Predicting the Relationship between Marriage and Discussion of Women's Issues in Congressional Representatives: A Multiplicative Interaction Model Approach

Athanasia Meimeteas
Harrison High School
Harrison, NY

ACKNOWLEDGMENTS

My Research Teachers: Ms. Allison Blunt and Mr. Randy Gunnell for supporting me through my research, encouraging me, and sharing their experience and knowledge.

My Mentor: Professor Christopher Stout of Oregon State University for guiding me through my project, providing me with resources to learn the analysis I performed, and aiding me in troubleshooting throughout the process.

ABSTRACT

This study uses a multiplicative interaction model to predict the relationship between marital status and Congressional representatives' likelihood to discuss women's issues. Based on Stout, Kretschmer, and Ruppaner's (2017) findings, heterosexual marriage in women alters the perception of self-interest by institutionalizing partnerships with men and decreasing the feeling of connection to other women, it was hypothesized that married women would discuss women's issues less often than unmarried women. Alternately, heterosexual married men would support women's issues more than unmarried men because of their connection to their partner. Twitter was chosen as the data source for this study as it provides frequent, unfiltered representative communication. 167,784 tweets from 408 representatives in the 115th Congress were coded for discussing women's issues — defined as the discussion of sexual harassment, sexual assault, pay inequality, the recognition of women, human trafficking, women's health issues, gender disparities in profession, and international women's issues. A multiplicative interaction modeled the frequency of discussion of women's issues based on gender and marriage. The hypothesis was supported. The results showed that unmarried Congresswomen were significantly more likely to discuss women's issues than married Congresswomen ($p<.01$). Additionally, married Congressmen were significantly more likely than unmarried Congressmen to discuss women's issues ($p<.1$). The implications were that unmarried Congresswomen and married Congressmen advocated for women's issues more than their counterparts did. Being able to predict the behavior of future candidates may alter the political choices of voting citizens by allowing them to make more informed decisions.

TABLE OF CONTENTS

FIGURES

INTRODUCTION

The term gender inequality acknowledges that men and women are not equal and that gender affects an individual's living experience, in particular, through unequal access to resources and opportunities. These differences can arise from a number of reasons, one of which is cultural norms.

Recent research exploring women's views on gender equality has focused on race, education, socioeconomic status, age, and political party. Little research has been done, however, on one of the most substantial social constructs, marriage, and how it affects views on gender inequality. Of the current research on marriage, the impact marriage has on an individual's or a family's financial stability has been explored. The U.S. economy encourages marriage by offering certain tax benefits and program entitlements for married couples (Murphy, 2019). Although patterns of marriage diverge with demographics, the Urban Institute report (2019) found that among other things, marital status altered eligibility for entitlement programs, tax rates, and the availability of social safety nets.

Marriage rates, however, are predicted to drop below 70 percent in the near future. Baby boomers have a marriage rate of 91 percent, late boomers have a marriage rate of 87 percent marriage, and Gen Xers have a marriage rate of 82 percent (Murphy, 2019). According to Pew Research (2019), up to 25 percent of millennials may never get married. Therefore, the growing rate of unmarried Americans may have ramifications the U.S. economy is unprepared to handle (Murphy, 2019). This research, however, did not address any social factors influenced by marital status.

One analysis of marriage's effect on social phenomena determined that gender linked fate in females — one's identification with a group and the perception that one's life chances are tied to the success of that group (Dawson, 1994) — weakens after marriage. This may be a result of women institutionalizing their relationship with men and feeling less connected to other women, or having weak gender linked fate (Stout, Kretschmer, & Ruppaner, 2017). Marriage can lead to a female's investment in family linked fate rather than gender linked fate as they begin to have children and run a household. Married women are inclined to feel less association with their gender than unmarried women. The findings of Stout's study suggests there are political ramifications of marriage for which we may be unaware.

To predict the relationship between marriage and a social issue, this study created a mathematical model. Mathematical models are often used in political science as a tool to predict and analyze the behavior of different groups. Such models have been used in the past to generate attention or financial redirection for politicians based on the predictions formed.

In this study, analysis was performed to determine whether married or unmarried Congressional representatives were more likely to discuss women's issues — as defined by discussion of sexual harassment, sexual assault, pay inequality, the recognition of women, human trafficking, women's health issues, gender disparities in profession, and international women's issues (Stout, Kretschmer, Bartula-Henkle, McLean, 2019). A mathematical model was created to predict the likelihood of representatives to discuss women's issues based on their marital status.

The model was created based on data retrieved from members of the 115th Congress, which occurred during the unraveling of Harvey Weinstein thus heightening sensitivity and awareness to the MeToo movement across the board. A data set of over 400 U.S. Congress representatives' Tweets was used in this study. Twitter was selected as the data source as it provides a frequent source of unfiltered representative communication. Other forms of outreach such as press releases and speeches are both rehearsed and less frequent. The sample of Congressional representatives was selected as it provides a targeted group of individuals who are intended to reflect the views of the public.

The intentions behind the model were to create predictions for future candidates' likelihood to speak out about women's issues in addition to showing current trends in this relationship. The results of the model could alter the public's political views and voting choices. Data-based predictions for candidates' likelihood to advocate for women's issues provide the public means by which to empirically assess the platforms candidates might pursue. Additionally, the model could help direct financial backing based on a candidate's ranking through the use of the model. An organization supporting women's rights may financially support a candidate with a higher likelihood to speak out about these issues.

The question examined in this study was, to what extent can a mathematical model predict the relationship between current marital status and congressional representatives' likelihood to discuss women's issues? A conditional hypothesis was developed for the mathematical model. It was hypothesized that the relationship between gender and frequency to discuss women's issues is contingent upon current marital status. If congressmen are married, they will be more likely to discuss women's issues. If congresswomen are married, they will be less likely to discuss women's issues.

METHODS AND MATERIALS

Culling the Data

A web scraper was used to collect tweets from as many members of Congress as possible from the 115th Congress. The use of data from just one Congress ensures consistency as few representatives quit or were removed from office during this time. Tweets were scraped for a total of 408 representatives accounting for 94% of the total U.S. House membership during the 115th Congress. This totaled to 167,784 tweets.

Both hand coding and computer assisted content coding analysis were used to assign each tweet with a score of 1 if it mentioned gender issues or 0 if it did not, creating a dummy variable. Computer assisted content coding analysis was used through the program Rtexttools. A sample of hand coded training set documents was used for the computer assisted content coding to code the remaining tweets with six different algorithms. The documents were then coded based on what the majority of the algorithms agreed was the predicted code and provided information to the researcher about where disagreement occurred among the algorithms. Additional content coding was used on computer assisted content coding where there was uncertainty.

3,500 tweets were hand coded for the presence or absence of a gender topic: discussion of sexual harassment, sexual assault, pay inequality, the recognition of women, human trafficking, women's health issues, gender disparities in professions, and international women's issues.

Tweets were not gathered for anti-choice discussion as this issue has two clear cut sides, pro-choice and anti-choice. Rtextttools was then used to classify the 164,285 remaining tweets.

Rtexttools was chosen as the code program because it provided the opportunity for the researcher to see if the algorithms agreed to determine their accuracy (Jurka et al, 2012; Collingwood and Wilkerson, 2011). The accuracy of the algorithms was found using 3000 of the 3500 hand-coded tweets as a training set and the other 500 as manually coded testers. This predicted the accuracy of the different algorithms (Collingwood and Wilkerson, 2011). When at least five of the six algorithms agreed, the program was over 90% accurate. Instances in which fewer than five algorithms agreed were revisited and hand coded (Stout et al, 2019).

Choosing the Variables

In order to evaluate all the variables desired to produce an accurate result, the possibility of overfitting was considered. This was accounted for by considering the ratio of observations and independent variables so there were not too many terms or the number of observations present.

Multicollinearity was also considered. It was unlikely in the given variables as they were unrelated demographics; the independent variables were correlated with the dependent variable, but not with other independent variables.

The dependent variable selected was "overall female appeal", or the frequency to tweet about women's issues. The independent variable for the regression was marriage. In these models, gender was hand sorted in order to assess male and female representatives separately. In the multiplicative interaction, the dependent variable remained overall female appeal. The independent variable was gender and the conditional variable was marriage.

The control variables selected in order to isolate the independent variable were:

1. If the representative was white
2. The representative's age
3. If the representative was a democrat
4. The partisan voting index of the representative's district
5. The percent of the representative's district that was female

6. The total tweets of the representative

It was important to control for race as each race of representatives may discuss these issues at a different capacity, regardless of their marital status, which may skew the results obtained (Bolzendahl & Meyers, 2004). Age was controlled for as it is likely that it has an effect on how often an individual speaks out due to changing views on gender equality over time. According to Pew Research (2017), Democrats are twice as likely as Republicans to say more work is needed on gender equality. Thus, democrat was added as a control. The partisan voting index (PVI) of the district was also controlled for as it measures how far a congressional district leans toward the Republican or Democratic party. For the same reason as the inclusion of the democrat variable, the PVI variable may lead to an inclination to elect representatives who advocate for these issues as the support for gender equality may be better associated with liberal views than conservative views. Due to the concept of gender linked fate, the percent female of each district was controlled for as districts with more women likely elect representatives who better advocate for women's issues. Lastly, the total tweets of the representative was controlled for. Each representative may tweet at different frequencies, so assessing their total tweets about women's issues without controlling for their total tweets might lead to inaccurate results.

Creation of Mathematical Models

It was determined that a t-test was not ideal or complex enough to represent the trend as it only reflects up to a bivariate relationship. Thus, there was no opportunity to include control variables. Similarly to a t-test, a two way ANOVA excludes the ability to add control variables to the model.

Instead, a linear regression was used. This was unsuccessful as women and men did not follow similar trends and therefore could not be grouped together in the regression with a successful result. The regression reflected a simplistic x-y relationship. In order to combat this, the OLS regression was run again once with just men and once with just women. This was still unable to capture the conditionality of the hypothesis. In order to reflect the conditional hypothesis, an interaction was the ideal statistical analysis test. It was possible to make predictions based on the multiplicative interaction model. In creating all of the models, the following assumptions and justifications were made in order to validate the models.

Assumptions and Justifications:

1. Assumption: All representatives use Twitter with equal purpose.

Justification: representatives were all tweeting about their social and political views.

2. Assumption: All marriages can be considered as equal.

Justification: There was no way to differentiate the dynamics within each marriage studied. Thus, the assumption must be made that they were all the same.

3. Assumption: Regardless of gender roles in a household, the likelihood of representatives to speak out was the same.

Justification: The roles of each partner in a household cannot be assessed differently for each representative evaluated.

4. Assumption: Why or why not representatives were married did not affect their likelihood to speak out.

Justification: The reason that each representative was or was not married was not possible to determine. The assumption must be made that the reason behind the choice to marry or not to marry will not have any bearing on their likelihood to speak out.

5. Assumption: representatives' activity on Twitter could be generalized to their discussion in real life.

Justification: There was no way to differentiate between the way representatives tweet and their day to day conversations in real life.

6. Assumption: All representatives were binary and gender was listed correctly.

Justification: Every representative collected in the dataset recognized as either male (0) or female (1).

7. Assumption: The use of Twitter was representative of discussion in real life.

Justification: There has been no research to support that twitter activity was not representative of the way representatives discussed similar topics in real life.

Linear Regression Models

Independent Variable: Married(1) not married(0)

Dependent Variable: Frequency of discussion of women's issues on Twitter (Overall female appeal)

A linear regression was used because the independent variables studied were categorical while the dependent variable was continuous. Linear regressions can be used in multivariate models to predict the relationship between x and y minimizing the total squared error. The coefficients of the regression are estimates of the actual population parameters. In all estimated equations written, the error term was assumed to be zero. In performing the OLS regressions, six definite assumptions were made:

1. The model was linear in the coefficients and the error term.
2. The average value of the random error term was zero.
3. There was no correlation between any of the independent variables measured and the error term.
4. Observations within the error term shared no correlation: if there was a correlation within the error term it was included within the actual model.
5. There was no heteroscedasticity, or the error term had constant variance.
6. No independent variable was a perfect linear function of another.
7. The error term was normally distributed. (Schneider et al, 2010).

Linear Regression 1: Uncontrolled

The regression was run uncontrolled for both males and females using marriage as the independent variable :

$$\hat{y} = a + b1xi$$

a represents the y-intercept. Here, overall female appeal.

$b1$ represents the married variable.

xi represents the score on marriage (0 or 1).

Linear Regression 2: Controlled

The model was also created using marital status as the main independent variable while controlling for other variables in order to assess marital status's isolated effect on the overall female appeal. The independent variables observed were the same as those in the "Choosing the Variables" section.

Both of these models were repeated, but with only female representatives included and again with only male representatives included.

$\hat{y} = a + \beta_1 xi + \beta_2 xi + \beta_3 xi + \beta_4 xi + \beta_5 xi + \beta_6 xi + \beta_7 xi$

a represents the y-intercept, the overall female appeal.

β_i accordingly from one to nine is the married coefficient, white coefficient, age coefficient, democrat coefficient, PVI coefficient, percent female coefficient, and total tweet coefficient.

xi represents the score on each variable.

Multiplicative Interaction Models

<u>Independent Variable:</u> Female(1) or male(0)

<u>Dependent Variable:</u> Frequency to Tweet about women's issues (overall female appeal)

<u>Conditional Variable:</u> Whether the representative was married (1) or not (0)

A simple x-y relationship was insufficient as the relationship was not linear so the linear regression failed to create a valid mathematical model to represent the relationship between the independent and dependent variables. A multiplicative interaction model was created instead. This method was chosen because it represents a conditional relationship where the relationship between x and y is contingent upon z. Additionally, it was selected because, unlike in additive models (like linear regressions), the coefficients in interaction models do not indicate the average effect of a variable. Therefore, the inclusion of an interaction term makes a change more likely. The terms for using an interaction model that have been established were followed:

1. The hypothesis is conditional.
2. All constitutive terms should be included in their interaction model specifications.
3. Constitutive terms should not be interpreted as if they are unconditional marginal effects.

4. The data cannot be interpreted without the use of confidence intervals. The lower term and higher term must be added in order to represent a representative's location on the graph.
5. Substantively meaningful marginal effects and standard errors should be calculated (Brambor, Clark, & Golder, 2005).

The control variables included in order to ensure that the results observed regarding marriage's effect on overall female appeal was not affected by other factors were the same as those mentioned in the "Choosing the Variables" section.

The interaction was run between the female and married variables. The separated values of Z added to the regression were:

1. Female
2. Married

Interaction 1: Uncontrolled

The uncontrolled interaction equation was modeled by:

$$\hat{y} = a + \beta_1 X + \beta_2 Z + \beta_3 X$$

In this equation,

a represents the y-intercept. In this case, the overall female appeal.

B_i represents the coefficient of female, married, and female##married.

X_i represents the score on female.

Z_i represents the score on married.

Interaction 2: Controlled

The interaction was then run again with controls by adding additional independent variables.

$$\hat{y} = a + \beta_1 X + \beta_2 Z + \beta_3 ZX + \beta_4 xi + \beta_5 xi + \beta_6 xi + \beta_7 xi + \beta_8 xi + \beta_9 xi$$

In this equation,

a represents the y-intercept. In this case, the overall female appeal when everything else is equal to zero.

β_i accordingly from one to nine is the female coefficient, married coefficient, female##married coefficient, white coefficient, age coefficient, democrat coefficient, PVI coefficient, percent female coefficient, total tweets coefficient

X represents the score on female.

Z represents the score on married.

ZX represents the score on female##married.

xi represents the score on every other variable.

Testing the Model

The actual value for overall female appeal was pulled from the dataset and compared with the predicted value for overall female appeal obtained through the interaction model. An unpaired t-test with equal variances was run. In this test, no significance signified that the results were similar to the actual value and predicted value. Subsequently, a scatterplot was created in order to visually represent how closely related the two variables were.

RESULTS

Linear Regressions

As seen in Figure 1, the uncontrolled linear regression, the variable married in the male and female sample was significant ($p<.01$). The married variable in the female sample was less significant, but still had significance ($p<.1$). Lastly, the married variable in the only male group was insignificant ($p>.1$).

	MALE AND FEMALE UNCONTROLLED		FEMALE UNCONTROLLED		MALE UNCONTROLLED	
VARIABLES	OVERALL FEMALE APPEAL	ERROR	OVERALL FEMALE APPEAL	ERROR	OVERALL FEMALE APPEAL	ERROR
Married	-0.02***	(0.00)	-0.02*	(0.01)	0.00	(0.00)
Constant	0.04***	(0.00)	0.08***	(0.01)	0.01***	(0.00)
Observations	408		80		328	
R-squared	0.03		0.04		0.00	

Standard error in parentheses
***p<0.01, **p<0.05, *p<0.1

Figure 1: Male and female linear regression uncontrolled, only female regression uncontrolled, only male regression uncontrolled. n=408 for the male and female sample. n=80 for the female sample. n=328 for the male sample.

For the male and female uncontrolled sample, the equation $\hat{y} = 0.04 - 0.02 married$ was modeled to predict for future representatives' likelihood to discuss women's issues. For the female uncontrolled group, the obtained equation was $\hat{y} = 0.08 - 0.02 married$. The male uncontrolled group resulted in a predicted equation of $\hat{y} = 0.00 + 0.01 married$.

In order to test the validity of what was found in figure 1, the linear regression was run with controls as seen in figure 2. For the male and female sample seen in figure 2, the married variable was extremely significant (p<.01). When looking at only females, married had no significance (p>.1). And when looking at males, the married variable was significant (p<.05).

	MALE AND FEMALE CONTROLLED		FEMALE CONTROLLED		MALE CONTROLLED	
VARIABLES	OVERALL FEMALE APPEAL	ERROR	OVERALL FEMALE APPEAL	ERROR	OVERALL FEMALE APPEAL	ERROR
Married	-0.02***	(0.00)	-0.00	(0.01)	0.01**	(0.00)
White	-0.00	(0.00)	0.01	(0.01)	-0.00	(0.01)
Age	0.00**	(0.00)	0.00	(0.00)	0.00	(0.00)
Democrat	0.01***	(0.00)	0.00	(0.02)	0.01***	(0.00)
PVI	0.00**	(0.00)	0.00*	(0.00)	0.00*	(0.00)
Percent Female	0.00	(0.00)	0.02**	(0.01)	-0.00	(0.00)
Total Tweet	0.00	(0.00)	-0.00	(0.00)	0.00	(0.00)
Constant	-0.06	(0.09)	-0.77***	(0.31)	0.04	(0.04)
Observations	408		80		328	
R-squared	0.21		0.26		0.23	

Standard error in parentheses
***p<0.01, **p<0.05, *p<0.1

Figure 2: Male and female linear regression controlled with n=408 , only female linear regression controlled n=80 , and the only male regression controlled with n=328.

For the male and female controlled group, the equation $\hat{y} = -0.06 - 0.02 married + 0.01 democrat$ was obtained to model the relationship.

$\hat{y} = -0.77 + 0.01 white + 0.02 percentfemale$ was the equation obtained from the female group.

The male controlled equation was $\hat{y} = 0.04 + 0.01 married + 0.01 democrat$.

Multiplicative Interaction Models

As seen in figure 3, the female term was significant (p<.01). For the married term, no significance was found (p>.1). The interaction between the female and married variables was extremely significant (p<.01).

	UNCONTROLLED	
VARIABLES	OVERALL FEMALE APPEAL	ERROR
Female	0.07***	(0.01)
Married	0.00	(0.00)
Female*Married	-0.02***	(0.01)
Constant	0.01***	(0.00)
Observations	408	
R-squared	0.39	

Standard error in parentheses
***p<0.01, **p<0.05, *p<0.1

Figure 3: Male and female interaction uncontrolled. n=408. For Female*Married (p<.01).

As seen in figure 4, the uncontrolled interaction, married males are .01% more likely than unmarried males to tweet about women's issues. Women are 2% less likely to tweet about women's issues if they are married. The interaction tested exclusively for gender and frequency to discuss women's issues based on marriage.

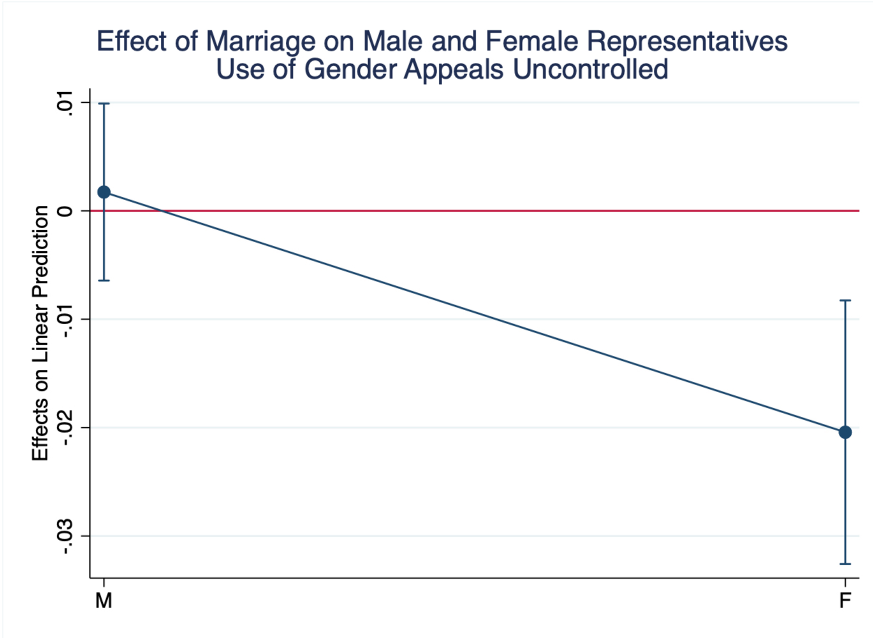

Figure 4: Graph of the Uncontrolled Interaction between Gender and Married Variables.

After eliminating any variables with coefficients of zero, the uncontrolled interaction equation was written as $\hat{y} = 0.01 - 0.02 female * married + 0.07 female$.

Figure 5 shows that the interaction had very high significance (p<.01). Additionally, the constitutive terms female and married both had significance. Female was found significant (p<.01) and married was slightly significant (p<.1).

	CONTROLLED	
VARIABLES	OVERALL FEMALE APPEAL	ERROR
1.Female	0.06***	(0.01)
1.Married	0.01*	(0.00)
Female*Married	-0.02***	(0.01)
White	-0.00	(0.00)
Age	0.00	(0.00)
Democrat	0.01**	(0.00)
PVI	0.00***	(0.00)
Percent Female	0.00	(0.00)
Total Tweet	-0.00	(0.00)
Constant	-0.10	(0.07)
Observations	408	
R-squared	0.47	

Standard error in parentheses
***p<0.01, **p<0.05, *p<0.1

Figure 5: Male and female interaction controlled. n=408. Female*Married (p<.01).

Figures 6 illustrates the multiplicative interaction results when it was run with controls. Married males were 1% more likely to discuss women's issues than unmarried males. Married women were 2% less likely to discuss women's issues than unmarried women.

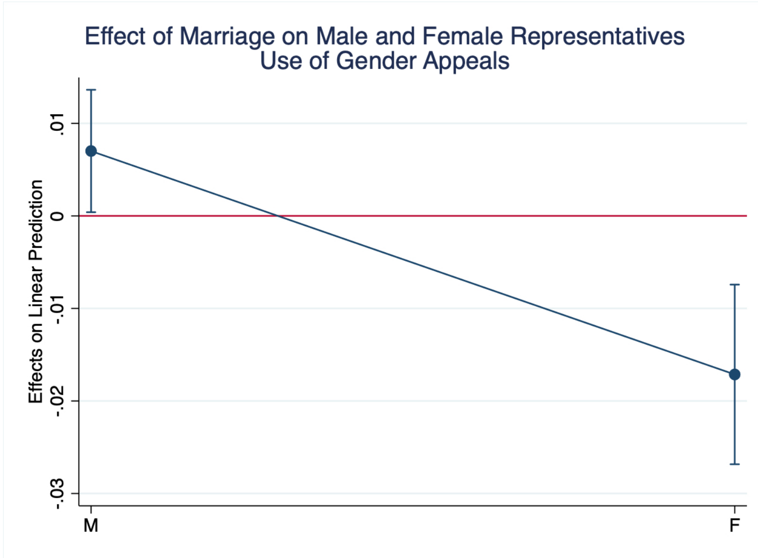

Figure 6: Graph of the Controlled Interaction between Gender and Married Variables.

The controlled interaction was written with the equation

$\hat{y} = -0.10 - 0.02 female * married + 0.06 female + 0.01 married + 0.01 democrat$. This equation was used to make predictions for future candidates.

Testing the Model

As seen in figure 7, the validity of the interaction model can be compared with the actual values of overall female appeal.The best fit line was observed to have a slope of around one, signifying that the predictions and actual values were similar.

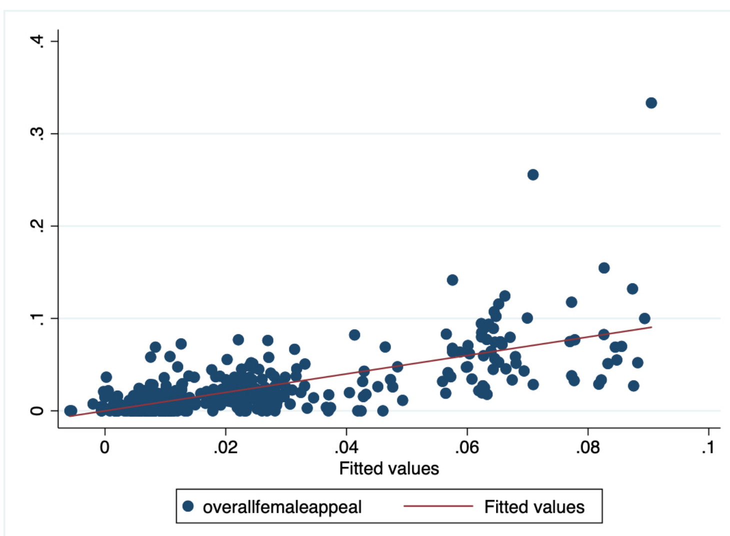

Figure 7: Graph of Measured Values of overall female appeal vs Estimated (Fitted) Values of overall female appeal.

In addition to the scatterplot created, a two-sample t-test was run between the actual values of overall female appeal and the predicted values assuming equal variance. Figure 8 shows that the standard deviation between these two variables was similar. Additionally, the two-tailed relationship was not significant (p>.1).

	overall female appeal	Predicted overall female appeal
Mean	0.02	0.02
Variance	0.00	0.00
Observations	419	408
Pooled Variance	0.00	
Hypothesized	0.00	
df	825	
t Stat	0.10	
P(T<=t) one-tailed	0.46	
t Critical one-tailed	1.65	
P(T<=t) two-tailed	0.92	
t Critical two-tailed	1.96	

Figure 8: Bivariate, unpaired t-test between actual overall female appeal and Predicted overall female appeal assuming equal variance. (p>.1) for the results.

DISCUSSION AND CONCLUSIONS

Analysis of Study

The hypothesis that unmarried congresswomen would discuss women's issues more than married congresswomen was supported. Additionally, married congressmen discussed women's issues more than unmarried congressmen supporting the hypothesis that the relationship between gender and frequency to discuss women's issues was dependent on current marital status. If congressmen were married, they were more likely to discuss women's issues. If congresswomen were married, they were less likely to discuss women's issues.

The results of the controlled interaction showed that married congresswomen were 2% less likely than unmarried congresswomen to speak out about women's issues. This result was significant (p<.01). Married congressmen were 1% more likely to speak out about women's issues than unmarried congressmen with ($p<.1$). When comparing the predictions made on overall female appeal using the interaction model with the actual overall female appeal, it was found that the t-test was insignificant ($p>.1$). Thus, the interaction model formed predictions of overall female appeal that were close to the actual values of overall female appeal as collected directly from Twitter.

Current research assesses the financial ramifications of marriage, but not the social ones. In particular, little has been done to assess how marriage affects gender equality. This model sought to represent the relationship between gender and discussion of women's issues, based upon marriage. In previous assessments, it was determined that the growing rate of millennials choosing not to get married will have effects on the economy that we are unprepared for (Murphy, 2019). However, the concepts of social inequality were not evaluated. The mathematical model was directed to examine one of these social effects, the frequency to advocate for women's issues. Congress members were the sample assessed. They serve as U.S. citizen representation as it is expected that they reflect the feelings of the general public in regard to their frequency to discuss these issues.

The social effect of discussion of women's issues was chosen because the desire to advocate for the equality of women may be a demonstration of gender linked fate. After Stout et al's findings (2017) that gender linked fate weakened in females after marriage due to them institutionalizing their relationships with males, it was expected that married congresswomen would speak out about women's issues less than unmarried congresswomen and that the opposite would hold true for congressmen. The reason that this trend may have occurred in congressmen was similar to why female gender linked fate weakens after marriage. Men may institutionalize their relationships with their wives and become more inclined to consider women's issues as they affect their partner. This trend may be further strengthened if congressmen have daughters (Stout, 2019). Adding another female family member continues to expose a congressman to women's issues.

The results of the study demonstrated that a relationship was present between marriage and the frequency to speak out about women's issues. This has political ramifications. The model can be used to generate attention or financial redirection based on the rating a candidate for congress receives on the model. When firms choose to fund a campaign, they may refer to this model in order to predict how likely the candidate they are supporting is to discuss these issues. This may alter who they choose to fund. The public can use this model as a measure to assess their political views and voting decisions based upon data. The multiplicative interaction model may alter the public's political views when future candidates are plugged into it. If the results of candidates' likelihood to speak out about these issues are not what a voting citizen hopes them to be, their voting choice may shift.

Evaluation

The use of Twitter as a proxy for discussion may not be wholly accurate. It is possible that representatives do not tweet about women's issues with the same frequency they discuss them in real life. This, however, could not be resolved as there was no way to assess each utterance of a congress member that contains a reference to women's issues. Another aspect of the study which could be strengthened is the generalizability of the model as it was created using data exclusively from the 115th Congress. This was in the heat of the MeToo movement following the unraveling of the Harvey Weinstein scandal. Thus, the frequency of representatives to tweet about women's issues in this Congress may not be reflective of all other Congresses. There may have been heightened advocacy for women's issues in this era.

Future Research

Research regarding marital statuses other than "married" and "not married" would add another dimension to the model created. In particular, marital statuses such as divorced and widowed may show important trends in the results. The validity of the current model may be further tested by scraping tweets of members of the 116th Congress in order to verify that the results obtained were not solely due to the MeToo movement that occurred contemporaneously with the 115th Congress. Lastly, tweets from one specific candidate or representative could be scraped and plugged into the model in order to test the model's accuracy in predicting their overall female appeal.

REFERENCES

Brambor, T., Clark, W., & Golder, M. (2006). Understanding interaction models: improving empirical analyses. *Political Analysis, 14*(1), 63-82. doi:10.1093/pan/mpi014

Bolzendahl, C. I., & Meyers, D. J. (2004). Feminist attitudes and support for gender equality: Opinion change in women and men, 1974-1998. *Social Forces*, *83*(2), 759-789. Retrieved from JSTOR database.

Clair, Robin Patric, et al. "# MeToo, sexual harassment: an article, a forum, and a dream for the future." *Journal of Applied Communication Research* (2019): 1-19.

Collingwood, Loren, and John Wilkerson. "Tradeoffs in accuracy and efficiency in supervised learning methods." *Journal of Information Technology & Politics* 9.3 (2012): 298-318.

Dawson, Michael. 1994. *Behind the Mule: Race and Class in African-American Politics*. Princeton: Princeton University Press.

Druckman, James N., Martin J. Kifer, and Michael Parkin. "Campaign communications in US congressional elections." *American Political Science Review* 103.3 (2009): 343-366.

Jurka, Timothy P., et al. "RTextTools: Automatic text classification via supervised learning." *R package version* 1.9 (2012).

Murphy, Meg. “NowUKnow: Why Millennials Refuse to Get Married.” *PreparedU View | Bentley University*, 3 Jan. 2019, www.bentley.edu/impact/articles/nowuknow-why-millennials-refuse-get-married.

Pietryka, Matthew T. "The roles of district and national opinion in 2010 congressional campaign agendas." *American Politics Research* 40.5 (2012): 805-843.

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis: part 14 of a series on evaluation of scientific publications. *Deutsches Arzteblatt international*, *107*(44), 776–782. doi:10.3238/arztebl.2010.0776

Sides, John. "The origins of campaign agendas." *British Journal of Political Science* 36.3 (2006): 407-436.

Stout, C. (2019, November 5). [Personal interview].

Stout, C.T., Kretschmer, K.N., & Ruppanner, L. (2017). *Gender linked fate, race/ethnicity, and the marriage gap in American politics.*

Stout, C., Kretschmer, K., Burtula-Henkle, M., & McLean, L. (n.d.). *The Weinstein effect: The role of MeToo movement in altering the communication strategies of members of Congress*.