panel

Joshua Ashkinaze

2023-03-28

Load packages

Load Data

```
##############################################
# LOAD DATA
##############################################
# read in data
df <- read_csv("https://raw.githubusercontent.com/josh-ashkinaze/attention/main/data/trend_merged_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`
```

```
df$year <- year(df$date)
df$month <- month(df$date)
df$week <- week(df$date)

# kwid is unique kw-id: (event, keyword, search_type)
df$kwid <- paste(paste(df$kw, df$event, "_"), df$search_type, "_")

# kwe is (keyword, event)
df$kwe <- paste(paste(df$kw, df$event, "_"))

df$search_type <- factor(df$search_type)
df$search_type <- relevel(df$search_type, ref = "web")
df$period <- factor(df$period)
df$period <- relevel(df$period, "control")
df$kw <- as.factor(df$kw)
df$event <- as.factor(df$event)
df$log_val <- log(df$value+1)
```

Modeling Linear Version

```
##############################################
# RANDOM FX VERSION
##############################################
# Make mixed model
model <- lmer(value ~ start_delta + year + month + period*search_type + (1 | event), data = df)



# 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, ~ period*search_type)
```

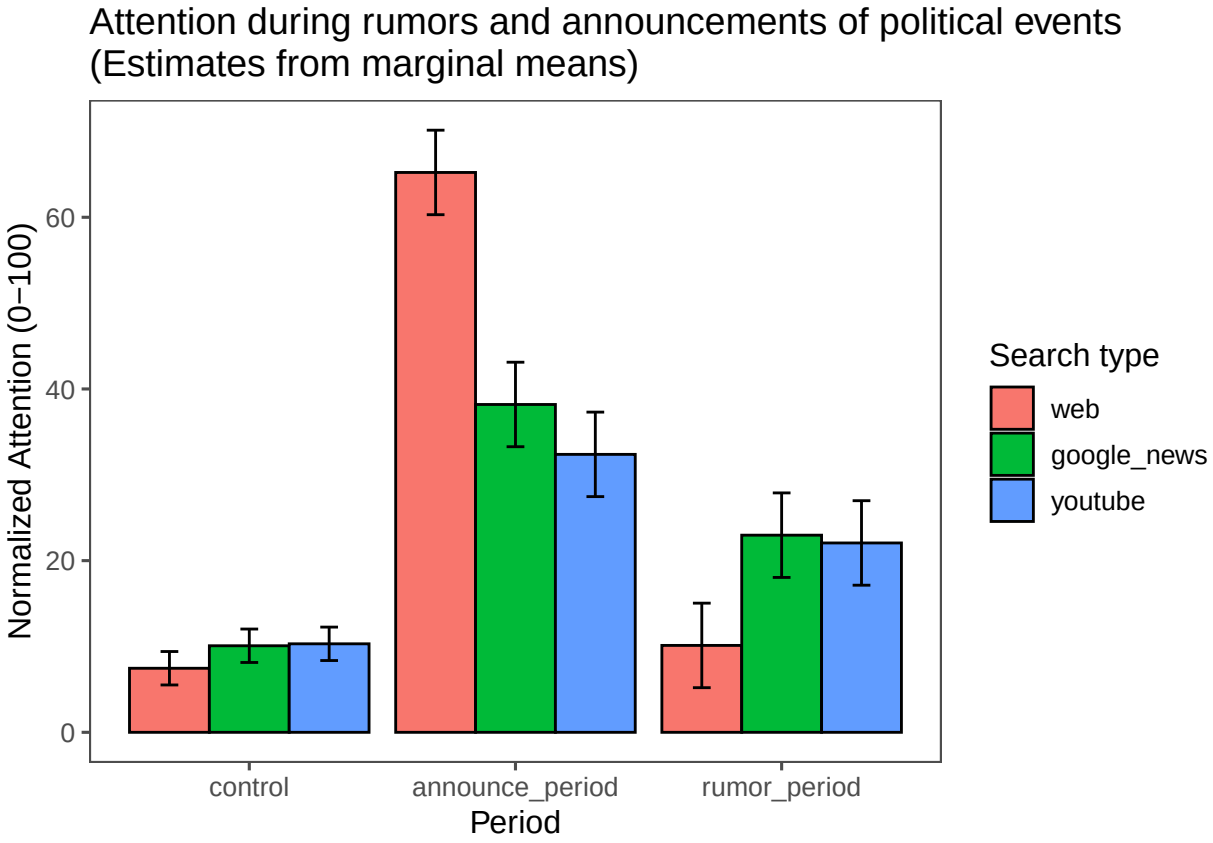
```
## Note: D.f. calculations have been disabled because the number of observations exceeds 3000.
## To enable adjustments, add the argument 'pbkrtest.limit = 16314' (or larger)
## [or, globally, 'set emm_options(pbkrtest.limit = 16314)' or larger];
## but be warned that this may result in large computation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

```
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 0.36 Inf   7.276  <.0001
##  youtube - web            2.846 0.36 Inf   7.904  <.0001
##  youtube - google_news    0.226 0.36 Inf   0.628  0.8050
##
## period = announce_period:
##  contrast              estimate   SE  df z.ratio p.value
##  google_news - web      -27.031 3.28 Inf  -8.240  <.0001
##  youtube - web          -32.844 3.28 Inf -10.012  <.0001
##  youtube - google_news   -5.812 3.28 Inf  -1.772  0.1791
##
## period = rumor_period:
##  contrast              estimate   SE  df z.ratio p.value
##  google_news - web       12.844 3.28 Inf   3.915  0.0003
##  youtube - web           11.938 3.28 Inf   3.639  0.0008
##  youtube - google_news   -0.906 3.28 Inf  -0.276  0.9588
##
## Degrees-of-freedom method: asymptotic
## P value adjustment: tukey method for comparing a family of 3 estimates
```

```
# Let's graph the Search Type X Period emmeans
em_df$lower <- em_df$asymp.LCL
em_df$upper <- em_df$asymp.UCL
ggplot(data=data.frame(em_df), aes(x=period, y=emmean, fill=search_type, 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 Attention (0-100)", fill="Search type") + theme_few() + ggtitle("Atten
```




```
##############################################
# PANEL MODEL VERSION
##############################################
# 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")
```

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


##############################################
# DISLAY MODELS
##############################################
stargazer(fem, model, type='text', se=list(fem_robust_se, NULL), p=list(fem_p_values, NULL))
```

```
##
## ================================================================================
##                                                  Dependent variable:
##                                              ---------------------------
##                                                         value
##                                                panel          linear
##                                                linear      mixed-effects
##                                                 (1)            (2)
## --------------------------------------------------------------------------------
## start_delta                                                     0.057***
##                                                                 (0.006)
##
## year                                                            2.041**
##                                                                 (0.897)
##
## month                                                           -0.047
##                                                                 (0.094)
##
## periodannounce_period                          58.520***        57.761***
##                                                 (4.610)          (2.335)
##
## periodrumor_period                               1.895            2.660
##                                                 (2.561)          (2.335)
##
## search_typegoogle_news                          2.620***         2.620***
##                                                 (0.413)          (0.360)
##
## search_typeyoutube                              2.846***         2.846***
##                                                 (0.414)          (0.360)
##
## periodannounce_period:search_typegoogle_news   -29.652***       -29.652***
##                                                 (6.439)          (3.300)
##
## periodrumor_period:search_typegoogle_news       10.223**         10.223***
##                                                 (4.696)          (3.300)
##
## periodannounce_period:search_typeyoutube        -35.690***       -35.690***
##                                                 (6.232)          (3.300)
##
## periodrumor_period:search_typeyoutube           9.091**          9.091***
##                                                 (4.602)          (3.300)
##
## Constant                                                        -4,112.901**
##                                                                 (1,810.850)
##
## --------------------------------------------------------------------------------
## Observations                                     16,314           16,314
## R2                                               0.062
## Adjusted R2                                      0.058
## Log Likelihood                                                  -70,835.280
## Akaike Inf. Crit.                                               141,698.600
## Bayesian Inf. Crit.                                             141,806.400
## F Statistic                          133.648*** (df = 8; 16242)
## ================================================================================
```

```
## Note:                                                *p<0.1; **p<0.05; ***p<0.01
```

Modeling - Negative Binoial

Make Model

```
library(glmmTMB)
```

```
## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.
## TMB was built with Matrix version 1.4.1
## Current Matrix version is 1.5.3
## Please re-install 'TMB' from source using install.packages('TMB', type = 'source') or ask CRAN for a
```

```
df$year2 <- scale(df$year)
df$start_delta2 <- scale(df$start_delta)
df$month2 <- scale(df$month)
model3 <- glmmTMB(value ~ start_delta2 + year2 + month2 + period*search_type, data = df, family=nbinom1)
```

Contrasts

```
# Interaction T1
em <- emmeans(model3, ~ period*search_type)
em_df <- as.data.frame(em)
pairs <- pairs(em, by = "search_type", type = "identity", rev = TRUE, adjust="tukey")
```

```
# Interaction T2
em3 <- emmeans(model3, ~ period | search_type)

# Print (rumor - control) pairwise contrasts
rumor_minus_control <- contrast(em3, list(rumor_minus_control = c(-1, 0, 1)), adjust = "none", type='res
pairs_rumor_contrasts <- pairs(rumor_minus_control, simple = "search_type", adjust = "tukey", rev=TRUE)
print(pairs_rumor_contrasts)
```

```
## contrast = rumor_minus_control:
##  contrast1             ratio    SE    df null t.ratio p.value
##  google_news / web     2.613 0.645 16301    1   3.888  0.0003
##  youtube / web         1.679 0.415 16301    1   2.098  0.0903
##  youtube / google_news 0.643 0.161 16301    1  -1.766  0.1810
##
## P value adjustment: tukey method for comparing a family of 3 estimates
## Tests are performed on the log scale
```

```
# Print (announce - control) pairwise contrasts
announce_minus_control <- contrast(em3, list(announce_minus_control = c(-1, 1, 0)), adjust = "none", typ
pairs_announce_contrasts <- pairs(announce_minus_control, simple = "search_type", adjust = "tukey", rev=
print(pairs_announce_contrasts)
```

```
## contrast = announce_minus_control:
##  contrast1             ratio    SE    df null t.ratio p.value
##  web / google_news      1.37 0.246 16301    1   1.752  0.1862
##  web / youtube          2.01 0.357 16301    1   3.916  0.0003
##  google_news / youtube  1.46 0.308 16301    1   1.815  0.1646
##
## P value adjustment: tukey method for comparing a family of 3 estimates
## Tests are performed on the log scale
```

```
announce_contrast_df <- data.frame(confint(announce_minus_control))
rumor_contrast_df <- data.frame(confint(rumor_minus_control))
contrast_df <- rbind(announce_contrast_df, rumor_contrast_df)

contrast_df <- contrast_df %>%
  mutate(contrast = recode(contrast,
                           "announce_minus_control" = "Announcement Attention\n÷\nControl Attention",
                           "rumor_minus_control" = "Rumor Attention\n÷\nControl Attention"))

ggplot(data=contrast_df, aes(x=contrast, y=ratio, fill=search_type, ymin=lower.CL, ymax=upper.CL)) +
  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="Shock Type", y="Ratio of Attention", fill="Search type") + theme_few() + ggtitle("Platform-le
```