -27.031 3.139 Inf  -8.610  <.0001
##  youtube - web          -32.844 3.139 Inf -10.462  <.0001
##  youtube - google_news   -5.812 3.139 Inf  -1.851  0.1531
##
## period = rumor_period:
##  contrast              estimate    SE  df z.ratio p.value
##  google_news - web       12.844 3.139 Inf   4.091  0.0001
##  youtube - web           11.938 3.139 Inf   3.802  0.0004
##  youtube - google_news   -0.906 3.139 Inf  -0.289  0.9551
##
## Degrees-of-freedom method: asymptotic
## P value adjustment: tukey method for comparing a family of 3 estimates
```

Graph

```r
# Let's graph the Search Type X Period emmeans
em_df$lower <- em_df$asymp.LCL
em_df$upper <- em_df$asymp.UCL
g <- ggplot(
  data = data.frame(em_df),
  aes(
    x = fct_relabel(period, .fun = relabel_func),
    y = emmean,
    fill = fct_relabel(search_type, .fun = relabel_func),
    ymin = lower,
    ymax = upper
  )
) +
  geom_bar(
    stat = "identity",
    position = position_dodge(width = 0.9),
    color = "black"
  ) +
  geom_errorbar(position = position_dodge(width = 0.9), width = 0.2) +
  labs(
    x = "Period",
    y = "Normalized Daily Search Volume",
    fill = "Search type",
    title = "People are more likely to turn to the web during announcements\nand more likely to turn to
    subtitle = paste(
      "Time series keyword search data for 26 U.S political events that had both\na rumor and official a
      nrow(df),
      " observations)\n\nPoint estimates and 95% CIs are estimated marginal means from mixed effects mod
    )
  ) +
  theme_nice() + scale_fill_manual(values = hex_color_list)
g
```

People are more likely to turn to the web during announcements and more likely to turn to platforms during rumors.

Time series keyword search data for 26 U.S political events that had both
a rumor and official announcement phase. (N = 16314 observations)

Point estimates and 95% CIs are estimated marginal means from mixed effects model.
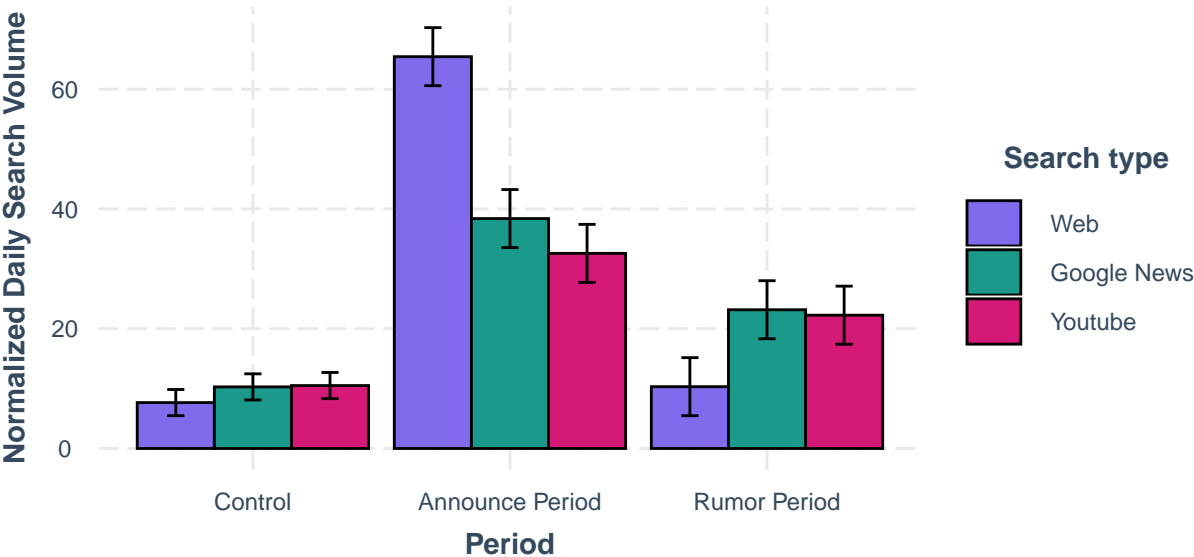

```
ggsave("model_results.png", dpi = 300)
```

```
## Saving 6.5 x 4.5 in image
```

Display models

```
models <- list(model.nested, fem)
model_names <- c("Nested Random Effects Model",
                 "Fixed Effect Model")

s <- stargazer(models,
        dep.var.labels = c("Normalized Attention"),
        model.names = TRUE,
        column.labels = model_names,
        type = 'text',
        se = list(NULL, fem_robust_se),
        p = list(NULL, fem_p_values))
```

```
##
## ============================================================================
##                                        Dependent variable:
##                              -----------------------------------------------
##                                        Normalized Attention
##                                   linear                       panel
##                                mixed-effects                   linear
##                           Nested Random Effects Model    Fixed Effect Model
##                                      (1)                        (2)
## ----------------------------------------------------------------------------
```

4

```
## start_delta                                       0.058***
##                                                   (0.006)
##
## year                                              1.949**
##                                                   (0.981)
##
## month                                             -0.058
##                                                   (0.099)
##
## periodannounce_period                             57.758***              58.520***
##                                                   (2.234)                (4.610)
##
## periodrumor_period                                 2.659                  1.895
##                                                   (2.234)                (2.561)
##
## search_typegoogle_news                             2.620***               2.620***
##                                                   (0.345)                (0.413)
##
## search_typeyoutube                                 2.846***               2.846***
##                                                   (0.345)                (0.414)
##
## periodannounce_period:search_typegoogle_news     -29.652***             -29.652***
##                                                   (3.158)                (6.439)
##
## periodrumor_period:search_typegoogle_news         10.223***              10.223**
##                                                   (3.158)                (4.696)
##
## periodannounce_period:search_typeyoutube         -35.690***             -35.690***
##                                                   (3.158)                (6.232)
##
## periodrumor_period:search_typeyoutube             9.091***               9.091**
##                                                   (3.158)                (4.602)
##
## Constant                                         -3,928.704**
##                                                  (1,979.393)
##
## ------------------------------------------------------------------------------------------------
## Observations                                      16,314                 16,314
## R2                                                                        0.062
## Adjusted R2                                                               0.058
## Log Likelihood                                   -70,195.290
## Akaike Inf. Crit.                                140,420.600
## Bayesian Inf. Crit.                              140,536.100
## F Statistic                                                              133.648*** (df = 8; 16242)
## ================================================================================================
## Note:                                                            *p<0.1; **p<0.05; ***p<0.01
```