Sources of Misperception - Preliminary Analyses

Patrick Kraft

January 3, 2020

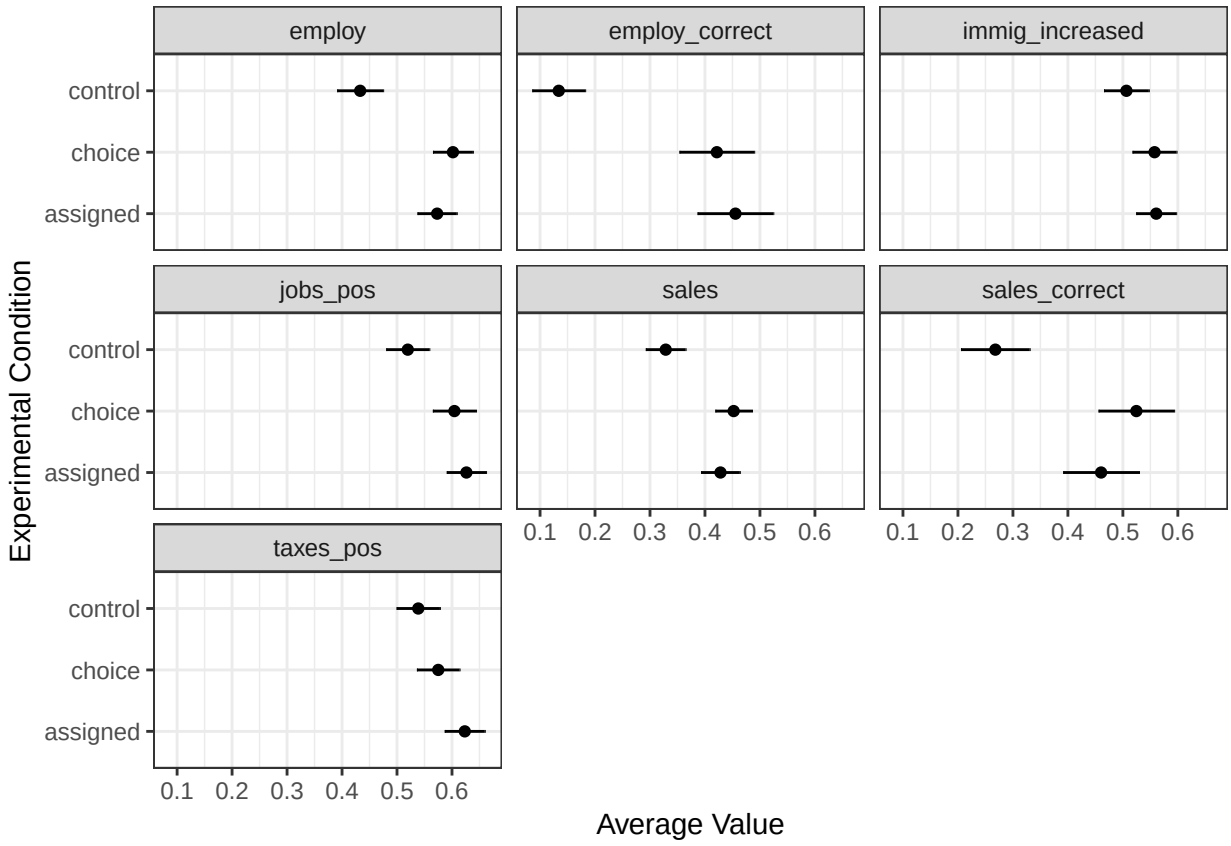
Load data, basic sociodemographics

```
### Packages
library(here)
library(readr)
library(tidyverse)

### Functions
se <- function(x, na.rm = T){
  sd(x, na.rm = na.rm)/sqrt(length(na.omit(x)))
}
cilo <- function(x, na.rm = T){
  mean(x, na.rm = na.rm) - qnorm(.975) * se(x, na.rm)
}
cihi <- function(x, na.rm = T){
  mean(x, na.rm = na.rm) + qnorm(.975) * se(x, na.rm)
}

### Load cleaned data
dat <- read_csv(here("data/immigration_20191219_clean.csv"))
```

Main analysis

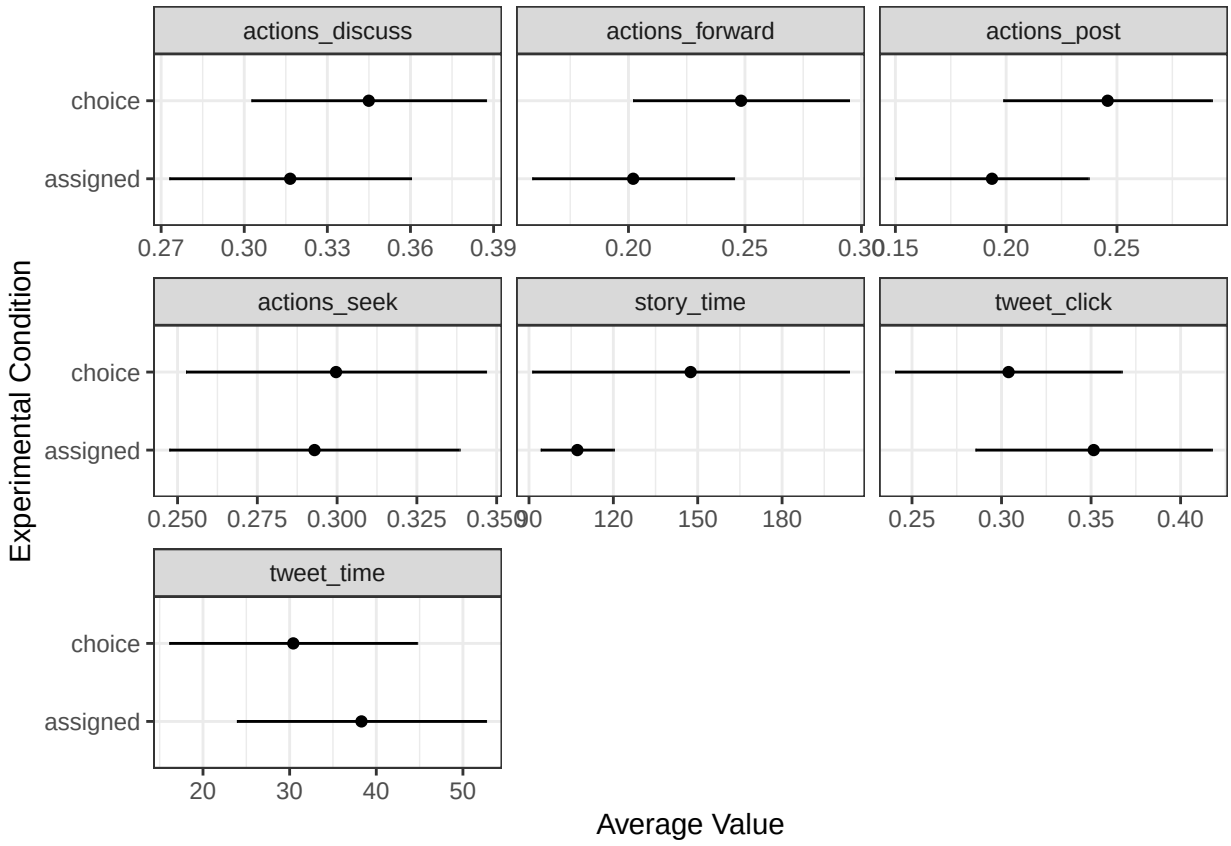
H1a-H1c: control vs. forced exposure vs. free choice

```
dat %>%
  select(condition, employ, employ_correct, sales, sales_correct,
         immig_increased, taxes_pos, jobs_pos) %>%
  gather(variable, value, -condition) %>%
  group_by(condition, variable) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T) %>%
  ggplot(aes(x = mean, xmin = cilo, xmax = cihi, y = condition)) +
  theme_bw() + labs(y = "Experimental Condition", x = "Average Value") +
  geom_point() + geom_errorbarh(height = 0) + facet_wrap(~variable)
```



Additional outcomes not measured in control condition

```
dat %>%
  filter(condition != "control") %>%
  select(condition, tweet_click, tweet_time, story_time,
         actions_discuss, actions_forward, actions_post, actions_seek) %>%
  gather(variable, value, -condition) %>%
  group_by(condition, variable) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T) %>%
  ggplot(aes(x = mean, xmin = cilo, xmax = cihi, y = condition)) +
  theme_bw() + labs(y = "Experimental Condition", x = "Average Value") +
  geom_point() + geom_errorbarh(height = 0) + facet_wrap(~variable, scales = "free_x")
```

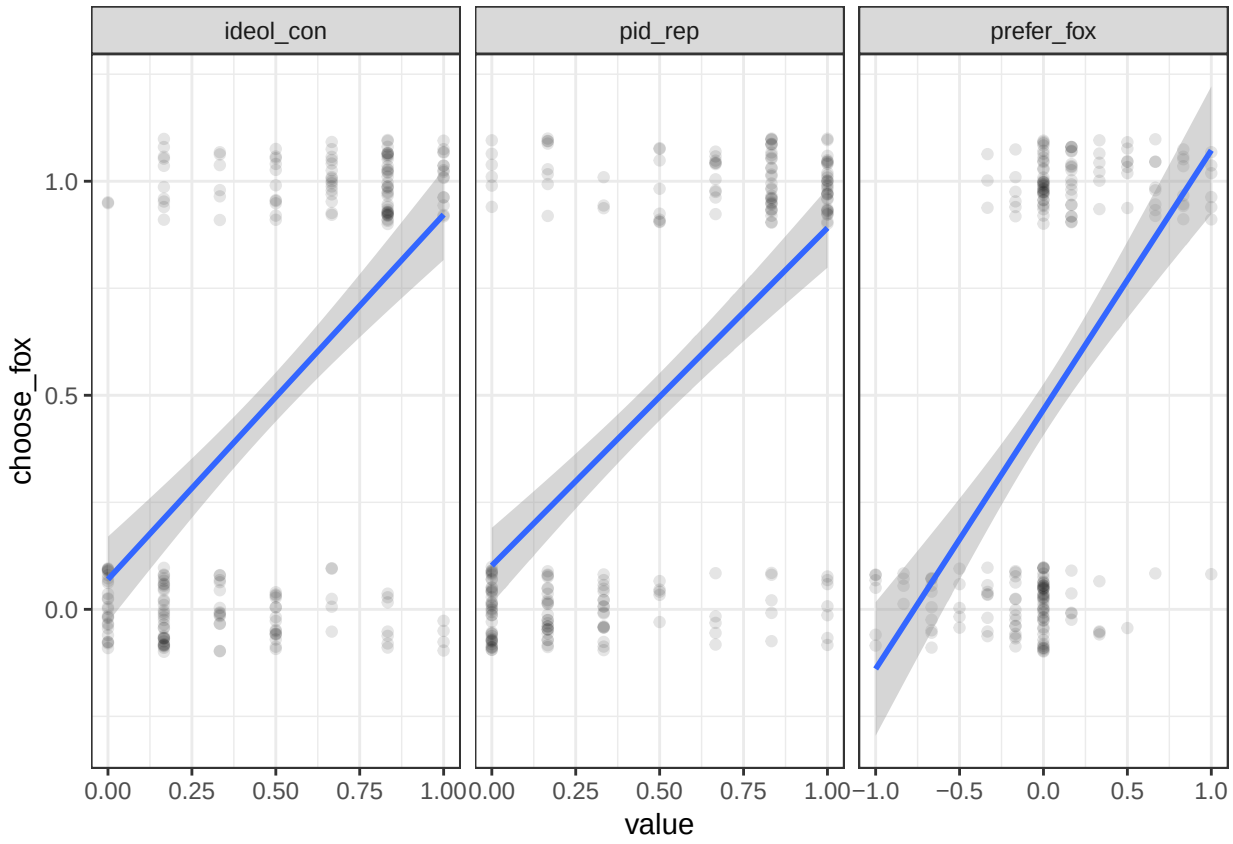


H2a-H2c: Determinants of media choice

```
dat %>%
  filter(condition == "choice") %>%
  mutate(prefer_fox = tv_fox - tv_msnbc,
         choose_fox = as.numeric(tweet == "fox")) %>%
  select(choose_fox, ideol_con, pid_rep, prefer_fox) %>%
  gather(variable, value, -choose_fox) %>%
  ggplot(aes(x = value, y = choose_fox)) +
  geom_jitter(alpha = .1, height = .1) + geom_smooth(method = "lm") +
  facet_grid(~variable, scales = "free_x") +
  theme_bw()
```

```
## Warning: Removed 4 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 4 rows containing missing values (geom_point).
```

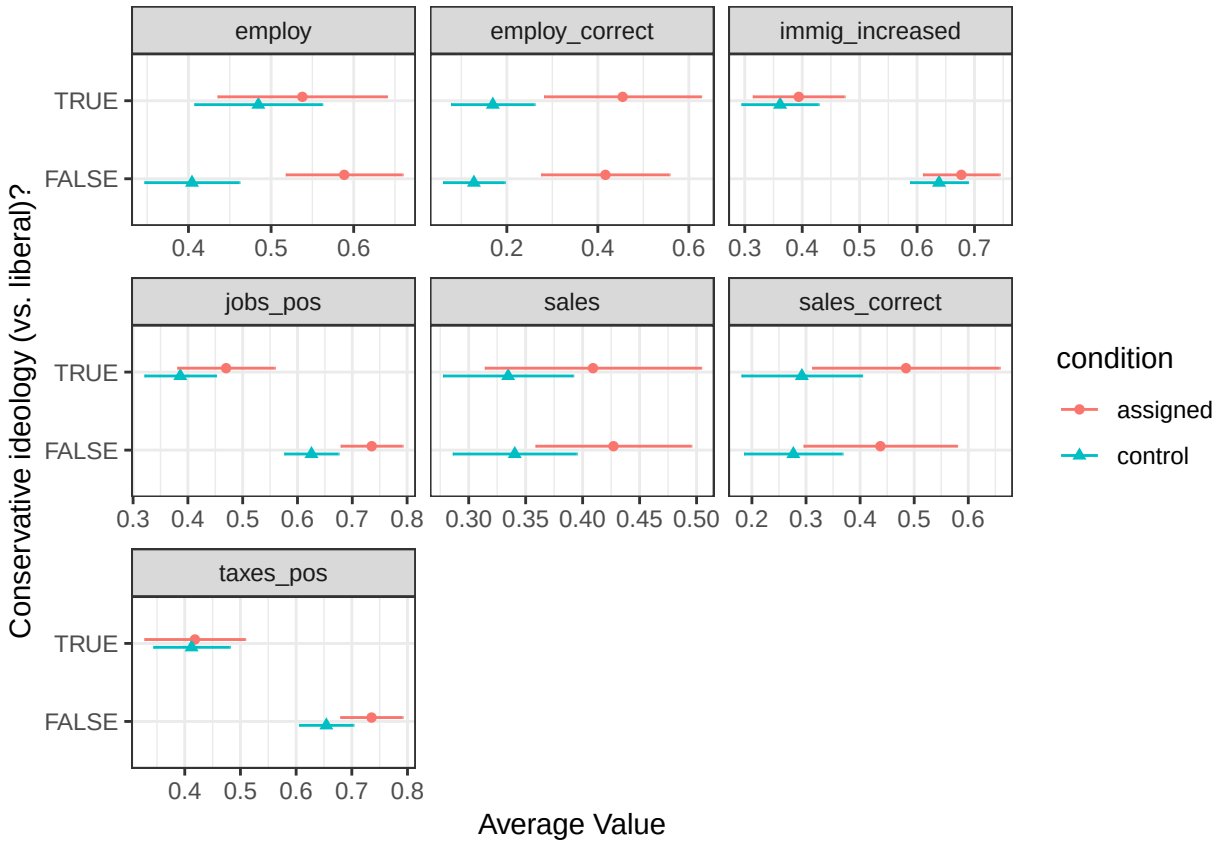


H3a-H3c: Effects of corrective information conditional on preferred media choice

Conditional effect of being assigned to Fox News

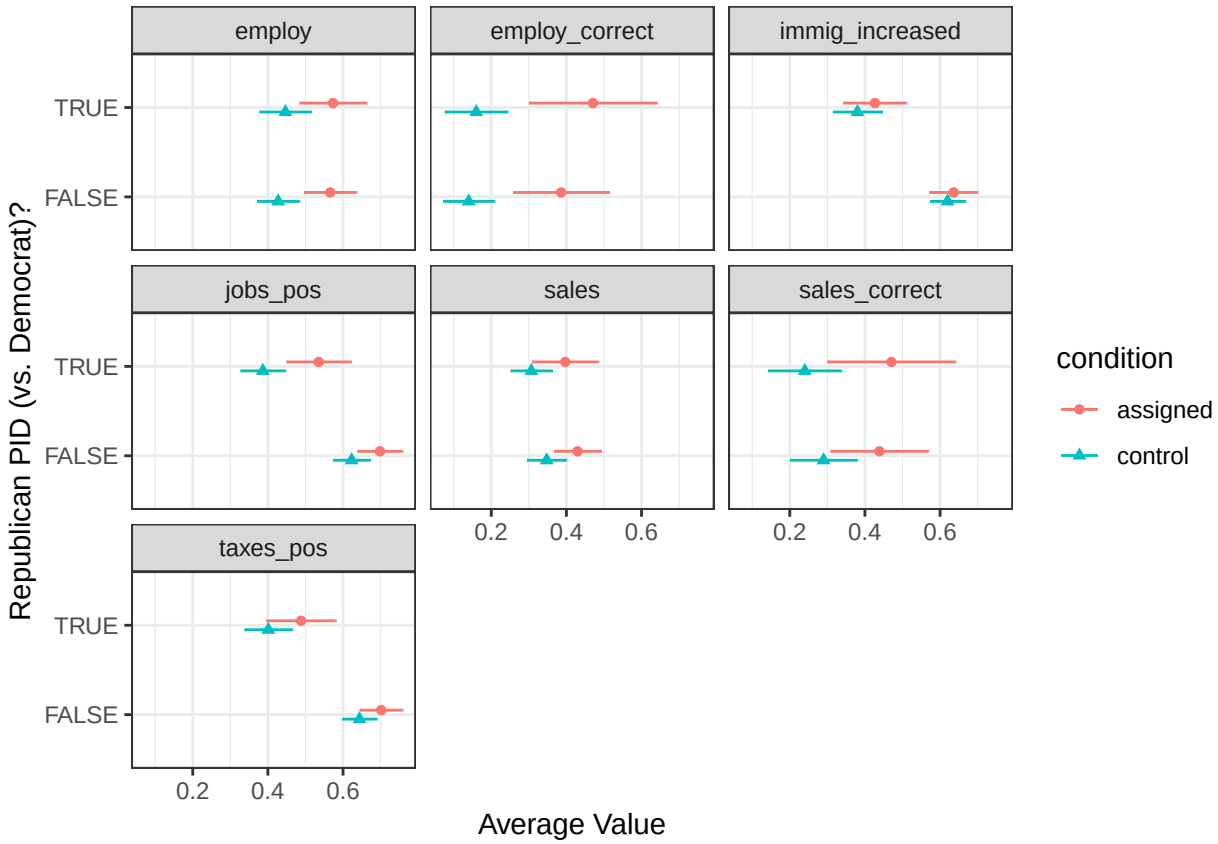
```
plot_df <- dat %>%
  filter((condition != "choice") & (tweet != "msnbc") & (ideol_con != .5)) %>%
  mutate(ideol_con = ideol_con>.5) %>%
  select(condition, employ, employ_correct, sales, sales_correct,
         immig_increased, taxes_pos, jobs_pos, ideol_con) %>%
  gather(variable, value, -condition, -ideol_con) %>%
  group_by(condition, variable, ideol_con) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = ideol_con,
                    shape = condition, col = condition)) +
  theme_bw() + labs(y = "Conservative ideology (vs. liberal)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))
  facet_wrap(~variable, scales = "free_x")
```

```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```



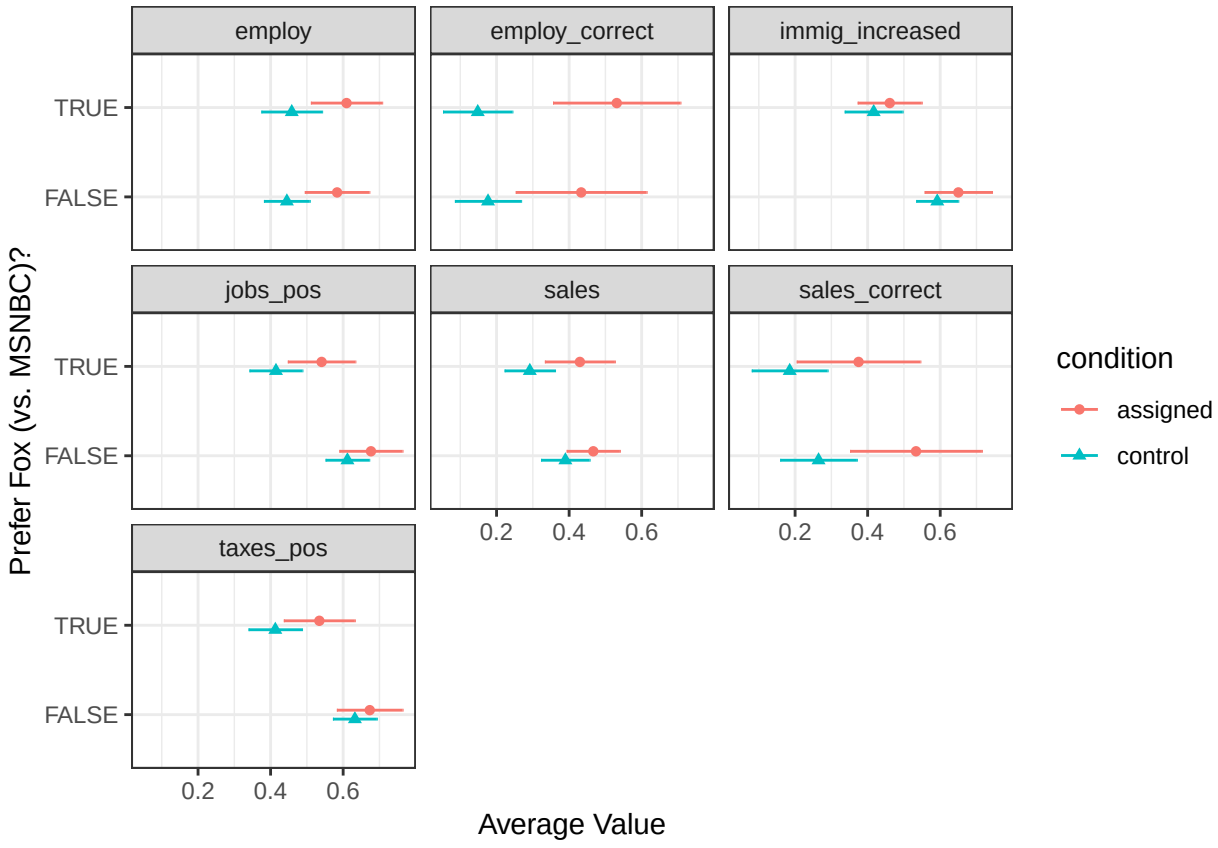

```
plot_df <- dat %>%
  filter((condition != "choice") & (tweet != "msnbc") & (pid_rep != .5)) %>%
  mutate(pid_rep = pid_rep>.5) %>%
  select(condition, employ, employ_correct, sales, sales_correct,
         immig_increased, taxes_pos, jobs_pos, pid_rep) %>%
  gather(variable, value, -condition, -pid_rep) %>%
  group_by(condition, variable, pid_rep) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = pid_rep,
                    shape = condition, col = condition)) +
  theme_bw() + labs(y = "Republican PID (vs. Democrat)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))
  facet_wrap(~variable)
```

```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```




```
plot_df <- dat %>%
  filter((condition != "choice") & (tweet != "msnbc") & ((tv_fox - tv_msnbc)!=0)) %>%
  mutate(prefer_fox = (tv_fox - tv_msnbc)>0) %>%
  filter(!is.na(prefer_fox)) %>%
  select(condition, employ, employ_correct, sales, sales_correct,
         immig_increased, taxes_pos, jobs_pos, prefer_fox) %>%
  gather(variable, value, -condition, -prefer_fox) %>%
  group_by(condition, variable, prefer_fox) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = prefer_fox,
                    shape = condition, col = condition)) +
  theme_bw() + labs(y = "Prefer Fox (vs. MSNBC)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))
  facet_wrap(~variable)
```

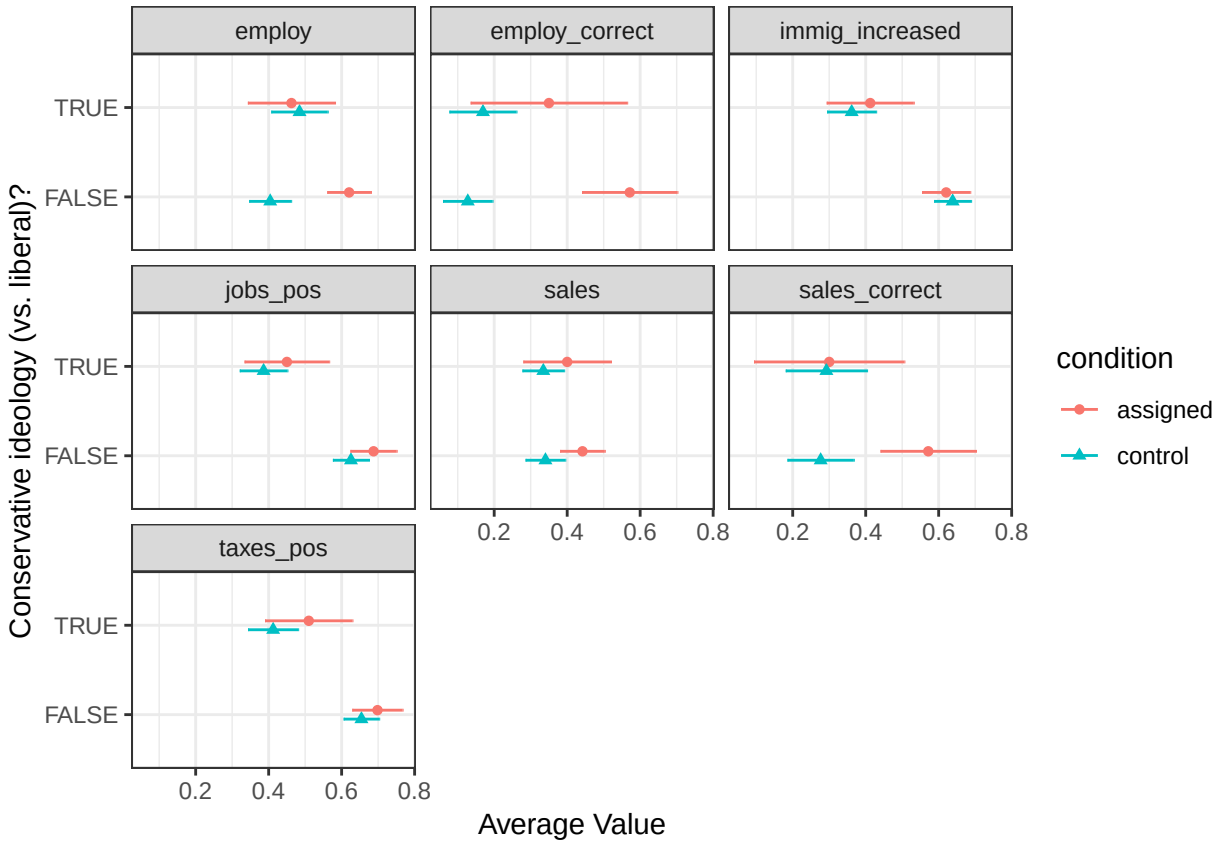
```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```



Conditional effect of being assigned to MSNBC

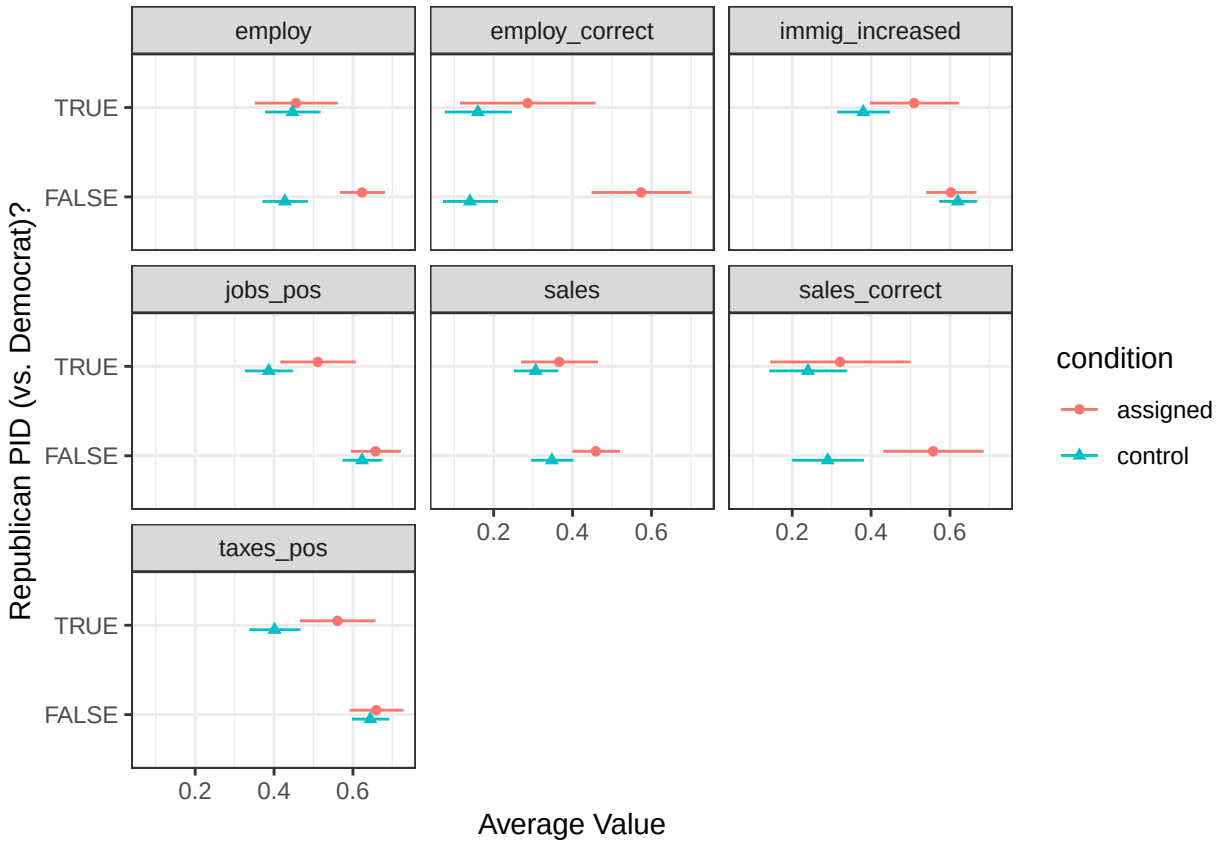
```
plot_df <- dat %>%
  filter((condition != "choice") & (tweet != "fox") & (ideol_con != .5)) %>%
  mutate(ideol_con = ideol_con>.5) %>%
  select(condition, employ, employ_correct, sales, sales_correct,
         immig_increased, taxes_pos, jobs_pos, ideol_con) %>%
  gather(variable, value, -condition, -ideol_con) %>%
  group_by(condition, variable, ideol_con) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = ideol_con,
                    shape = condition, col = condition)) +
  theme_bw() + labs(y = "Conservative ideology (vs. liberal)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))
  facet_wrap(~variable)
```

```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```



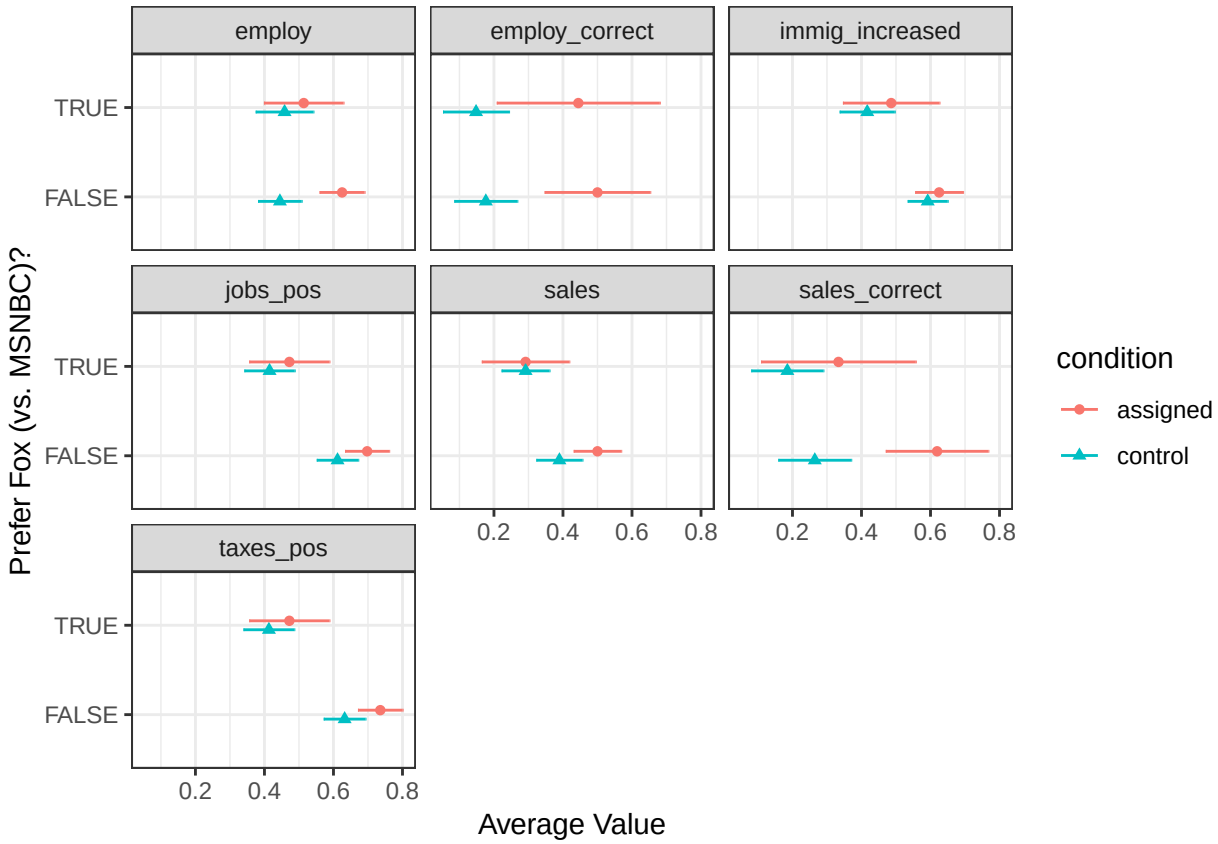

```
plot_df <- dat %>%
  filter((condition != "choice") & (tweet != "fox") & (pid_rep != .5)) %>%
  mutate(pid_rep = pid_rep>.5) %>%
  select(condition, employ, employ_correct, sales, sales_correct,
         immig_increased, taxes_pos, jobs_pos, pid_rep) %>%
  gather(variable, value, -condition, -pid_rep) %>%
  group_by(condition, variable, pid_rep) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = pid_rep,
                    shape = condition, col = condition)) +
  theme_bw() + labs(y = "Republican PID (vs. Democrat)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))
  facet_wrap(~variable)
```

```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```




```
plot_df <- dat %>%
  filter((condition != "choice") & (tweet != "fox") & ((tv_fox - tv_msnbc)!=0)) %>%
  mutate(prefer_fox = (tv_fox - tv_msnbc)>0) %>%
  filter(!is.na(prefer_fox)) %>%
  select(condition, employ, employ_correct, sales, sales_correct,
         immig_increased, taxes_pos, jobs_pos, prefer_fox) %>%
  gather(variable, value, -condition, -prefer_fox) %>%
  group_by(condition, variable, prefer_fox) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = prefer_fox,
                    shape = condition, col = condition)) +
  theme_bw() + labs(y = "Prefer Fox (vs. MSNBC)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$condition == "control"))
  facet_wrap(~variable)
```

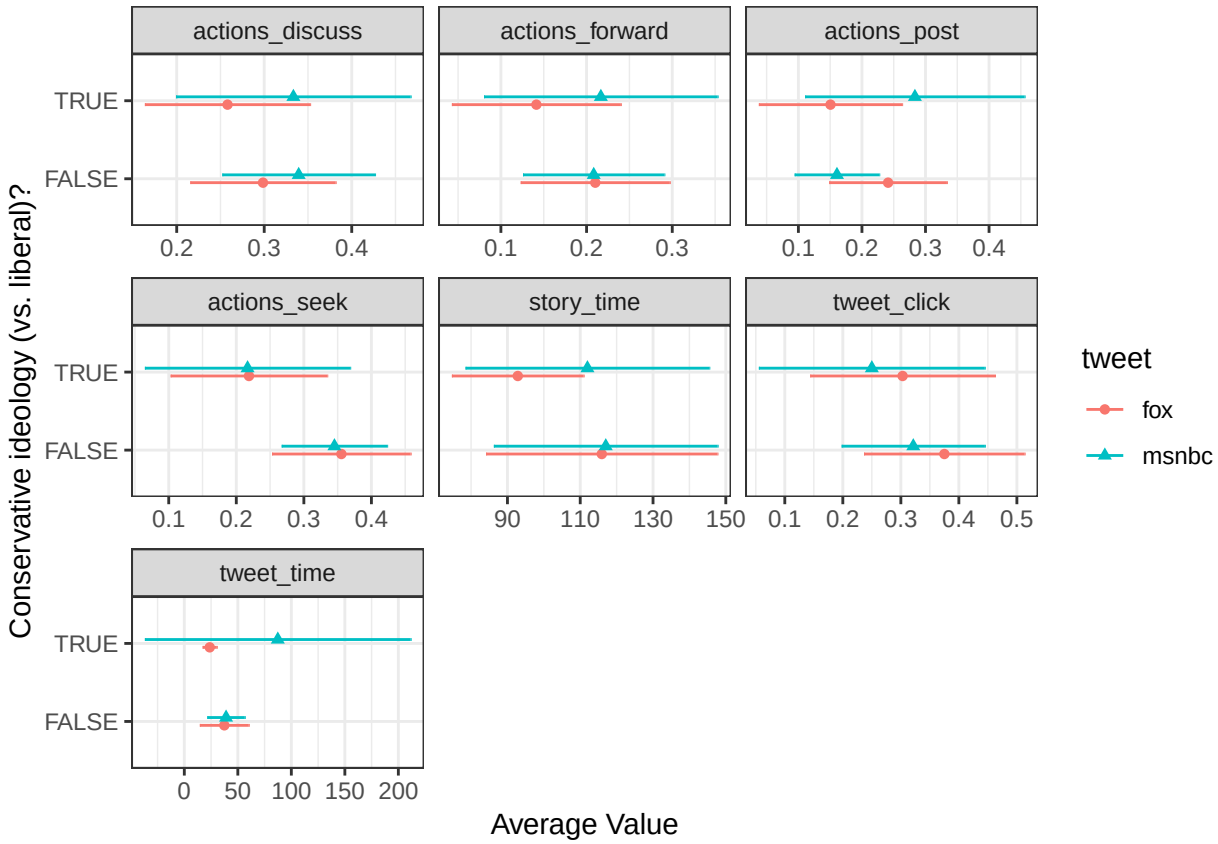
```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```



Exploratory analysis: predictors of engagement with article

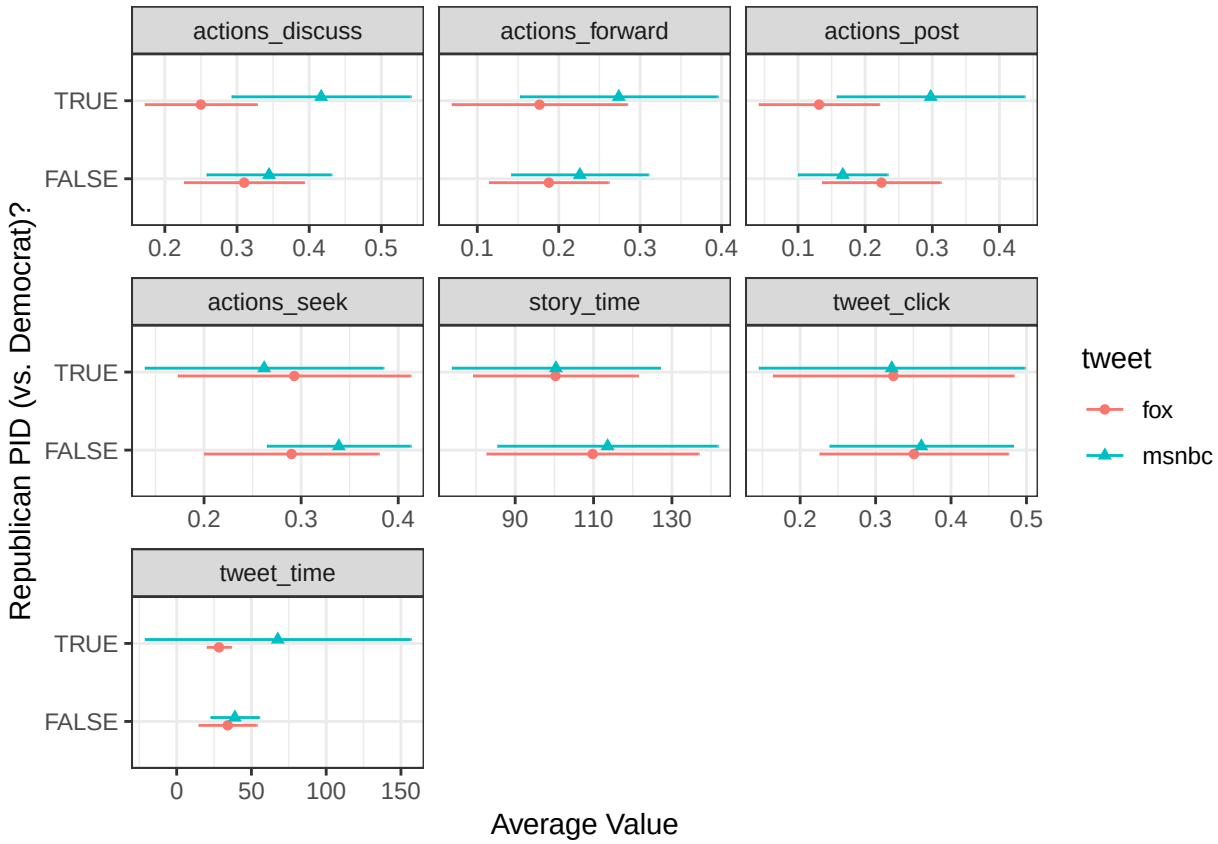
```
plot_df <- dat %>%
  filter((condition == "assigned") & (ideol_con != .5)) %>%
  mutate(ideol_con = ideol_con>.5) %>%
  select(ideol_con, tweet, tweet_click, tweet_time, story_time,
         actions_discuss, actions_forward, actions_post, actions_seek) %>%
  gather(variable, value, -tweet, -ideol_con) %>%
  group_by(tweet, variable, ideol_con) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = ideol_con,
                    shape = tweet, col = tweet)) +
  theme_bw() + labs(y = "Conservative ideology (vs. liberal)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$tweet == "fox"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$tweet == "fox"))) +
  facet_wrap(~variable, scales = "free_x")
```

```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```



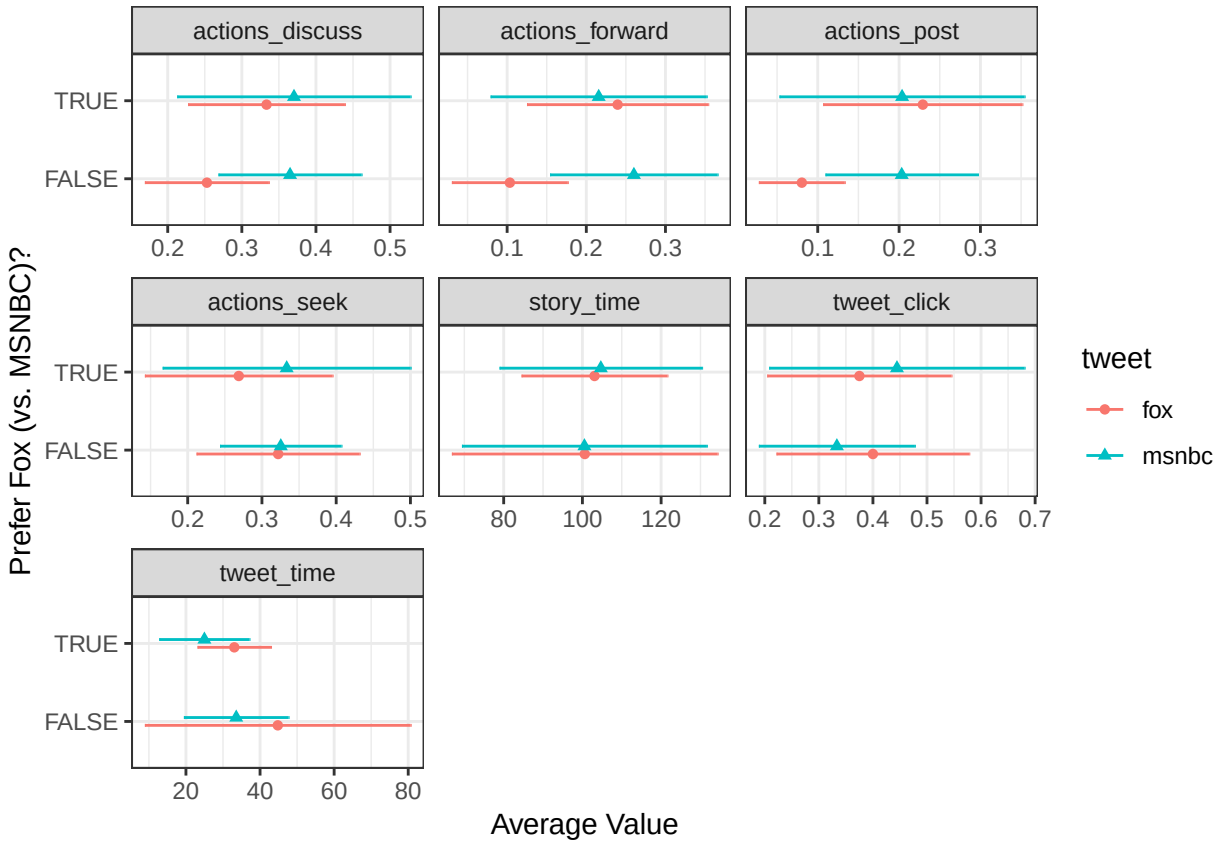

```
plot_df <- dat %>%
  filter((condition == "assigned") & (pid_rep != .5)) %>%
  mutate(pid_rep = pid_rep>.5) %>%
  select(pid_rep, tweet, tweet_click, tweet_time, story_time,
         actions_discuss, actions_forward, actions_post, actions_seek) %>%
  gather(variable, value, -tweet, -pid_rep) %>%
  group_by(tweet, variable, pid_rep) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = pid_rep,
                    shape = tweet, col = tweet)) +
  theme_bw() + labs(y = "Republican PID (vs. Democrat)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$tweet == "fox"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$tweet == "fox"))) +
  facet_wrap(~variable, scales = "free_x")
```

```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```




```
plot_df <- dat %>%
  filter((condition == "assigned") & ((tv_fox - tv_msnbc)!=0)) %>%
  mutate(prefer_fox = (tv_fox - tv_msnbc)>0) %>%
  filter(!is.na(prefer_fox)) %>%
  select(prefer_fox, tweet, tweet_click, tweet_time, story_time,
         actions_discuss, actions_forward, actions_post, actions_seek) %>%
  gather(variable, value, -tweet, -prefer_fox) %>%
  group_by(tweet, variable, prefer_fox) %>%
  summarize_all(list(mean = ~mean, cilo = ~cilo, cihi = ~cihi), na.rm = T)

ggplot(plot_df, aes(x = mean, xmin = cilo, xmax = cihi, y = prefer_fox,
                    shape = tweet, col = tweet)) +
  theme_bw() + labs(y = "Prefer Fox (vs. MSNBC)?", x = "Average Value") +
  geom_point(position = position_nudge(y = .05 - .1 * (plot_df$tweet == "fox"))) +
  geom_errorbarh(height = 0, position = position_nudge(y = .05 - .1 * (plot_df$tweet == "fox"))) +
  facet_wrap(~variable, scales = "free_x")
```

```
## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used

## Warning in if (params$y != 0) {: the condition has length > 1 and only the
## first element will be used
```



Next steps:

Estimate average choice-specific treatment effect (ACTE) following Knox et al. (2019)