Exercise_4

Liliana Tretyakova

2023-04-04

Analysis of patent examiner network and processing time

This R markdown file presents an analysis of patent examiner network and its relationship with processing time of patent applications. The data used in this analysis includes a sample of patent application data and network data of patent examiners.

Prepare

First, we load the required libraries such as tidyverse, lubridate, arrow, gender, and wru. Then, we load the data into R from the local directory.

```
# Load required libraries
library(tidyverse)
library(lubridate)
library(arrow)
library(gender)
library(wru)

# Load data
applications <- read_parquet("C:/Users/ulyan/OneDrive - McGill University/Documents/MMA/Winter II 2023/
edges_sample <- read_csv("C:/Users/ulyan/OneDrive - McGill University/Documents/MMA/Winter II 2023/Org
```

Add gender variable We predict gender based on the first name of each examiner using the gender library. We join the predicted gender data back to the main applications dataset.

```
# Get unique examiner first names
examiner_names <- applications %>% distinct(examiner_name_first)

# Predict gender based on first names
examiner_names_gender <- examiner_names %>%
  do(results = gender(.$examiner_name_first, method = "ssa")) %>%
  unnest(cols = c(results), keep_empty = TRUE) %>%
  select(examiner_name_first = name, gender, proportion_female)

# Join gender data back to the main applications dataset
applications <- applications %>%
  left_join(examiner_names_gender, by = "examiner_name_first")
```

Add race variable We predict the race of each examiner based on their last name using the wru library. We join the predicted race data back to the main applications dataset.

```
# Get unique examiner last names
examiner_surnames <- applications %>% select(surname = examiner_name_last) %>% distinct()

# Predict race based on last names
examiner_race <- predict_race(voter.file = examiner_surnames, surname.only = T) %>% as_tibble()
```

```
## Proceeding with last name predictions...
```

```
## i All local files already up-to-date!
```

```
## 701 (18.4%) individuals' last names were not matched.
```

```
# Select the race with the highest probability for each last name
examiner_race <- examiner_race %>%
  mutate(max_race_p = pmax(pred.asi, pred.bla, pred.his, pred.oth, pred.whi)) %>%
  mutate(race = case_when(
    max_race_p == pred.asi ~ "Asian",
    max_race_p == pred.bla ~ "black",
    max_race_p == pred.his ~ "Hispanic",
    max_race_p == pred.oth ~ "other",
    max_race_p == pred.whi ~ "white",
    TRUE ~ NA_character_
  ))

# Join race data back to the main applications dataset
applications <- applications %>%
  left_join(examiner_race, by = c("examiner_name_last" = "surname"))
```

Add tenure variable We calculate the tenure of each examiner by extracting the examiner IDs and application dates. We calculate the earliest and latest dates for each examiner and their tenure in days. We then join the tenure data back to the main applications dataset.

```
# Extract examiner IDs and application dates
examiner_dates <- applications %>%
  select(examiner_id, filing_date, appl_status_date)

# Convert dates to a consistent format
examiner_dates <- examiner_dates %>%
  mutate(start_date = ymd(filing_date), end_date = as_date(dmy_hms(appl_status_date)))

# Calculate the earliest and latest dates for each examiner and their tenure in days
examiner_dates <- examiner_dates %>%
  group_by(examiner_id) %>%
  summarise(
    earliest_date = min(start_date, na.rm = TRUE),
    latest_date = max(end_date, na.rm = TRUE),
    tenure_days = interval(earliest_date, latest_date) %/% days(1)
  ) %>%
  filter(year(latest_date)<2018)
```

```r
# Join tenure data back to the main applications dataset
applications <- applications %>% left_join(examiner_dates, by = "examiner_id")
```

Create Advice Networks and Calculate Centrality Scores

This part of the R code uses the igraph and dplyr libraries to create an advice network from the edges_sample data and calculate centrality scores for all examiners.

Load the required libraries - igraph and dplyr

```r
library(igraph)
library(dplyr)
```

Create a list of unique examiner IDs from both the ego_examiner_id and alter_examiner_id columns using the unique function.Then, create an igraph object from the edges_sample data, specifying vertex names as the unique examiner IDs. This code creates an igraph object g from the edges_sample data, with the ego_examiner_id and alter_examiner_id columns as edges, and the unique examiner IDs as vertices.

```r
unique_examiner_ids <- unique(c(edges_sample$ego_examiner_id, edges_sample$alter_examiner_id))

g <- graph_from_data_frame(edges_sample[, c("ego_examiner_id", "alter_examiner_id")], directed = TRUE, v
```

```
## Warning in graph_from_data_frame(edges_sample[, c("ego_examiner_id",
## "alter_examiner_id")], : In ‘d’ ‘NA’ elements were replaced with string "NA"
```

```
## Warning in graph_from_data_frame(edges_sample[, c("ego_examiner_id",
## "alter_examiner_id")], : In ‘vertices[,1]’ ‘NA’ elements were replaced with
## string "NA"
```

Calculate the degree, betweenness, and closeness centralities for the entire dataset using the degree, betweenness, and closeness functions in igraph.This creates a data frame centrality_entire with the examiner IDs, degree centrality, betweenness centrality, and closeness centrality for all examiners in the dataset.

```r
centrality_entire <- data.frame(
  examiner_id = V(g)$name,
  degree_centrality = degree(g, mode = "out"),
  betweenness_centrality = betweenness(g, directed = TRUE),
  closeness_centrality = closeness(g, mode = "out")
)
```

Convert examiner_id in centrality_entire to double using the as.numeric function.After that join the centrality data back to the main applications dataset using the left_join function in dplyr.This adds the degree centrality, betweenness centrality, and closeness centrality columns to the applications dataset based on the examiner_id column.

```r
centrality_entire$examiner_id <- as.numeric(centrality_entire$examiner_id)
```

```
## Warning: NAs introduced by coercion
```

```
applications <- applications %>%
  left_join(centrality_entire, by = "examiner_id")
```

1. Create variable for application processing time

This section of the code creates a new variable in the applications dataset that measures the number of days from the application filing date until the final decision on it, which could either be a patent issue or abandonment.

```
# Calculate the processing time
applications <- applications %>%
  mutate(
    final_decision_date = coalesce(patent_issue_date, abandon_date),
    app_proc_time = as.numeric(difftime(final_decision_date, filing_date, units = "days"))
  )
```

2. Linear regression models

Remove rows with missing values in degree, betweenness, or closeness centrality.

```
applications_clean <- applications %>%
  filter(!is.na(degree_centrality),
         !is.na(betweenness_centrality),
         !is.na(closeness_centrality))
```

```
# Estimate the linear regression model with degree_centrality as the independent variable
degree_model <- lm(
  app_proc_time ~ degree_centrality + gender + race + tenure_days,
  data = applications_clean
)

# Print the summary of the model
summary(degree_model)
```

```
## 
## Call:
## lm(formula = app_proc_time ~ degree_centrality + gender + race + 
##     tenure_days, data = applications_clean)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -2537.6  -442.5  -119.0   305.7  4933.2 
## 
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)    
## (Intercept)       1.543e+03  8.002e+00 192.856  < 2e-16 ***
## degree_centrality 1.509e-01  2.523e-02   5.980 2.24e-09 ***
## gendermale        2.716e+01  1.818e+00  14.937  < 2e-16 ***
## raceblack         4.762e+00  4.762e+00   1.000  0.31739    
## raceHispanic      1.599e+01  5.749e+00   2.781  0.00542 ** 
## raceother         9.462e+00  3.615e+01   0.262  0.79349    
## racewhite         -6.491e+01  1.925e+00 -33.726  < 2e-16 ***
## tenure_days       -4.627e-02  1.294e-03 -35.768  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 646.5 on 594077 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.005387,   Adjusted R-squared:  0.005376
## F-statistic: 459.7 on 7 and 594077 DF,  p-value: < 2.2e-16
```

The degree_model includes degree_centrality, gender, race, and tenure_days as independent variables, and app_proc_time as the dependent variable. The adjusted R-squared value of the model is 0.005376, which means that only about 0.54% of the variation in app_proc_time can be explained by the model. This is quite low, indicating that the model does not fit the data well.

```r
# Betweenness centrality linear regression model
betweenness_model <- lm(
  app_proc_time ~ betweenness_centrality + gender + race + tenure_days,
  data = applications_clean
)
summary(betweenness_model)
```

```
##
## Call:
## lm(formula = app_proc_time ~ betweenness_centrality + gender +
##     race + tenure_days, data = applications_clean)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2536.4  -442.3  -118.9   305.6  4934.9
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1.563e+03  7.946e+00 196.700  < 2e-16 ***
## betweenness_centrality  2.052e-03  1.191e-04  17.229  < 2e-16 ***
## gendermale              2.579e+01  1.819e+00  14.182  < 2e-16 ***
## raceblack               6.620e+00  4.761e+00   1.391  0.16437
## raceHispanic            1.726e+01  5.747e+00   3.004  0.00267 **
## raceother               1.046e+01  3.614e+01   0.289  0.77232
## racewhite              -6.388e+01  1.924e+00 -33.194  < 2e-16 ***
## tenure_days            -4.987e-02  1.298e-03 -38.404  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 646.3 on 594077 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.005824,   Adjusted R-squared:  0.005813
## F-statistic: 497.2 on 7 and 594077 DF,  p-value: < 2.2e-16
```

The betweenness_model includes betweenness_centrality, gender, race, and tenure_days as independent variables, and app_proc_time as the dependent variable. The adjusted R-squared value of the model is 0.005813, which is still low, suggesting that betweenness_centrality is not a good predictor of app_proc_time.

```
# Closeness centrality linear regression model
closeness_model <- lm(
  app_proc_time ~ closeness_centrality + gender + race + tenure_days,
  data = applications_clean
)
summary(closeness_model)
```

```
##
## Call:
## lm(formula = app_proc_time ~ closeness_centrality + gender +
##     race + tenure_days, data = applications_clean)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2563.8  -440.9  -118.6   305.4  5009.8
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.494e+03  7.975e+00 187.299  < 2e-16 ***
## closeness_centrality -1.175e+02  2.315e+00 -50.732  < 2e-16 ***
## gendermale            2.728e+01  1.814e+00  15.036  < 2e-16 ***
## raceblack             2.128e+01  4.761e+00   4.470 7.81e-06 ***
## raceHispanic          1.735e+01  5.735e+00   3.025  0.00248 **
## raceother            -4.591e+00  3.607e+01  -0.127  0.89871
## racewhite            -6.012e+01  1.922e+00 -31.288  < 2e-16 ***
## tenure_days          -3.257e-02  1.316e-03 -24.746  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 645.1 on 594077 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.009618,   Adjusted R-squared:  0.009607
## F-statistic: 824.2 on 7 and 594077 DF,  p-value: < 2.2e-16
```

The closeness_model includes closeness_centrality, gender, race, and tenure_days as independent variables, and app_proc_time as the dependent variable. The adjusted R-squared value of the model is 0.009607, which is slightly better than that of the betweenness_model, but still relatively low. This suggests that while closeness_centrality may have some predictive power for app_proc_time, it is not a strong predictor on its own.

```
##
## Call:
## lm(formula = app_proc_time ~ degree_centrality + betweenness_centrality +
##     closeness_centrality + gender + race + tenure_days, data = applications_clean)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2564.5  -440.7  -118.8   305.3  5009.5
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             1.509e+03  8.104e+00 186.189  < 2e-16 ***
## degree_centrality      -2.006e-01  2.617e-02  -7.666 1.77e-14 ***
## betweenness_centrality  9.939e-04  1.222e-04   8.132 4.22e-16 ***
## closeness_centrality   -1.181e+02  2.423e+00 -48.759  < 2e-16 ***
## gendermale              2.648e+01  1.816e+00  14.586  < 2e-16 ***
## raceblack               2.131e+01  4.763e+00   4.473 7.70e-06 ***
## raceHispanic            1.736e+01  5.737e+00   3.025  0.00248 **
## raceother              -5.187e+00  3.607e+01  -0.144  0.88565
## racewhite              -6.002e+01  1.923e+00 -31.205  < 2e-16 ***
## tenure_days            -3.484e-02  1.338e-03 -26.039  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 645 on 594075 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.009804,   Adjusted R-squared:  0.009789
## F-statistic: 653.6 on 9 and 594075 DF,  p-value: < 2.2e-16
```

The combined model (including degree, betweenness, and closeness centralities) has an adjusted R-squared of 0.009789, while the closeness_model has an adjusted R-squared of 0.009607. Although the combined model has a slightly higher adjusted R-squared, the improvement is marginal.

3. Does this relationship differ by examiner gender?

The part of the code consists of four linear regression models in R, each with a different independent variable (degree centrality, betweenness centrality, closeness centrality, or a combination of all three) and interaction with gender. The dependent variable in each model is app_proc_time, which represents the time it takes for an application to be processed.

```r
# Degree centrality model with interaction
degree_gender_interaction <- lm(
  app_proc_time ~ degree_centrality * gender + race + tenure_days,
  data = applications_clean
)
summary(degree_gender_interaction)
```

```
##
## Call:
## lm(formula = app_proc_time ~ degree_centrality * gender + race +
##     tenure_days, data = applications_clean)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2538.4  -442.7  -118.7   305.7  4939.9
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   1.533e+03  8.059e+00 190.228  < 2e-16 ***
## degree_centrality             6.129e-01  5.092e-02  12.037  < 2e-16 ***
## gendermale                    3.675e+01  2.037e+00  18.043  < 2e-16 ***
## raceblack                     5.074e+00  4.762e+00   1.065  0.28670
## raceHispanic                  1.807e+01  5.752e+00   3.142  0.00168 **
## raceother                     9.810e+00  3.614e+01   0.271  0.78607
## racewhite                    -6.512e+01  1.925e+00 -33.837  < 2e-16 ***
## tenure_days                  -4.578e-02  1.295e-03 -35.363  < 2e-16 ***
## degree_centrality:gendermale -6.103e-01  5.842e-02 -10.447  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 646.4 on 594076 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.00557,    Adjusted R-squared:  0.005557
## F-statistic:   416 on 8 and 594076 DF,  p-value: < 2.2e-16
```

The first model, degree_gender_interaction, shows that degree centrality is a statistically significant predictor of app_proc_time, with an estimated coefficient of 0.613. Gender is also a significant predictor, with male examiners taking longer to process applications than female examiners (coefficient of 36.75). There is a statistically significant interaction effect between degree centrality and gender, indicating that the relationship between degree centrality and app_proc_time depends on the gender of the examiner.

```
# Betweenness centrality model with interaction
betweenness_gender_interaction <- lm(
  app_proc_time ~ betweenness_centrality * gender + race + tenure_days,
  data = applications_clean
)
summary(betweenness_gender_interaction)
```

```
##
## Call:
## lm(formula = app_proc_time ~ betweenness_centrality * gender +
##     race + tenure_days, data = applications_clean)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2535.0  -442.4  -118.6   305.4  4931.7
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       1.568e+03  7.957e+00 197.055  < 2e-16 ***
## betweenness_centrality           -4.636e-05  2.182e-04  -0.212  0.83178
## gendermale                        2.081e+01  1.870e+00  11.127  < 2e-16 ***
## raceblack                         5.567e+00  4.761e+00   1.169  0.24230
## raceHispanic                      1.623e+01  5.748e+00   2.824  0.00474 **
## raceother                         1.180e+01  3.613e+01   0.326  0.74405
## racewhite                        -6.385e+01  1.924e+00 -33.182  < 2e-16 ***
## tenure_days                      -5.014e-02  1.299e-03 -38.613  < 2e-16 ***
## betweenness_centrality:gendermale  2.973e-03  2.592e-04  11.472  < 2e-16 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 646.3 on 594076 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.006045,   Adjusted R-squared:  0.006031
## F-statistic: 451.6 on 8 and 594076 DF,  p-value: < 2.2e-16
```

The second model, betweenness_gender_interaction, shows that betweenness centrality is not a statistically significant predictor of app_proc_time, with an estimated coefficient of -0.000046. Gender is again a significant predictor, with male examiners taking longer to process applications than female examiners (coefficient of 20.81). There is a statistically significant interaction effect between betweenness centrality and gender, indicating that the relationship between betweenness centrality and app_proc_time depends on the gender of the examiner.

```r
# Closeness centrality model with interaction
closeness_gender_interaction <- lm(
  app_proc_time ~ closeness_centrality * gender + race + tenure_days,
  data = applications_clean
)
summary(closeness_gender_interaction)
```

```
## 
## Call:
## lm(formula = app_proc_time ~ closeness_centrality * gender + 
##     race + tenure_days, data = applications_clean)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -2565.0  -440.8  -118.5   305.4  5002.4 
## 
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                      1.490e+03  8.082e+00 184.301  < 2e-16 ***
## closeness_centrality            -1.072e+02  4.025e+00 -26.629  < 2e-16 ***
## gendermale                       3.160e+01  2.283e+00  13.844  < 2e-16 ***
## raceblack                        2.072e+01  4.765e+00   4.349 1.37e-05 ***
## raceHispanic                     1.581e+01  5.757e+00   2.745  0.00604 ** 
## raceother                       -5.814e+00  3.607e+01  -0.161  0.87194    
## racewhite                       -6.022e+01  1.922e+00 -31.335  < 2e-16 ***
## tenure_days                     -3.237e-02  1.318e-03 -24.556  < 2e-16 ***
## closeness_centrality:gendermale -1.518e+01  4.861e+00  -3.122  0.00179 ** 
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
## 
## Residual standard error: 645.1 on 594076 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.009635,   Adjusted R-squared:  0.009621 
## F-statistic: 722.4 on 8 and 594076 DF,  p-value: < 2.2e-16
```

The third model, closeness_gender_interaction, shows that closeness centrality is a statistically significant predictor of app_proc_time, with an estimated coefficient of -107.2. Gender is also a significant predictor, with male examiners taking longer to process applications than female examiners (coefficient of 31.6). There is a statistically significant interaction effect between closeness centrality and gender, indicating that the relationship between closeness centrality and app_proc_time depends on the gender of the examiner.

```r
# Combined model with interaction
combined_gender_interaction <- lm(
  app_proc_time ~ (degree_centrality + betweenness_centrality + closeness_centrality) * gender + race +
  data = applications_clean
)
summary(combined_gender_interaction)
```

```
## 
```

```
## Call:
## lm(formula = app_proc_time ~ (degree_centrality + betweenness_centrality + 
##     closeness_centrality) * gender + race + tenure_days, data = applications_clean)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -2564.8  -440.7  -118.3   305.2  5002.2 
## 
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                       1.499e+03  8.351e+00 179.545  < 2e-16 ***
## degree_centrality                 2.636e-01  5.402e-02   4.880 1.06e-06 ***
## betweenness_centrality           -1.236e-03  2.222e-04  -5.563 2.66e-08 ***
## closeness_centrality             -1.034e+02  4.281e+00 -24.146  < 2e-16 ***
## gendermale                        3.630e+01  2.733e+00  13.281  < 2e-16 ***
## raceblack                         1.954e+01  4.766e+00   4.100 4.13e-05 ***
## raceHispanic                      1.634e+01  5.757e+00   2.839 0.004527 ** 
## raceother                        -4.789e+00  3.606e+01  -0.133 0.894363    
## racewhite                        -6.039e+01  1.924e+00 -31.395  < 2e-16 ***
## tenure_days                      -3.454e-02  1.342e-03 -25.744  < 2e-16 ***
## degree_centrality:gendermale     -6.084e-01  6.167e-02  -9.866  < 2e-16 ***
## betweenness_centrality:gendermale  3.139e-03  2.638e-04  11.900  < 2e-16 ***
## closeness_centrality:gendermale  -1.936e+01  5.144e+00  -3.764 0.000167 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
## 
## Residual standard error: 644.9 on 594072 degrees of freedom
##   (236232 observations deleted due to missingness)
## Multiple R-squared:  0.01018,    Adjusted R-squared:  0.01016 
## F-statistic:   509 on 12 and 594072 DF,  p-value: < 2.2e-16
```

Based on the output from the fourth (combined) model with gender interactions, it seems that the relationship between centrality measures and app_proc_time does indeed differ by examiner gender. The interaction terms for all three centrality measures with gender (degree_centrality:gendermale, betweenness_centrality:gendermale, and closeness_centrality:gendermale) are statistically significant with p-values less than 0.05:

1. degree_centrality:gendermale: Estimate = -6.084e-01, p-value < 2e-16
2. betweenness_centrality:gendermale: Estimate = 3.139e-03, p-value < 2e-16
3. closeness_centrality:gendermale: Estimate = -1.936e+01, p-value = 0.000167

These results suggest that the relationship between centrality measures and application processing time does differ between male and female examiners. In other words, the effect of centrality on app_proc_time is not consistent across examiner gender.

4. Discussion

The findings of this exercise suggest that there is a relationship between the centrality of patent examiners and the processing time of patent applications. Specifically, the results indicate that degree centrality, betweenness centrality, and closeness centrality are all weak predictors of application processing time, with adjusted R-squared values ranging from 0.005376 to 0.009789. However, when examining the relationship between centrality and processing time by examiner gender, the results suggest that this relationship is not consistent across gender. The interaction terms between gender and each of the three centrality measures are all statistically significant, indicating that the effect of centrality on processing time is different for male and female examiners.

These findings have important implications for the USPTO. First, the relatively weak relationship between centrality measures and application processing time suggests that other factors, beyond examiner centrality, are likely driving variations in processing time. Second, the results showing differences in the effect of centrality on processing time by gender raise concerns about potential inequities in the agency's decision-making process. This finding suggests that the USPTO may need to examine its policies and practices around examiner mobility, promotion, and attrition, and how they may differ by gender and other demographic characteristics. Addressing any potential inequities in these areas could help to reduce processing time and improve the agency's ability to support innovation and economic growth.