Study S3

January 31, 2025

Items

Variable Names 4
Primary Analyses 5
- Race Feedback when Racial Minority Underrepresented in Candidate Set 5
- Race Feedback when Racial Minority Overrepresented in Candidate Set 6
- Race Feedback * Racial Minority Underrepresented 7

Robustness 8

```r
# Set this to TRUE if you have API access, FALSE if using CSV
USE_API <- T

if(USE_API) {
  ## Pull directly from Qualtrics API
  qual_data <- fetch_survey(surveyID='SV_3jZLUQYdzXtaJCu',
                    label = T,
                    convert = F,
                    start_date = "2024-07-05",
                    force_request = T)
} else {
  # Read the processed data directly from CSV
  d0 <- read.csv('StudyS3.csv', check.names = F)
  num_excluded <- unique(d0$num_excluded_total)
}

# Define the categories
URM <- c(
  "Cherelle Parker",
  "Eric Adams",
  "Justin Bibb",
  "London Breed",
  "Karen Bass",
  "Brandon Johnson",
  "Todd Gloria",
  "Eric Johnson",
  "Vi Lyles",
  "Victoria Woodards",
  "LaToya Cantrell",
  "Yadira Ramos-Herbert",
  "Francis Suarez",
  "Rex Richardson",
  "Yemi Mobolade",
  "Andre Dickens",
  "Regina Romero",
  "Quinton Lucas",
  "Cavalier Johnson",
  "Keith James",
  "Shawyn Patterson-Howard",
  "Tishaura Jones"
)

if(USE_API) {
  d0 <- qual_data |>
    filter(!is.na(`choice-7`), !is.na(workerId), Finished==1) |>
    mutate(
      race_pick = case_when(`choice-7` %in% URM ~ 1,
                            TRUE ~ 0),
           race_feedback = ifelse(cond=="treat", 1, 0),
      majority_pool = case_when(pool == 'Non-URM' ~ 1,
                        TRUE ~ 0),
      gender_code = case_when(gender=="Man" ~ 1, TRUE ~ 0),
      race_code = case_when(race=="White / Caucasian" ~ 1, TRUE ~ 0),
```

```
      base_race = rowSums(across(`choice-1`:`choice-6`, ~ . %in% URM))
    ) |>
    select(race_pick:race_code, gender, race, age, `choice-1`:`choice-7`)
  # Calculate the number of excluded participants
  num_excluded <- nrow(qual_data) - nrow(d0)

  # Save num_excluded in d0
  d0$num_excluded_total <- num_excluded  # As a column

  # Write the API-pulled data into a CSV file
  write.csv(d0, 'StudyS3.csv', row.names = FALSE, quote = TRUE)
}

# Create the pool-specific dataframes
d0_majority_pool <- d0 |>
  filter(majority_pool==1)

d0_minority_pool <- d0 |>
  filter(majority_pool==0)
```

Variable Names

Variable	Description
race_feedback	Binary indicator of whether a participant was randomly assigned to race feedback condition.
race_pick	Binary indicator of whether a participant selected a racial minority mayor for their seventh selection.
majority_pool	Binary indicator of whether a participant was randomly assigned to white-dominated mayor pool.
choice-1 to choice-7	The selected mayors.
gender	Self-selected gender.
race	Self-selected race.
age	Self-entered age.
gender_code	Dummy code for gender (male = 1).
race_code	Dummy code for race (white = 1).

Primary Analyses

Race Feedback when Racial Minority Underrepresented in Candidate Set

```
##
## Call:
## lm(formula = race_pick ~ race_feedback, data = d0_majority_pool)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3044 -0.3044 -0.1773 -0.1773  0.8227
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.17726    0.02216   7.999 6.54e-15 ***
## race_feedback  0.12709    0.03470   3.663 0.000272 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4235 on 596 degrees of freedom
## Multiple R-squared:  0.02209,	Adjusted R-squared:  0.02045
## F-statistic: 13.46 on 1 and 596 DF,  p-value: 0.0002653
```

```
##                    2.5 %    97.5 %
## (Intercept)   0.13373796 0.2207771
## race_feedback 0.05894699 0.1952336
```

Race Feedback when Racial Minority Overrepresented in Candidate Set

```
##
## Call:
## lm(formula = race_pick ~ race_feedback, data = d0_minority_pool)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8200  0.1800  0.1800  0.3154  0.3154
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.82000    0.02226   36.84  < 2e-16 ***
## race_feedback -0.13544    0.03500   -3.87 0.000121 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4269 on 596 degrees of freedom
## Multiple R-squared:  0.02462,    Adjusted R-squared:  0.02299
## F-statistic: 15.05 on 1 and 596 DF,  p-value: 0.0001166
```

```
##                    2.5 %      97.5 %
## (Intercept)    0.7762917  0.86370826
## race_feedback -0.2041690 -0.06670345
```

Race Feedback * Racial Minority Underrepresented

```
##
## Call:
## lm(formula = race_pick ~ race_feedback * majority_pool, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8200 -0.3044 -0.1773  0.3154  0.8227
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  0.82000    0.02226  36.845  < 2e-16 ***
## race_feedback               -0.13544    0.03500  -3.870 0.000115 ***
## majority_pool               -0.64274    0.03141 -20.466  < 2e-16 ***
## race_feedback:majority_pool  0.26253    0.04928   5.327 1.19e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4252 on 1192 degrees of freedom
## Multiple R-squared:  0.2791, Adjusted R-squared:  0.2773
## F-statistic: 153.8 on 3 and 1192 DF,  p-value: < 2.2e-16
```

```
##                                  2.5 %      97.5 %
## (Intercept)                  0.7763362  0.86366384
## race_feedback               -0.2040992 -0.06677331
## majority_pool               -0.7043593 -0.58112563
## race_feedback:majority_pool  0.1658380  0.35921508
```

Robustness

```r
## robust to demographic controls
### when racial minorities are underrepresented

r3 <- lm(race_pick ~ race_feedback + gender_code + race_code + age,
↪  data=d0_majority_pool)

# Display the robust_summary with robust standard errors
robust_summary(r3)
```

```
##
## Call:
## lm(formula = race_pick ~ race_feedback + gender_code + race_code +
##     age, data = d0_majority_pool)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3116 -0.3021 -0.1788 -0.1700  0.8303
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.1841118  0.0699436   2.632 0.008702 **
## race_feedback  0.1270400  0.0348947   3.641 0.000296 ***
## gender_code    0.0032855  0.0353841   0.093 0.926053
## race_code      0.0024909  0.0408005   0.061 0.951340
## age           -0.0002316  0.0013487  -0.172 0.863700
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4246 on 593 degrees of freedom
## Multiple R-squared:  0.02216,    Adjusted R-squared:  0.01556
## F-statistic: 3.359 on 4 and 593 DF,  p-value: 0.00986
```

```r
robust_confint(r3)
```

```
##                      2.5 %      97.5 %
## (Intercept)    0.046744605 0.321479085
## race_feedback  0.058507688 0.195572239
## gender_code   -0.066207949 0.072778891
## race_code     -0.077640260 0.082622042
## age           -0.002880524 0.002417258
```

```r
## logistic regression
# Fit the logistic regression model
r4 <- glm(race_pick ~ race_feedback, family = binomial, data=d0_majority_pool)

# Odds ratio
tidy_r4 <- tidy(r4, exponentiate = TRUE, conf.int = T)
print(tidy_r4)
```

```
## # A tibble: 2 x 7
##   term           estimate std.error statistic  p.value conf.low conf.high
##   <chr>             <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 (Intercept)       0.215     0.151    -10.1  3.80e-24    0.159     0.287
## 2 race_feedback     2.03      0.197      3.60 3.19e- 4    1.39      3.00
```

```r
summary(r4)
```

```
##
## Call:
## glm(formula = race_pick ~ race_feedback, family = binomial, data = d0_majority_pool)
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.5350     0.1514 -10.137  < 2e-16 ***
## race_feedback   0.7084     0.1968   3.599 0.000319 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 660.20  on 597  degrees of freedom
## Residual deviance: 646.86  on 596  degrees of freedom
## AIC: 650.86
##
## Number of Fisher Scoring iterations: 4
```

```r
## robust to demographic controls
### when racial minorities are overrepresented

r5 <- lm(race_pick ~ race_feedback + gender_code + race_code + age,
↪  data=d0_minority_pool)

# Display the robust_summary with robust standard errors
robust_summary(r5)
```

```
##
## Call:
## lm(formula = race_pick ~ race_feedback + gender_code + race_code +
##     age, data = d0_minority_pool)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8716  0.1233  0.1881  0.2899  0.3817
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.911678   0.069413  13.134  < 2e-16 ***
## race_feedback -0.136900   0.035395  -3.868 0.000122 ***
## gender_code   -0.045837   0.035584  -1.288 0.198208
## race_code     -0.008292   0.043551  -0.190 0.849062
## age           -0.001384   0.001437  -0.963 0.335791
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.427 on 593 degrees of freedom
## Multiple R-squared:  0.02898,    Adjusted R-squared:  0.02243
## F-statistic: 4.425 on 4 and 593 DF,  p-value: 0.001565
```

```r
robust_confint(r5)
```

```
##                     2.5 %      97.5 %
## (Intercept)    0.77535373  1.048002334
## race_feedback -0.20641421 -0.067385407
## gender_code   -0.11572367  0.024049942
## race_code     -0.09382567  0.077241461
## age           -0.00420504  0.001437474
```

```r
## logistic regression
# Fit the logistic regression model
r6 <- glm(race_pick ~ race_feedback, family = binomial, data=d0_majority_pool)

# Odds ratio
tidy_r6 <- tidy(r6, exponentiate = TRUE, conf.int = T)
print(tidy_r6)
```

```
## # A tibble: 2 x 7
##   term          estimate std.error statistic  p.value conf.low conf.high
##   <chr>            <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 (Intercept)      0.215     0.151     -10.1 3.80e-24    0.159     0.287
## 2 race_feedback    2.03      0.197       3.60 3.19e- 4    1.39      3.00
```

```r
summary(r6)
```

```
##
## Call:
## glm(formula = race_pick ~ race_feedback, family = binomial, data = d0_majority_pool)
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.5350     0.1514 -10.137  < 2e-16 ***
## race_feedback   0.7084     0.1968   3.599 0.000319 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 660.20  on 597  degrees of freedom
## Residual deviance: 646.86  on 596  degrees of freedom
## AIC: 650.86
##
## Number of Fisher Scoring iterations: 4
```

```r
### interaction model

r7 <- lm(race_pick ~ race_feedback*majority_pool + gender_code + race_code + age,
  data=d0)

# Display the robust_summary with robust standard errors
robust_summary(r7)
```

```
##
## Call:
## lm(formula = race_pick ~ race_feedback * majority_pool + gender_code +
##     race_code + age, data = d0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8471 -0.2962 -0.1477  0.3017  0.8601
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.8686581  0.0519233  16.730  < 2e-16 ***
## race_feedback              -0.1361216  0.0351647  -3.871 0.000114 ***
## majority_pool              -0.6443327  0.0315002 -20.455  < 2e-16 ***
## gender_code                -0.0212988  0.0250307  -0.851 0.394991
## race_code                  -0.0034564  0.0295889  -0.117 0.907027
## age                        -0.0007856  0.0009817  -0.800 0.423681
## race_feedback:majority_pool  0.2642208  0.0494410   5.344 1.09e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4255 on 1189 degrees of freedom
## Multiple R-squared:  0.2799, Adjusted R-squared:  0.2763
## F-statistic: 77.04 on 6 and 1189 DF,  p-value: < 2.2e-16
```

```r
robust_confint(r7)
```

```
##                                    2.5 %       97.5 %
## (Intercept)                  0.766786539  0.970529642
## race_feedback               -0.205113311 -0.067129926
## majority_pool               -0.706134881 -0.582530481
## gender_code                 -0.070408078  0.027810400
## race_code                   -0.061508559  0.054595793
## age                         -0.002711624  0.001140327
## race_feedback:majority_pool  0.167219473  0.361222037
```

```r
## logistic regression
# Fit the logistic regression model
r8 <- glm(race_pick ~ race_feedback*majority_pool, family = binomial, data=d0)

# Odds ratio
tidy_r8 <- tidy(r8, exponentiate = TRUE, conf.int = T)
print(tidy_r8)
```

```
## # A tibble: 4 x 7
##   term                  estimate std.error statistic  p.value conf.low conf.high
##   <chr>                    <dbl>     <dbl>     <dbl>    <dbl>    <dbl>     <dbl>
## 1 (Intercept)             4.56       0.150     10.1  6.10e-24   3.42      6.18
## 2 race_feedback           0.476      0.195     -3.80 1.46e- 4   0.323     0.696
## 3 majority_pool           0.0473     0.213    -14.3  2.11e-46   0.0308    0.0712
## 4 race_feedback:majori~   4.26       0.277      5.23 1.70e- 7   2.48      7.37
```

```r
summary(r8)
```

```
##
## Call:
## glm(formula = race_pick ~ race_feedback * majority_pool, family = binomial,
##     data = d0)
##
## Coefficients:
##                             Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   1.5163     0.1503  10.090  < 2e-16 ***
## race_feedback                -0.7415     0.1953  -3.798 0.000146 ***
## majority_pool                -3.0514     0.2133 -14.303  < 2e-16 ***
## race_feedback:majority_pool   1.4499     0.2772   5.230  1.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1658.0  on 1195  degrees of freedom
## Residual deviance: 1301.2  on 1192  degrees of freedom
## AIC: 1309.2
##
## Number of Fisher Scoring iterations: 4
```