Observational Study on Determinants of COVID-19 Vaccination Uptake

Lavinia Xu
WASHINGTON UNIVERSITY IN ST. LOUIS

Introduction

This study examines factors that influence COVID-19 vaccination uptake in the United States, using data from a September 2022 Pew Research Center survey. As vaccination rates play a crucial role in managing public health, understanding the demographic and social factors that drive vaccination behavior has become increasingly essential. Research suggests that trust in the perceived credibility of elected officials and scientists, and demographic influences like age, political beliefs, and religious affiliations are all important components shaping individuals' vaccination choices.

To analyze the influences, we employed proportional odds model, a statistical approach well-suited to evaluating ordinal outcomes — in this case, fully vaccinated, partially vaccinated, and unvaccinated groups. This model enables us to evaluate the associations between respondents' vaccination status and variables such as confidence in officials across different sectors while preserving the ordered nature of vaccination categories. By investigating various factors, this study aims to provide a comprehensive understanding of how public trust, personal beliefs, and demographic characteristics have an impact on vaccination behavior.

Research Goal

- To apply statistical modeling to identify key predictors of COVID-19 vaccination status
- To generate insights that provide guidance on future public health strategies, aiming to enhance vaccination uptake

Methods

Proportional odds model is a statistical technique designed to analyze ordinal response variables, where the outcome has a natural order but unknown intervals between categories.

To understand the model algorithms for proportional odds model, we must first define logit. The *logit* is a transformation that represents the natural logarithm of the odds of a particular event occurring. In statistics, the logit function is used to map probabilities (ranging between 0 and 1) onto a continuous scale that spans from negative to positive infinity. The following equation depicts the probability p of an event occurring:

$$\text{logit}\,(p) = \log\left(\frac{p}{1-p}\right)$$

Proportional odds model estimates a series of cumulative logits. For each category j of the response variable Y, the model calculates the cumulative probability of being in that category or lower. These probabilities are then linked to predictor variables $\boldsymbol{x}$ through a logistic function. In its mathematical form, the model estimates:

$$\text{logit}\,[P(Y \leq j)] = \log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \alpha_j + \boldsymbol{\beta}'\boldsymbol{x},$$

where α_j are intercepts for each category and $\boldsymbol{\beta}$ represents the coefficients for predictor variables.

One key assumption that the model relies on is *Proportional Odds Assumption*, stating that the effect of predictor variables on the odds is consistent across all levels of the ordinal outcome.

Results

```
Call:
polr(formula = VACCINATED ~ CONF_GOV + AGE + RELIG +
PARTY +
    GENDER, data = ordinal_data, weights = WEIGHT, Hess
= TRUE)

Coefficients:
               Value Std. Error t value
CONF_GOVHM -0.4113    0.06815  -6.036
CONF_GOVL  -0.4420    0.05858  -7.544
AGE2       -0.1835    0.06488  -2.828
AGE3       -0.5960    0.07210  -8.267
AGE4       -1.3438    0.08309 -16.172
RELIGCTHLC -0.5519    0.06682  -8.260
RELIGUNFFL -0.1350    0.06002  -2.249
RELIGOTHER -0.3753    0.09464  -3.965
PARTYREPUB  0.1774    0.09650   1.839
PARTYDEMOC -1.3632    0.09976 -13.664
GENDERM    -0.4372    0.04860  -8.996

Intercepts:
    Value    Std. Error t value
Y|P  -0.6895   0.1177    -5.8560
P|N  -0.3008   0.1176    -2.5570

Residual Deviance: 13136.93
AIC: 13162.93
```

Figure 1(a). Ordinal Model with CONF_GOV

```
Call:
polr(formula = VACCINATED ~ CONF_EDU + AGE + RELIG +
PARTY +
    GENDER, data = ordinal_data, weights = WEIGHT, Hess
= TRUE)

Coefficients:
               Value Std. Error t value
CONF_EDUH  -1.2974    0.10404 -12.470
CONF_EDUM  -1.0109    0.08206 -12.319
CONF_EDUL  -0.6487    0.08652  -7.497
AGE2       -0.1753    0.06520  -2.688
AGE3       -0.6052    0.07271  -8.322
AGE4       -1.3708    0.08388 -16.343
RELIGCTHLC -0.5796    0.06725  -8.618
RELIGUNFFL -0.1465    0.06031  -2.429
RELIGOTHER -0.4007    0.09564  -4.189
PARTYREPUB  0.1635    0.09737   1.679
PARTYDEMOC -1.2962    0.10053 -12.894
GENDERM    -0.4625    0.04897  -9.443

Intercepts:
    Value    Std. Error t value
Y|P  -1.2229   0.1344    -9.0982
P|N  -0.8279   0.1341    -6.1739

Residual Deviance: 12992.42
AIC: 13020.42
```

Figure 1(b). Ordinal Model with CONF_EDU

```
Call:
glm(formula = I(VACCINATED == "N") ~ CONF_SCI + AGE +
RELIG +
    PARTY + GENDER, family = binomial, data = data_YP)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.49272    0.22075   6.762 1.36e-11 ***
CONF_SCIH   -3.03104    0.14587 -20.780  < 2e-16 ***
CONF_SCIM   -2.06333    0.12029 -17.153  < 2e-16 ***
CONF_SCIL   -0.89461    0.12382  -7.225 5.02e-13 ***
AGE2        -0.12661    0.11205  -1.130    0.259
AGE3        -0.59519    0.11709  -5.083 3.71e-07 ***
AGE4        -1.41364    0.12474 -11.333  < 2e-16 ***
RELIGCTHLC  -0.59729    0.08781  -6.802 1.03e-11 ***
RELIGUNFFL   0.07714    0.07719   0.999    0.318
RELIGOTHER  -0.18846    0.11939  -1.579    0.114
PARTYREPUB   0.17592    0.16289   1.080    0.280
PARTYDEMOC  -1.28505    0.16949  -7.582 3.41e-14 ***
GENDERM     -0.49454    0.06293  -7.859 3.88e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9146.9  on 10121  degrees of freedom
Residual deviance: 7088.5  on 10109  degrees of freedom
AIC: 7114.5

Number of Fisher Scoring iterations: 6
```

Figure 1(c). Logistic Regression with CONF_SCI

We used two proportional odds models and one binary logistic regression model to investigate factors influencing COVID-19 vaccination status, focusing on public confidence in government officials, educational leaders, and scientific authorities, along with demographic factors such as age, religious affiliation, political beliefs, and gender. To address potential collinearity as suggested by **Figure 2(a)** and **Figure 2(b)**, each model included only one type of confidence variable, enabling us to separately evaluate each confidence factor's impact on vaccination likelihood.

In the first proportional odds model in **Figure 1(a)**, we analyzed the effect of confidence in government officials on vaccination status. The model revealed that individuals with low or moderate (high) confidence in government officials had greater odds of vaccination compared to those with no confidence, as indicated by the model's negative coefficients. This association suggests that confidence in government influences individuals' vaccination decisions. The second proportional odds model in **Figure 1(b)** examined confidence in educational leaders, showing a clear linear trend: as confidence in these officials increased, so did the likelihood of vaccination. This more consistent trend underscores the positive role that educational trust may play in encouraging vaccination behavior. For the third model shown in **Figure 1(c)**, we employed a binary logistic regression to assess the effect of confidence in scientific authorities, combining fully and partially vaccinated individuals due to the small size of the partially vaccinated group. This model found confidence in scientific authorities had the strongest impact on vaccination status, with significant negative coefficients across all confidence levels, indicating that higher trust in scientists strongly correlates with vaccination uptake.

Demographic variables across all models provided additional insights. Age was positively associated with vaccination, as older age groups consistently showed higher vaccination rates compared to the youngest group. This finding aligns with our graph shown in **Figure 2(c)**. Religious affiliation was also significant, with Catholics more likely to be vaccinated than Protestants, while the unaffiliated and other religious groups showed no substantial differences. Political affiliation revealed that Democrats were significantly more likely to get vaccinated than those with no lean, while Republicans exhibited an opposite association. Gender differences were also statistically significant, with men showing a higher likelihood of vaccination compared to women.

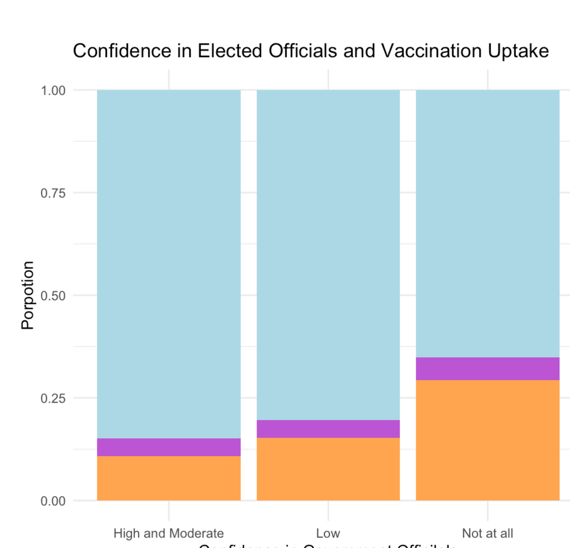

Figure 2(a)

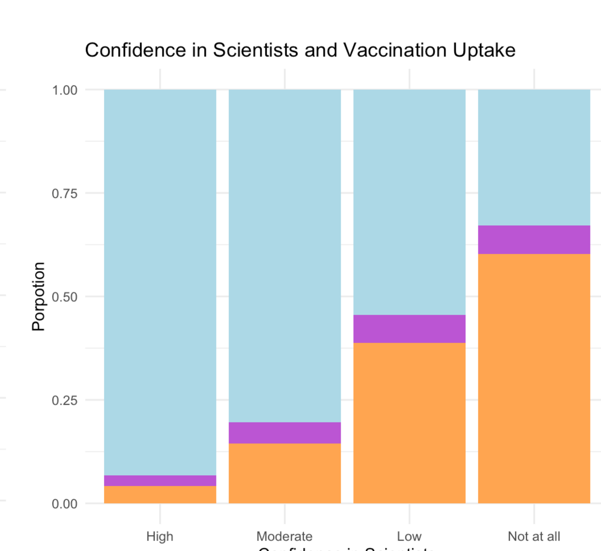

Figure 2(b)

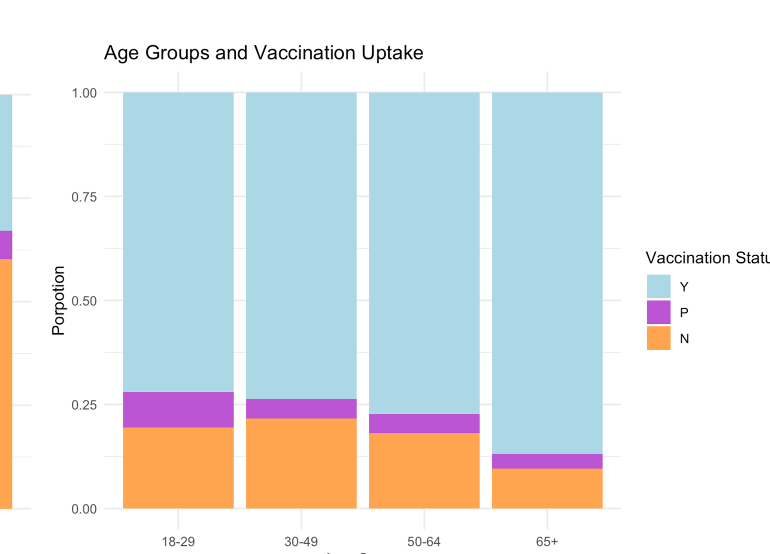

Figure 2(c)

Conclusion

This study underscores the importance of public trust in scientific, educational, and government authorities in promoting COVID-19 vaccination uptake in the United States. Confidence in scientists, or more specifically medical scientists, appeared to be the most influential factor, with higher trust correlating with a profoundly greater likelihood of vaccination. Demographic factors, including age, religious affiliation, political beliefs, and gender, also showed statistical significance for impacting vaccination status.

The findings suggest that building and maintaining public trust in scientific institutions is crucial not only for current vaccination efforts but also for future public health initiatives. Policymakers may want to prioritize transparent, science-based communication and engage trusted community leaders to bridge gaps in trust. Additionally, efforts to depoliticize health messaging and foster bipartisan support for science could improve vaccine acceptance across diverse groups. By addressing both trust and demographic influences, public health strategies can better promote vaccination rate and prepare for future health challenges in a polarized environment.

Future Research

In future research, I aim to investigate the underlying causes of varying confidence levels in scientific and governmental authorities across different demographic and political groups. By examining the effects of media exposure, education levels, and community engagement, I hope that I can gain a clearer understanding on how trust is developed and sustained. Additionally, I plan to conduct longitudinal studies to observe shifts in confidence over time, which may reveal insights into sustaining public trust in health initiatives. Exploring the effectiveness of targeted interventions, such as community-driven health campaigns, may also provide strategies for enhancing vaccine acceptance in diverse communities.

Acknowledgments

I would like to express my deepest gratitude to Professor Jimin Ding for her invaluable guidance, support, and insights throughout this research process. Her mentorship has been instrumental in shaping my understanding and approach. I am also deeply thankful to everyone who offered their support and encouragement along the way—your kindness made this journey possible.

References

Agresti, Alan. *Analysis of Ordinal Categorical Data*. Wiley, 2010.

Funk, Cary, and Alec Tyson. “Lack of Preparedness Among Top Reactions Americans Have to Public Health Officials’ COVID-19 Response.” *Pew Research Center*, 5 Oct. 2022, Accessed 2024.