

Who's Masking?: Exploring Mask-Wearing Behavior by U.S. County

Rena Cohen and Chloe Shawah

December 10, 2020

Introduction

As the number of COVID-19 cases soars to unprecedented heights around the United States, public health experts and many political figures continue to emphasize mask wearing as one of the most effective ways to slow the spread of the pandemic. But, as a New York Times survey from July 2020 shows, mask wearing adherence varies widely in counties around the nation. What predictors might explain this variation in mask wearing, and how might public health officials use this information to develop more effective mask-wearing interventions? To what extent can mask wearing predict the spread of the virus on a county level? In this paper, we will address these questions by building several models to best predict mask-wearing behavior at the county level, assess the potential effects of state and county-wide mask mandates, and explore whether mask-wearing data from July has any relationship to current rates of COVID-19 (as of December 2020). Throughout this process, we will pay particular attention to the relationship between political party and mask wearing, which a recent study by researchers at the University of Chicago claims to be “the single most important predictor of local mask use,” a statement we would like to investigate further (Milosh 2020).

Research and Hypothesis

Before delving into our model building and exploratory processes, we made several hypotheses about what variables might be useful in predicting mask usage based on literature. As aforementioned, we hypothesized that political party would be an important predictor of mask-wearing, as President Trump has consistently questioned public health guidance about wearing masks, sometimes rejecting it outright (Victor, 2020). Indeed, numerous surveys have already shown that Republicans wear masks at lower rates than Democrats (generally about a 20%) (Igielnik 2020). We hypothesized that this relationship would hold on a countywide level, with Republican counties being less likely to wear masks on average. Perhaps it would be even more amplified than the relationship between partisanship and mask-wearing on an individual level, for one would assume that a Republican in a Democratic county would be more likely to wear a mask than a Republican in a Republican county, thus compounding these disparities in rate of mask-wearing. While our data does not include responses on an individual level, this would be an interesting question for further research.

As well as political party affiliation within the county, we hypothesized that several other predictors could have some bearing on mask-wearing. Researchers at the National Institute of Health have suggested that age and location (i.e. rural vs. urban setting) affect mask wearing behavior, so we included the percentage of seniors in a county (hypothesized to have a positive association with mask-wearing since COVID-19 most severely affects the elderly) and various measures of population density in our mask-wearing model (hypothesized to be positive, as urban residents are more likely to come into close contact with many people in day-to-day life) (Haisher 2020). Race could also be an important variable, as data from the Pew Research Center has shown differences in mask wearing frequency by race among individuals, and COVID-19 has disproportionately impacted communities of color according to the CDC; accordingly, we hypothesize that counties with higher proportions of minority residents would be more likely to wear masks (Igielnik 2020, CDC 2020). Finally, given that this same Pew Research Center survey showed differences in mask-wearing behavior by education level, we hypothesized that counties with a greater rates of college education would

have higher rates of mask-wearing. Other demographic variables in our model included poverty (predicted to have a negative association with mask wearing) and the percentage of adults who are female (predicted to have a positive association with mask wearing).

Along with these demographic characteristics, we hypothesized that legal and public health conditions would impact mask-wearing frequency: in particular, that counties with more cases and deaths would have greater rates of mask wearing, as residents would be more likely to fear getting the virus. However, the relationship between cause and effect is muddy here: does less mask wearing predict higher rates of the virus or do higher rates of the virus predict higher mask wearing rates? We will not be making causal claims in this study, but it will be interesting to examine this relationship nevertheless. Finally, we hypothesize that state and countywide mask mandates will increase the frequency of mask wearing, a claim we will devote ample time to further investigating.

Data Collection and Wrangling

Mask-wearing frequency, which is our y-variable for much of this study, is from the aforementioned New York Times survey, which was conducted by the survey firm Dynata on behalf of the Times from July 2 to July 14. Aggregated at the county level, it sorts 250,000 individual responses into 3,000 U.S. counties. The survey asked respondents how often they wore a mask (choices were always, frequently, sometimes, rarely, or never) and presents the percentage of people who gave each answer for every county, which we combined into a single weighted average representing the probability that a randomly selected person is wearing a mask in the county; for more information about the rationale behind this, see the Appendix. It is important to remember that this data likely carries a good amount of uncertainty stemming both from the possible variation in the interpretation of the survey response choices by individuals and because only about 80 individuals per county were surveyed (not to mention possible undercoverage/non-response bias; we do not know the sampling procedure, but the NYT does write that “overall the large number of responses provide rough comparisons across many areas”).

Our predictor variables, which included a combination of countywide demographic measurements and coronavirus-specific values, came from a variety of sources detailed in the table below. Whenever possible, we used the most recent data we could; notably, this meant that we chose to use the percentage of people who voted for Trump in 2020 as rather than in 2016 as a measure of partisanship, as we figured this would more closely reflect the political party affiliations of county residents in July 2020 when the NYT mask-wearing survey was administered. Whenever applicable, we transformed raw counts into rates so that predictors would be comparable across counties with different total population sizes.

Compiling and cleaning this data was a (mostly) smooth process, thanks to the FIPS county code, which combines a 2 digit state code with a 3 digit county code to create a unique, standardized 5 digit code for every US county. When the merging process was over, we had 127 rows containing missing data, representing about 4% of the total number of counties in our dataset. The most concerning omission was the lack of political data for Alaska; after doing more research, we learned that Alaska treats election districting in a different way than it treats FIPS county codes, a process that has led to unfortunate complications for many a data scientist (Tonmccg, 2020). Because our class datasets for election results likewise did not include this data for Alaska at the county level, we simply carried forward our analysis without it, with the understanding that any results we came to should not be generalized to Alaska. Other than the large group of unused Alaskan counties, we did not think that removing the rest of the counties containing missing data would greatly impact our results. For more information about how we came to this conclusion, see the data cleaning section of our Appendix.

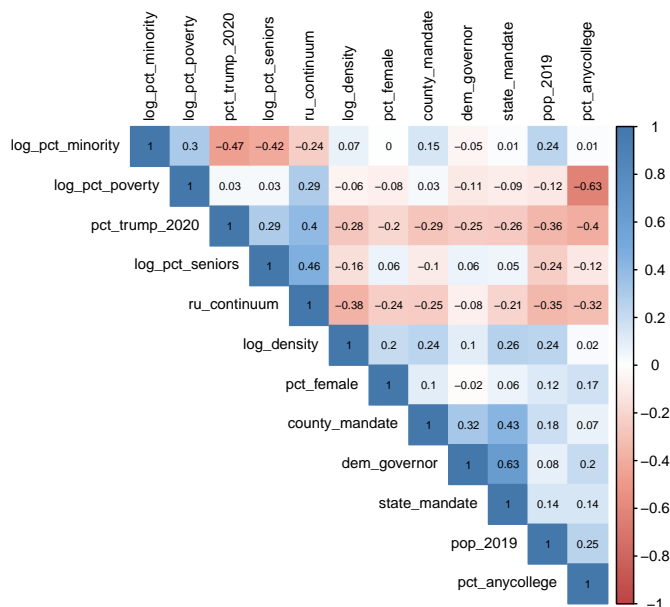
Exploratory Data Analysis

After cleaning our data, our first step was to visualize our main response variable, mask-wearing score, to ensure that assumptions of our later models would be met. We found the distribution of mask-wearing score

to be approximately normal (see Appendix), which means that we will leave it untransformed for our later models. We then repeated this process for our predictors, log-transforming where needed. A final list of our predictors and their transformations is listed below:

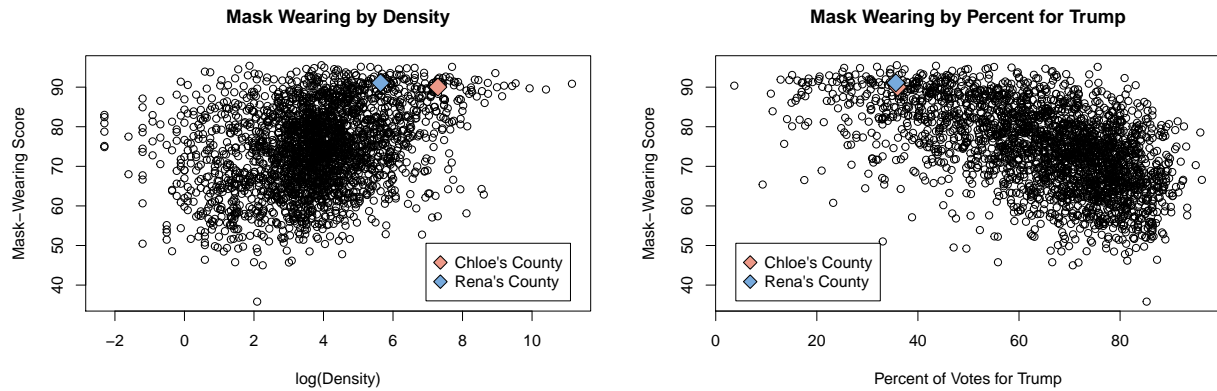
pop_2019:	population in 2019
log_density:	population density, log transformed
ru_continuum:	discrete score from 1 to 10 on the rural-urban continuum with 1 being the most urban
log_pct_seniors:	percent of adults 65+ in 2019, log transformed
log_pct_minority:	percent of people from minority backgrounds in 2019, log-transformed
log_pct_poverty:	percent of people estimated to be living in poverty in 2018, log-transformed
pct_anycollege:	percent of adults who attended at least some college in 2018
pct_female:	percent of females in 2019
pct_trump2020:	percent of votes for Trump in 2020 election
dem_governor:	indicator of whether the county is in a state with a Democrat governor
state_mandate:	indicator of whether the county is in a state with a state-wide mask mandate
county_mandate:	indicator of whether the county has a county-wide mask mandate

To ensure that none of these predictors were too colinear, we computed a correlation table to assess their correlations:

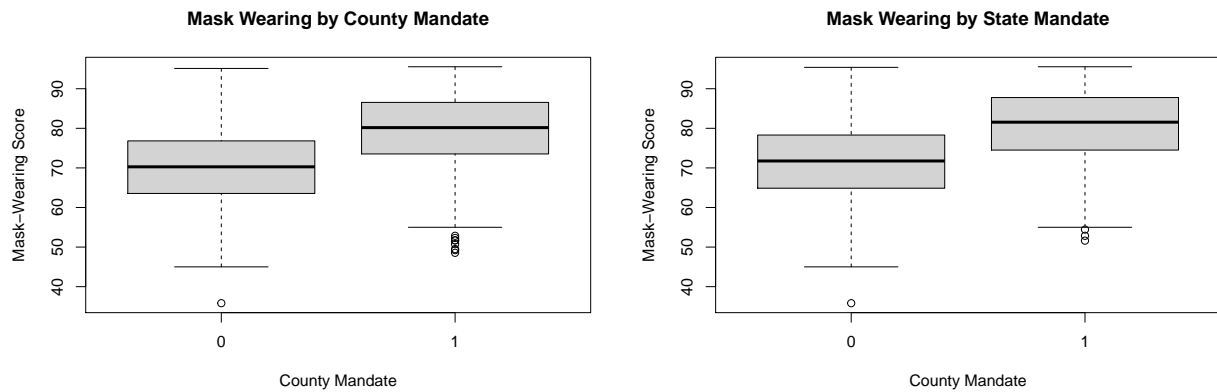


The strongest correlations we see between predictors is between percent poverty and percent college at -0.63 and between Democrat governor and state mandate at 0.63. Both of these associations do not surprise us, and are not so high that we are concerned about their impacts for our modeling. Another takeaway from this correlation table, however, is the extent to which political predictors like votes for Trump are correlated with seemingly apolitical predictors such as percent minority (-0.47). This will be important for us to keep in mind later in our analysis when we try to determine the effect of politics on mask-wearing behavior.

We then visualized the relationship between several of our predictors and mask-wearing score with scatter-plots and boxplots. Two note-worthy plots were those for density and percent Trump votes.



It appears as though there might be a slight positive correlation between density and mask-wearing score, and a fairly strong negative correlation between percent Trump votes and mask-wearing. This shows us that both demographic characteristics and political characteristics of a county have associations with the county's mask-wearing behavior. (But it is also important to remember that density and percent Trump had a -0.28 correlation.) We then decided to visualize the association of mandates with mask-wearing compliance.



Based on the side-by-side boxplots above, we see that there appears to be a difference in mask-wearing behavior for counties with state-wide and county-wide mandates: a mandate looks like it is associated with an increase in mask-wearing. To determine if these differences are statistically significant, we ran two t-tests for a difference in means with $H_0 : \mu_1 = \mu_2$.

	No Mandate	Mandate	Difference	t-statistic	df	p-value
County-level Policy	70.29	79.38	-9.09	-24.52	2502.70	0.00
State-level Policy	71.34	80.26	-8.92	-23.08	1966.38	0.00

These results are unbelievably significant, so we reject the null hypothesis that there is no difference in the mask-wearing behavior in counties with and without a mandate. Yet county and state-level mandates are often a result of local politics and there has been writing, mentioned in our introduction, about how different Republicans are less likely to wear masks than Democrats. Can these differences in mask-wearing behavior be explained entirely by a county's political leaning, or can we explain this difference using our demographic predictors? This question motivates our next section, where we will look more deeply into mask compliance by county through the lens of trying to decide if mask wearing is inherently political.

Modeling and Analysis

Is Mask Wearing Inherently Political?

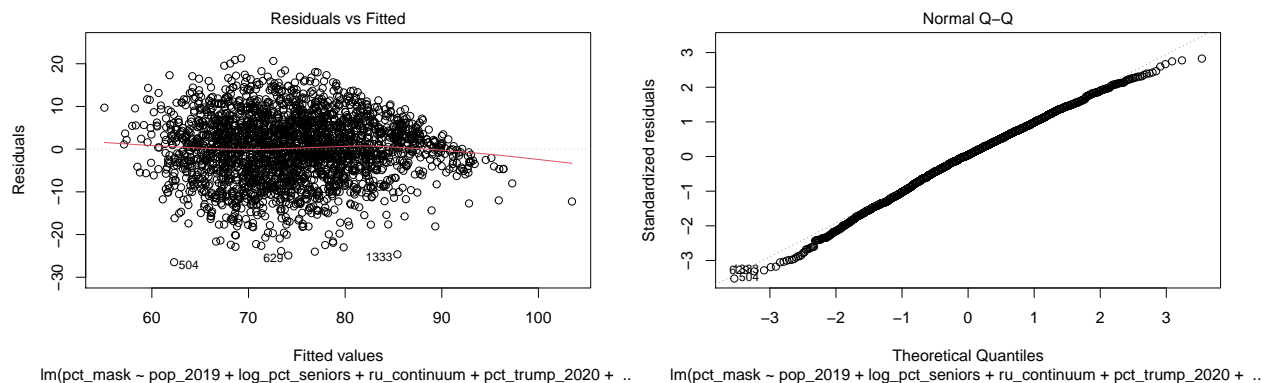
To decide if mask-wearing behavior is inherently political, we will divide our predictors, removing state and county mandates, into two categories: political and apolitical. Even though these predictors can be correlated across categories as we saw in our correlation table, this does not undermine our goal of determining whether demographic factors can explain mask wearing better than political factors.

```
political:  pct_trump_2020, dem_governor
apolitical: pop_2019, log_pct_seniors, ru_continuum, log_pct_minority, pct_anycollege, log_density,
           pct_female, log_pct_poverty
```

First, we will run 3 models with these two categories of predictors *combined* to predict mask-wearing score so we have an idea of our base ability explain mask-wearing behavior.

- 1) **linear full model**: an OLS model with main effects only
- 2) **linear interaction model**: an OLS model with main effects and all interaction terms
- 3) **linear selected model**: an OLS model using step-wise selection in both directions with an intercept-only model as an lower bound and starting point and the interaction model as an upper bound.

Included here are the assumption checks for the first model **linear full model**. The normality and linearity assumption appear to be reasonable based on the residual and Q-Q plots, however we do see that the variance of residuals appears lower for very high \hat{y} , but does not seem severe enough to warrant a more robust method than OLS. The plots for the second two models are very similar to these plots for **linear full model**.



Here is a table of the coefficients for `full linear model`:

Variable	Coefficient
Intercept	78.23
pop_2019	0.00
log_pct_seniors	8.76
ru_continuum	-1.24
pct_trump_2020	-0.18
log_pct_minority	2.98
dem_governor	3.97
pct_anycollege	-0.06
log_density	0.77
pct_female	-0.33
log_pct_poverty	-1.90

It is consistent with our hypotheses that counties with more seniors, more minorities, Democratic governors, and higher density would wear masks at higher rates (positive coefficients), and that counties that are more rural, have higher poverty levels, and cast a higher percent of votes for Trump would wear masks at lower rates (negative coefficients). The coefficients for `pct_female` and `pct_anycollege` had opposite signs than our hypotheses would suggest.

Because we feel as though our assumptions have been met, we will report our R^2 values for these 3 models. At the end of this section, we will include a full table with all of this section's models and their train and test RMSE.

Model Name	R^2
linear full model	0.4751
linear interaction model	0.5329
linear selected model	0.5309

This tells us that a few things: adding the interaction terms does improve the power of our model, and cutting our full interaction model with 55 predictors to 31 well-chosen predictors only decreases the R^2 of our model slightly.

We compared the coefficients for the main effects in the `linear full model` and `linear selected model` and noticed that when we added the interaction terms, some of the coefficients flipped signs for the main effect: `log_pct_seniors`, `pct_trump_2020`, and `pct_female`. For `pct_trump_2020`, this coefficient was negative in `linear full model`, but flipped to positive in the `linear selected model` while almost all of the interaction terms involving `pct_trump_2020` were negative. This is a little surprising: when we control for many interaction terms such as `pct_trump_2020:pct_anycollege`, an increase in Trump votes while holding all other predictors constant — including the interaction terms involving `pct_trump_2020` — is associated with an increase in mask-wearing score. Very interesting! Perhaps this indicates that the relationship between politics and mask-wearing is more complex than we had thought.

Next, we will fit two more models:

- 4) `linear political model`: an OLS model with just our political predictors and their interactions
- 5) `linear apolitical model`: an OLS model with just our apolitical predictors and their interactions

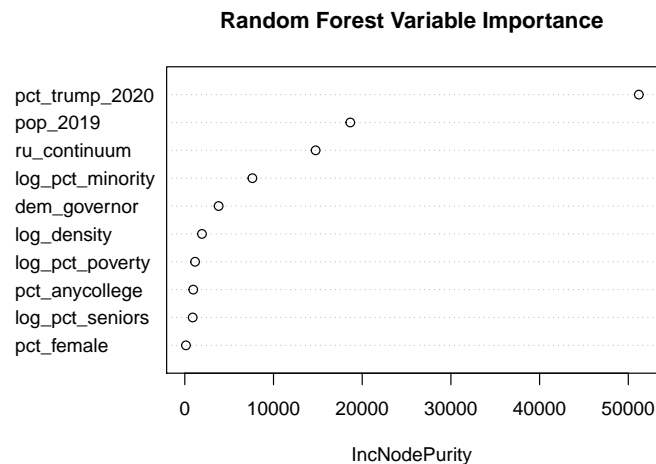
Model Name	R^2
linear political model	0.3222
linear apolitical model	0.4287

Based on these two R^2 values, it seems like our apolitical model is better able to explain variance in mask-wearing behavior by county than using our two political predictors. Even though we only have two political predictors, this still helps us debunk the theory that mask wearing behavior is completely tied to support for President Trump because we are able to explain just as much variability, if not more, using only our demographic variables.

For our final models for this section, we will fit 2 well-tuned random forest models and 1 decision tree that will have an added advantage of being able to catch nonlinear trends in our data. (A decision tree will be used for the political predictors as there are only two). They will be:

- 6) **rf full**: a random forest with all predictors
- 7) **rf apolitical**: a random forest with only apolitical predictors
- 8) **dt political**: a decision tree with the two political predictors

After fitting the **rf full** model, we checked the variable importance plot pictured below.



We found that **pct_trump_2020** was most often used to split the nodes. This does not contradict our early discussion of apolitical factors being able to explain more than political factors, but certainly adds some nuance: on its own, **pct_trump_2020** may be the most the most important predictor of mask-wearing score, but at least in the linear models, when we include lots of other, less individually important demographic predictors, we can get a model that is just as good if not better. We also noticed that **pop_2019** is the second most important predictor according to this plot, even though it was deemed unimportant in some of the linear models above; this indicates to us that there may be a nonlinear relationship between **pop_2019** and **pct_mask**.

Final RMSE Table:

From this final RMSE table, we can make a few observations. The model that performed best on the test data was the linear model with all interaction terms. Some of our models appear to be a little bit overfit, but nothing drastic. Finally, and most importantly, in both the linear models and decision tree/random forest models, the apolitical models outperformed the political models on train and test data.

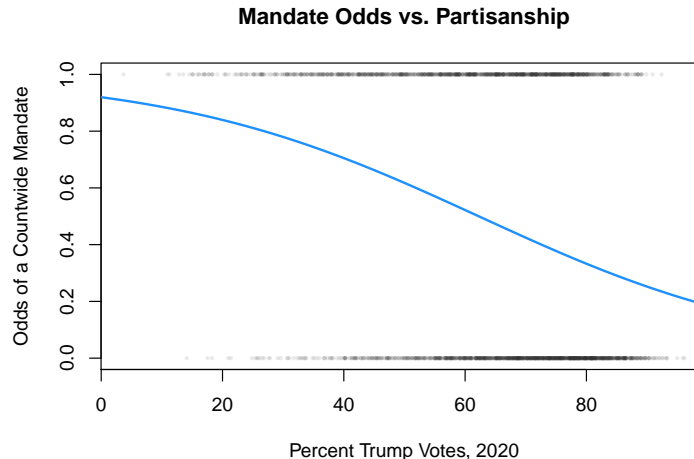
This tells us, tentatively, that we can explain more variability in mask-wearing behavior between counties with demographic predictors as we can with political predictors. Politics may inform people's choice to wear a mask, but it seems that their identity and environment have a significant role in this decision as well.

Model Name	Train RMSE	Test RMSE
linear full model	7.437	7.953
linear interaction model	7.011	7.609
linear selected model	7.032	7.648
linear political model	8.451	8.634
linear apolitical model	7.745	8.120
rf full model	7.601	8.127
dt political model	8.497	8.683
rf apolitical model	7.959	8.465

Examining the Efficacy of Mask Mandates In Promoting Mask Wearing Behavior

Mask mandates at the county and state levels have been widely proposed as an effective and low-cost strategy for reducing the spread of COVID-19. This strategy assumes that wearing a mask truly does reduce the spread of COVID-19 (very reasonable given the preponderance of scientific evidence supporting this claim) (Neff, 2020). It also makes a second assumption: that people will follow a mask mandate if it is in place. This assumption is more questionable, as there are reasons to believe that a mask mandate may have no effect on some individuals who are skeptical about masks (most of whom tend to be Republican, as discussed in a recent Pew Research Study) given that in many states, law enforcement is unwilling to enforce mask mandates (Kessel 2020, Mann 2020) . In this portion of the project, we wish to further investigate the relationship between mask mandates and mask wearing behavior. Is there evidence that counties with mandates have higher mask-wearing behavior after taking into account partisanship (which has become a proxy for attitudes towards mask wearing)? If so, might counties with different political leanings respond to county-wide mask mandates in different ways? While we cannot make causal claims with any of our data, we hope to identify patterns that may inform the often-contentious discussions about local mask mandates as a strategy for COVID-19 mitigation.

While counties with local mask mandates did have higher mask-wearing adherence on average (as shown in a two-sample t-test in our EDA section), counties with a larger share of Republicans were also much less likely to have mask mandates in general. A logistic regression predicting the existence of a mask mandate in early July from the percent of the county who voted for Trump in the 2020 presidential election showed that partisanship is a very significant predictor of a county-wide mask mandate ($z = -13.87$, $p < 0.0001$). An increase of 1 percent in 2020 votes for Trump is associated with an odds ratio of 0.962, or a multiplicative decrease of 2.616%. As shown in the plot below, this model meant that a county that was 80% Democratic had about an 82% chance of having a mask mandate, a county equally split between the counties had about a 57% chance of a mandate, and a county that was 80% Republican had about a 27% chance of a mandate.



Clearly, partisanship is a strong predictor of whether a county has a mask mandate in the first place (which makes sense given that leaders in local governments who make masking policies naturally represent the political beliefs of their constituents). Can the difference in mask-wearing by mandate be explained away by the political leaning of the county? To address this question, we conducted an ESS F test to compare a linear model predicting mask wearing from just partisanship on the training set (`lm_trump_quad`) to a linear model with both partisanship and a mask mandate as a predictor (`lm_trump_quad_mandate`). In both these cases, partisanship was fitted using a polynomial of degree 2, as this seemed to better account for regression assumptions and explain non-linearities in the data. We found that the addition of the mask mandate was statistically significant ($F = 356.78$, $p < 0.0001$, as shown below): After controlling for partisanship, the predicted increase in `pct_mask` in the presence of a mandate was 6.392.

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	2404	179670.93				
2	2403	156980.63	1	22690.31	347.33	0.0000

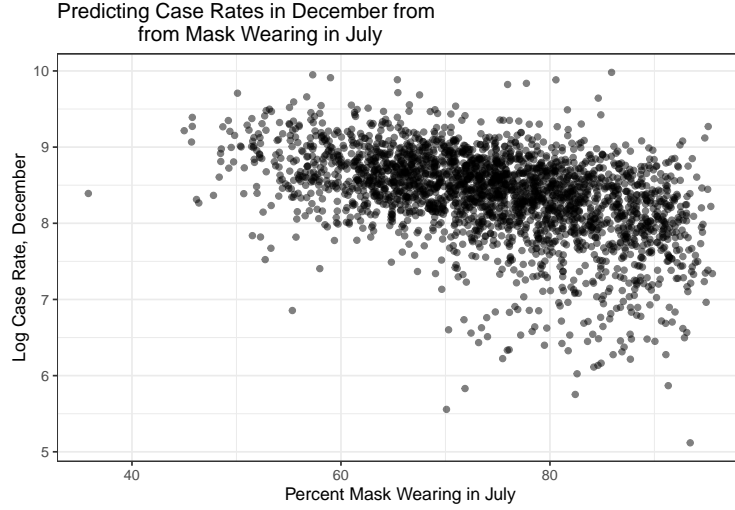
In order to consider whether the relationship between mandates and mask-wearing varied based on partisanship, we considered adding an interaction term between these two variables (`lm_trump_quad_mandate_interact`). The addition of this term was not significant in an ESS-F test ($F_2 = 1.5243$, $p = 0.218$), suggesting that the effect of a mask mandate does not depend on partisanship.

One other question that arose was whether county-wide mandates still increased mask wearing when a statewide mandate was already in place. In order to account for this, we refit our previous model adding in a term for statewide mandate, and refit again with an interaction term between state and county mandates. In each of these models, the addition of the new terms added significant predictive power to the model as determined by an appropriate ESS F Test ($F = 144.35$, and $F = 42.791$ for each of the successive tests). Focusing more closely on the interaction term, the 95% bootstrap confidence interval for the predicted increase in mask-wearing for a county with a county-wide mandate when a state mandate is already in place is (0.5037, 3.0058). This suggests that, while there is a slight benefit to having both a county and a state mask mandate in place, it is not nearly as large as the increase in mask wearing associated with going from having no mandates to a single one.

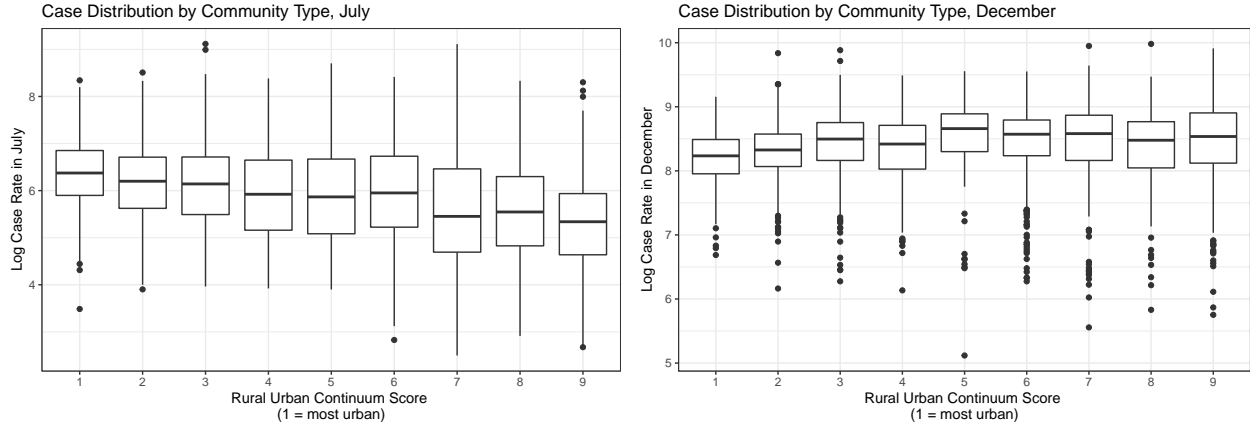
Future COVID-19 Spread and Mask Wearing Behavior

Thus far, we have concentrated our efforts on exploring what variables contribute to mask-wearing behavior as measured by the NYT survey, with the underlying idea that mask-wearing behavior is a useful thing to know about because it directly impacts the spread of the pandemic. As a final exploration in our project, we wanted to put this claim to the test and investigate whether mask wearing behavior as measured by a survey in July could provide any predictive power in determining COVID-19 case rates in a county in December. While the link between masks and COVID-19 is causal (i.e. wearing a mask directly increases or reduces your chances of obtaining the virus), this is in no way a causal model, as behaviors may have changed greatly within the past 6 months and we are dealing with observational data to begin with. Rather, we intend to treat `pct_mask` as a proxy for the “carefulness” of a county, as measured early in the pandemic.

As expected, the current case rate is quite right-skewed; most counties have between 0 and 10,000 cases per 100,000 residents, but some counties have as many as 20,000 cases per 100,000 residents. For this reason, we log transformed `case_rate_december`, our y variable for this investigation. An initial plot of the transformed case rate variable against `pct_mask` shows a weak negative association, albeit with many influential points, as there seemed to be several counties with very high mask-wearing scores that had an especially low case rate in December.



Rather than fitting an OLS model, we decided to begin with a mixed effects model to predict cases by state using random intercepts. This is because different states have enacted different policies (i.e. shutdowns, prevalence/availability of tests, etc.) that could help explain differing numbers in current cases. We also considered adding density to this initial model (as one would hypothesize that more people coming into contact with one another more frequently would predict the spread of the virus), but decided not to, as it did not add any predictive power with the ESS F-Test. This was surprising, as density was a significant predictor in our earlier models for `pct_mask` as measured in July — what had changed? As illustrated by the boxplots below, while urban counties had far higher case rates than rural counties back in July, now in December, the case rates have more or less evened out and population density is no longer a useful predictor of what counties will have higher rates of spread.



Based on this model, the average case rate among states is estimated to be 3,603.785 per 100,000 residents. Next, we fit another random intercepts model which included `pct_mask` as a fixed effect; in this model, on average, the an additional 1 percent in mask wearing within a county in July was associated with an ($e^{(-0.003855)}$) multiplicative change, or 0.4% multiplicative decrease in the case rate per 100,000 people. This result, while quite small in magnitude, added significant explanatory power to the model ($\chi^2_1 = 11.393$, $p = 0.0007 < 0.05$); it also held when we calculated the RMSE for our two models on test data (it decreased RMSE from 0.3925893 to 0.3905824). This suggests that how careful a county was about wearing masks in July is an important predictor for case rates in December even after controlling for statewide differences, suggesting that fostering cultures of county-wide mask wearing should be a priority.

To conclude this section, we wanted to take a closer look at the counties that this mixed effects model had mis-predicted the most (i.e. the counties with the largest residuals). Counties that had lower rates than their mask wearing score would suggest tended to be small, rural communities in the Southwestern United States; while their success in controlling case rates could simply be due to chance given their small size (there are plenty of small towns that have been unsuccessful, so it could be due to random chance alone), it could be worthwhile to do further research on these counties in order to determine if they instituted any policy strategies at a local level that could be applicable to other scenarios. Likewise, further investigating the counties that comparatively did the worst (which tended to be midsize cities, spread out around the nation) given their mask wearing score could also be an interesting case study about what policy practices were ineffective.

Here are the counties with the most negative residuals (indicating COVID-19 rates below what their mask-wearing status would predict).

	county_name	state	pct_mask	pop_2019	ru_continuum	residuals	log_cases_dec
2029		AK	82.42	1592.00	9.00	-5.28	5.75
1461	King County	TX	70.30	272.00	9.00	-4.43	6.60
2229	San Juan County	WA	91.33	17582.00	9.00	-4.40	5.87
1149	Borden County	TX	78.55	654.00	8.00	-4.24	6.64
135	Sierra County	CA	76.10	3005.00	8.00	-4.18	6.34
1229	Daggett County	UT	55.35	950.00	9.00	-4.02	6.85

And here were the counties with the most positive residuals (indicating COVID-19 rates above what their mask-wearing status would predict).

	county_name	state	pct_mask	pop_2019	ru_continuum	residuals	log_cases_dec
1348	Lincoln County	AR	80.58	13024.00	3.00	3.52	9.88
2080	Rockland County	NY	89.97	325789.00	1.00	3.57	8.84
822	Chattahoochee County	GA	77.75	10907.00	2.00	3.70	9.84
262	Luce County	MI	84.62	6229.00	7.00	3.76	9.64
1676	Malheur County	OR	61.75	30571.00	6.00	4.00	9.01
226	Crowley County	CO	85.90	6061.00	8.00	4.96	9.98

Conclusions and Discussion

Fitting our initial model to predict `pct_mask` from all the main effects, we found that the coefficients of the β terms universally supported our various hypothesis, other than the coefficient for `college`, which was negative (albeit with small magnitude). This suggests that the trends in mask wearing behavior that researchers have noted at an individual level likewise hold at a county level and suggests that mask-wearing campaigns specifically focused on younger white people in low-income, low-density areas could be merited. Focusing on the association of `pct_trump_2020` with mask wearing, our evidence supports research claims that partisanship is the single most important predictor of mask wearing, as shown in our variable importance plot for the random forest model. This makes sense, as partisanship directly contributes to rhetoric relating to efficacy of masks in a way that demographic variables do not. However, it is far from the *only* important predictor, as models fitted using solely demographic predictors out-predicted models using only political predictors on both the train and test set.

Having established the relative importance of partisanship in predicting mask-wearing behavior, we further investigated the relationship between mask mandates and mask-wearing. Although Democratic counties were far more likely to have a mask mandate at all (in July), there was not an association between political party and the degree of response to a mask mandate. In other words, among the Republican counties that did have mandates, the response was similar to Democratic counties with mandates (an increase of about 6% in the prevalence of mask wearing). If there was already a state mandate in place, the addition of a

countywide mandate was not predicted to be associated with as large of a bump in mask wearing (about 2% vs. 6% from before), but there was still a significant increase associated with having the two mandates in place. Overall, our findings suggested that county and statewide mandates are associated with meaningful increases in mask-wearing behavior regardless of the politics of the state or county in question.

Armed with a better understanding of which counties were wearing masks in July, we flipped our response `pct_mask` into a predictor in order to determine whether there was an association between mask wearing in July and case rates in December. Although there have been many shifts in the way the country has experienced the virus within the past six months (notably, population density is no longer a significant predictor of COVID-19 case rates the way it was in July), mask-wearing behavior in July still provided explanatory in power when added to a model to predict December case rates from random intercepts based on states. This finding suggests that further research into measures of the “carefulness” of a county, which could extend beyond mask-wearing rates and into to public health messaging, local stay-at-home orders, and business regulations, could be a useful way to predict the rate of COVID-19 spread within counties.

While we hope our conclusions about mask-wearing and partisanship still have meaningful implications, it is important to recognize that there are many limitations to our data and the type of claims that we can make. First of all, none of our findings should be thought of as causal claims; rather, the fact that significant associations exist between political background, demographic characteristics, and mask wearing should serve as an impetus for further investigation social science researchers. While there *is* a proven causal relationship between mask-wearing and the spread of COVID-19 as shown in numerous scientific experiments, our final model that used `pct_mask` to predict case rates in December should not be interpreted as causal (Cheng 2020). This is due to (1) the extreme lag in time between these two measurements and (2) the fact that this was not an experiment, and that people who wear masks are likely more careful about their social distancing conduct in other ways as well. Finally, our findings are limited due to the nature of using survey data to as our response variable, as surveys may suffer from undercoverage or response bias. One interesting extension to this project would be to collect data about the actual prevalence of mask wearing (as measured by manually taking and recording observations at a public location, as researchers in Hong Kong did) and seeing how this number compares to the self-reported frequency of mask wearing in various US counties. This would help give a more realistic understanding of the relationship between mask wearing and the spread of COVID-19 at a county level.

Bibliography

- Centers for Disease Control and Prevention. *Disparities in Incidence of COVID-19 Among Underrepresented Racial/Ethnic Groups in Counties Identified as Hotspots During June 5–18, 2020 – 22 States, February–June 2020*. (2020, August 20). <https://www.cdc.gov/mmwr/volumes/69/wr/mm6933e1.htm>
- Cheng, V., Wong, S., Chuang, V., So, S., Chen, J., Sridhar, S., . . . Yuen, K. (2020, July). *The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7177146/>
- Haischer, M., Beilfuss, R., Hart, M., Opielinski, L., Wrucke, D., Zirgaitis, G., . . . Hunter, S. (2020, October 15). *Who is wearing a mask? Gender-, age-, and location-related differences during the COVID-19 pandemic*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7561164/>
- Igielnik, R. (2020, July 27). *Most Americans say they regularly wore a mask in stores in the past month; fewer see others doing it*. <https://www.pewresearch.org/fact-tank/2020/06/23/most-americans-say-they-regularly-wore-a-mask-in-stores-in-the-past-month-fewer-see-others-doing-it/>
- Kessel, P., & Quinn, D. (2020, November 02). *Both Republicans and Democrats cite masks as a negative effect of COVID-19, but for very different reasons*. <https://www.pewresearch.org/fact-tank/2020/10/29/both-republicans-and-democrats-cite-masks-as-a-negative-effect-of-covid-19-but-for-very-different-reasons/>
- Mann, B. (2020, July 08). *More States Require Masks In Public As COVID-19 Spreads, But Enforcement Lags*. <https://www.npr.org/2020/07/08/888499285/more-states-require-masks-in-public-as-covid-19-spreads-but-enforcement-lags>
- Milosh, M., Painter, M., Sonin, K., Van Dijke, D., & Wright, A. (2020, August 03). *Unmasking Partisanship: Polarization Undermines Public Response to Collective Risk*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3664779
- Neff, Todd F. (2020, August 18). *The science says: Wear a mask*. <https://www.uchealth.org/today/wear-a-mask-the-science-that-supports-masks/>
- Tonmccg. (n.d.). *Alaska County Level Results Issues*. https://github.com/tonmccg/US_County_Level_Election_Results_08-20/issues/2
- Ueki, H., Furusawa, Y., Iwatsuki-Horimoto, K., Imai, M., Kabata, H., Nishimura, H., & Kawaoka, Y. (2020, October 28). *Effectiveness of Face Masks in Preventing Airborne Transmission of SARS-CoV-2*. <https://msphere.asm.org/content/5/5/e00637-20>
- Victor, D., Serviss, L., & Paybarah, A. (2020, October 02). *In His Own Words, Trump on the Coronavirus and Masks*. <https://www.nytimes.com/2020/10/02/us/politics/donald-trump-masks.html>

Appendix

Data Cleaning and Compilation Process

We compiled data from a total of 8 different sources in order to obtain the combination of demographic, political, and public health data that we desired. We began with the data from the NYT, which was described in our main paper. Most of the cleaning was simply a matter of renaming variables, making raw counts into rates, creating indicators for countywide and statewide mandates, and merging based on county FIPS code. Once we had compiled our predictor variables and done exploratory data analysis, we transformed them within the dataset in order to help us keep track of them, renaming accordingly (so a variable like density, which required a log transformation, became `log_density` in the dataset). For more information on each of our predictors and what dataset they came from, see the table below:

Name	Source
<code>countyfp</code>	New York Times
<code>always</code>	New York Times
<code>frequently</code>	New York Times
<code>sometimes</code>	New York Times
<code>rarely</code>	New York Times
<code>never</code>	New York Times
<code>cases_02</code>	New York Times
<code>cases_14</code>	New York Times
<code>cases_27</code>	New York Times
<code>pop_2019</code>	United States Census Bureau
<code>ru_continuum</code>	United States Census Bureau
<code>density</code>	county_level_election.csv from class
<code>pct_less_than_hs</code>	2014-18 American Community Survey
<code>pct_hs</code>	2014-18 American Community Survey
<code>pct_some_college</code>	2014-18 American Community Survey
<code>pct_college</code>	2014-18 American Community Survey
<code>pct_poverty</code>	U.S. Census Bureau, Small Area Income and Poverty Estimates (SAIPE) Program
<code>pct_female</code>	U.S. Census Bureau
<code>pct_black</code>	U.S. Census Bureau
<code>pct_native</code>	U.S. Census Bureau
<code>pct_hispanic</code>	U.S. Census Bureau
<code>pct_seniors</code>	U.S. Census Bureau
<code>pct_trump_2016</code>	county_level_election.csv from class
<code>pct_trump_2020</code>	Scraped by GitHub user <code>tonmcg</code> from Fox News, Politico, and New York Times
<code>dem_governor</code>	National Governor's Association
<code>state_mandate</code>	Axios
<code>county_mandate</code>	Harris Institute of Public Policy

Handling Missing Data

As mentioned in our paper, we had 127 rows with at least one missing predictor value. 28 of these rows were from Alaska, all of which were missing county-level data for both the 2016 and 2020 elections due to the fact that Alaska reports election results using boroughs instead of counties. Because partisanship was such a key variable in many of our models and because this data was not easily accessible for Alaska (even the cleaned data sets we used in class did not have it), we decided to exclude Alaska from our models entirely, understanding that any conclusions we came to would not be able to be reasonably generalized to this state.

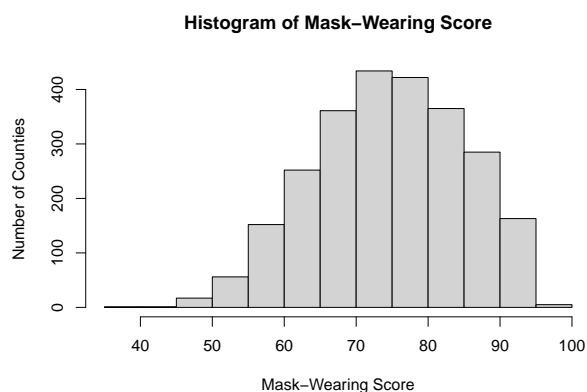
Beyond the total omission of political data from Alaska, the next most worrying part of our predictor data was the fact that case rates were missing for five extremely populous counties in New York City with over a million residents each. Upon further investigation, we realized that this was because (for reasons unknown to us) New York City reports its COVID-19 data in most sources as a full unit rather than as the 5 different

counties it can be broken up into. In order to remedy this, we manually inputted values for current December case rates by searching for them online, allowing us to include them in our mixed model for predicting current case rates from mask-wearing in July. Unfortunately, we could not easily find COVID-19 data for these counties available in July at the county level, but we ultimately did not end up using this predictor in any of our final models, limiting the adverse impact of this missing data.

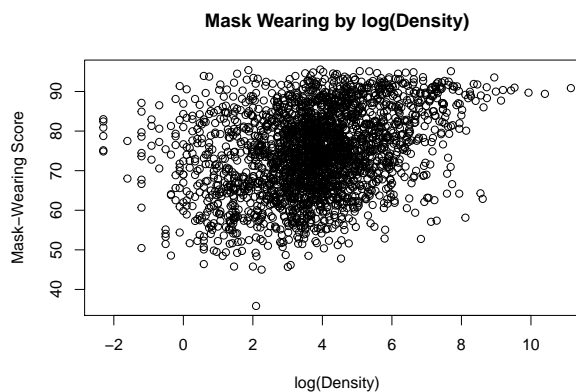
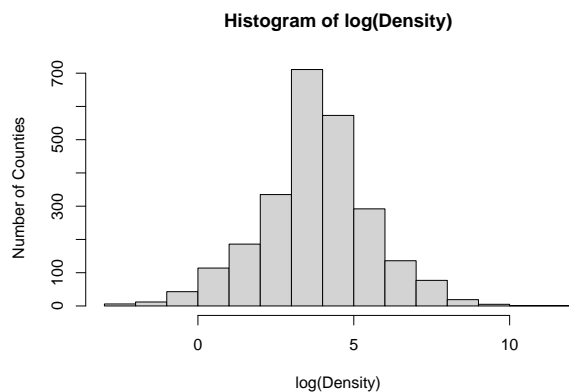
The rest of the missing data primarily came in the form of missing case/death rates from July in small, rural counties; more than likely, these counties simply were not publishing information about their case rates at that time, or they did not yet have any cases at all. After conducting a two sample t-test, we determined that these counties were more rural, male, Republican, older, and slightly more college educated, more white, and less likely to wear masks than the counties in the full data set. There was not a statistically significant difference in the poverty rate. Although this finding implicates that the data we ended up using slightly underestimates small, rural, Republican counties, the entire premise of analyzing data at the county level inherently *overrepresents* these counties, because they have far smaller populations than urban, Democratic counties. To illustrate, even though counties with missing data (not including NY or AK) represented 2.96% of the *counties* in our dataset, they contained just 0.11% of the current US *population*. While this could have been addressed by weighting our data by county population, doing so negatively affected many of our diagnostic plots (i.e. made them nonlinear) and prevented our mixed model from converging. All that is to say, we did not end up weighting by population, so our missing county-level data slightly underestimates the influence of rural counties, but puts us a little closer to the reality of the American population at large.

Exploratory Data Analysis

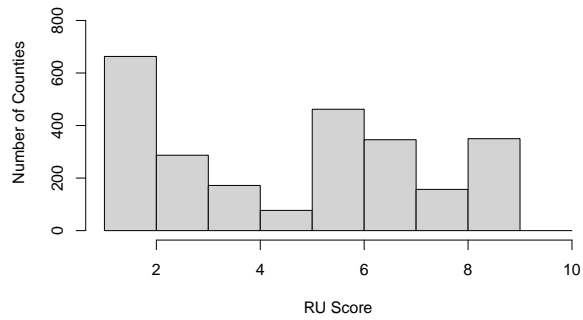
Distribution of Response Variable:



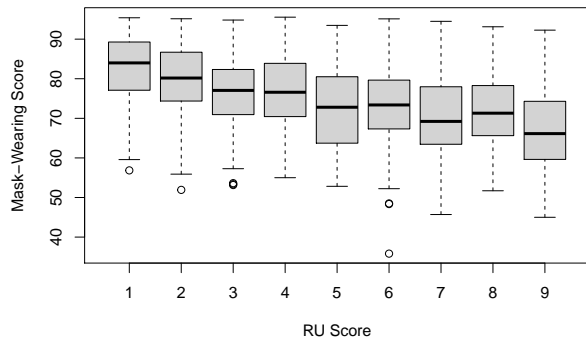
Distributions of Predictors with Necessary Transformations:



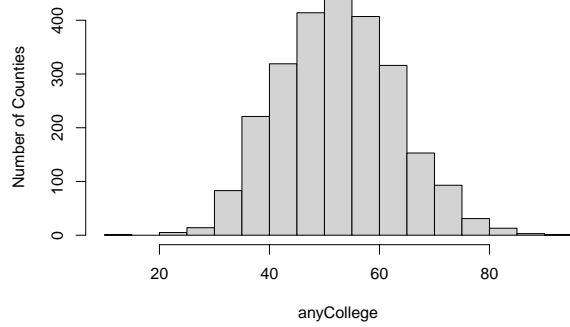
Histogram of Rural-Urban Continuum Score



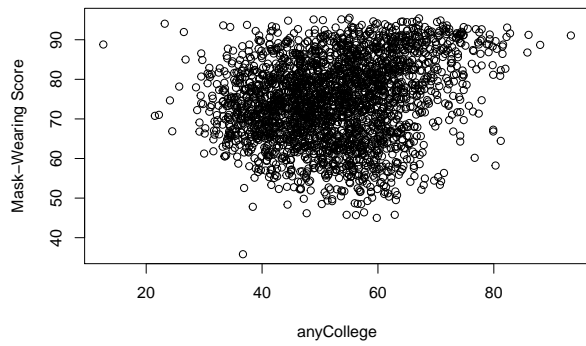
Mask Wearing by Rural-Urban Continuum Score



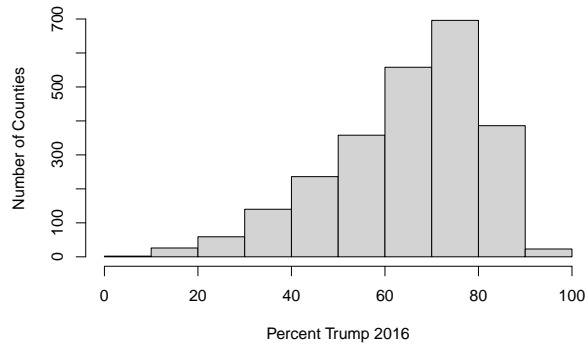
Histogram of anyCollege



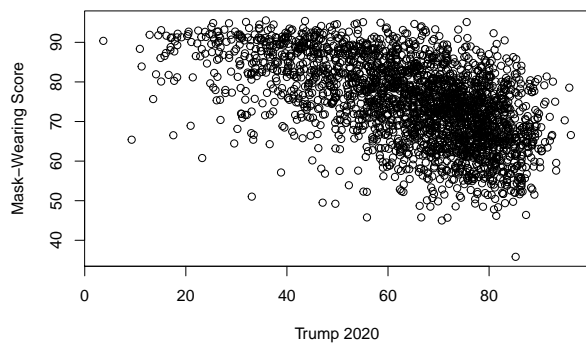
Mask Wearing by anyCollege

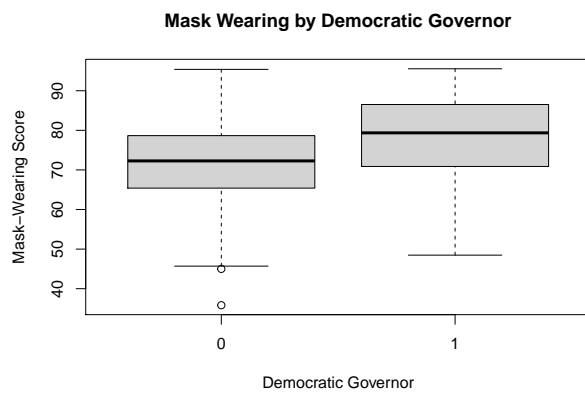
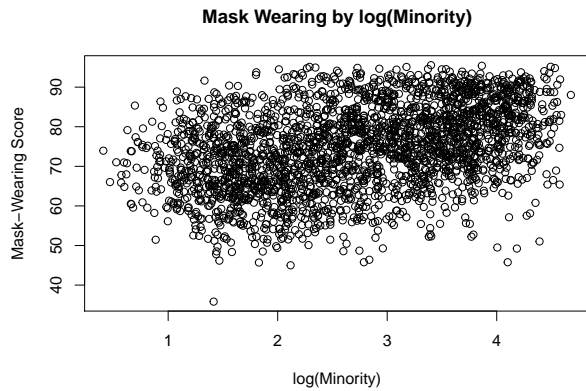
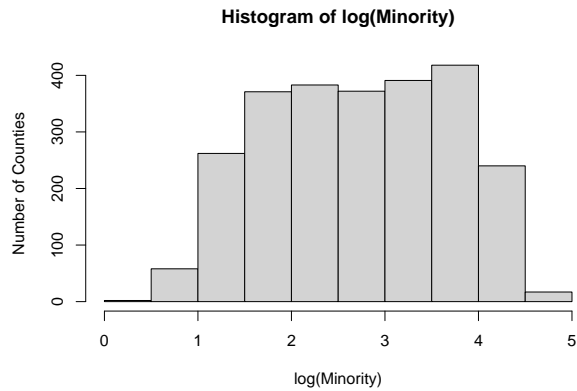


Histogram of Percent Voted for Trump 2016 (untransformed)



Mask Wearing by pct_Trump_2020





Section 1: Is Mask-wearing Political?

Coefficient Comparison between linear full model and linear selected model:

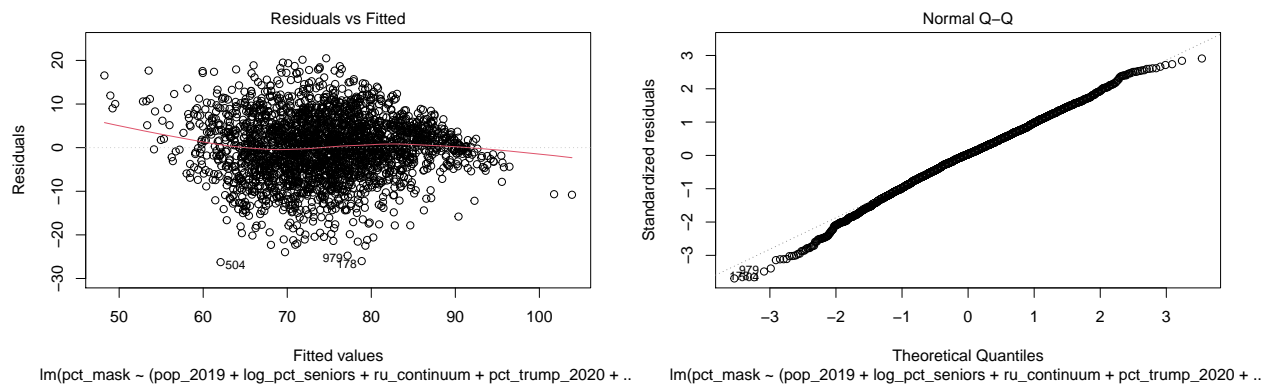
Predictor	Full Model	Interaction Model
pop_2019	0	2e-05
log_pct_seniors	8.76441	-1.05173
ru_continuum	-1.23539	-6.58297
pct_trump_2020	-0.18145	0.1561
log_pct_minority	2.97637	22.39559
dem_governor	3.97165	10.48379
pct_anycollege	-0.06162	-0.16007
log_density	0.77165	2.29168
pct_female	-0.32704	0.25898
log_pct_poverty	-1.90413	-11.16163

Summary of Full Model

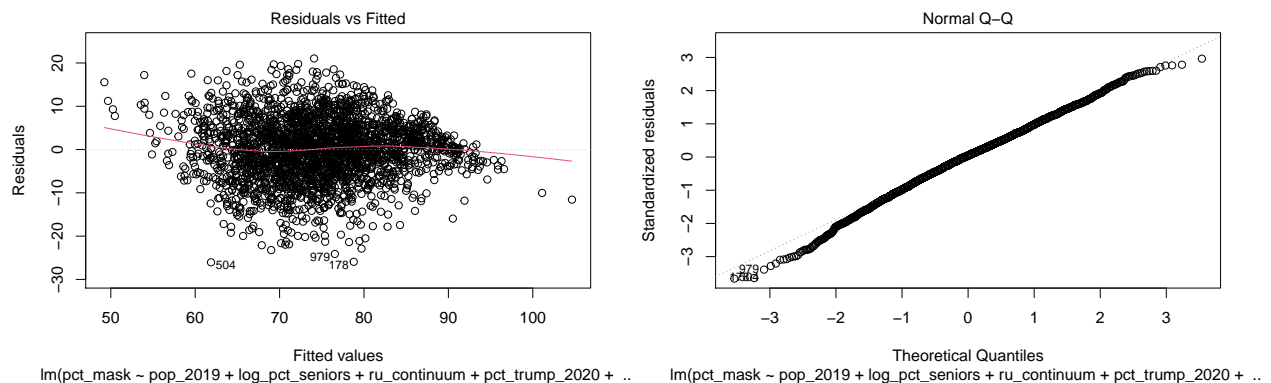
```
##
## Call:
## lm(formula = pct_mask ~ pop_2019 + log_pct_seniors + ru_continuum +
##     pct_trump_2020 + log_pct_minority + dem_governor + pct_anycollege +
##     log_density + pct_female + log_pct_poverty, data = mask_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -26.4808 -4.7311 0.3221 5.1499 21.2564
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.823e+01  4.832e+00  16.190 < 2e-16 ***
## pop_2019      1.051e-06  5.113e-07   2.056 0.039848 *
## log_pct_seniors 8.764e+00  7.749e-01  11.311 < 2e-16 ***
## ru_continuum  -1.235e+00  7.529e-02 -16.409 < 2e-16 ***
## pct_trump_2020 -1.815e-01  1.329e-02 -13.651 < 2e-16 ***
## log_pct_minority 2.976e+00  2.075e-01  14.342 < 2e-16 ***
## dem_governor   3.972e+00  3.254e-01  12.204 < 2e-16 ***
## pct_anycollege -6.162e-02  2.172e-02  -2.837 0.004593 **
## log_density     7.717e-01  1.005e-01   7.678 2.31e-14 ***
## pct_female     -3.270e-01  7.326e-02  -4.464 8.41e-06 ***
## log_pct_poverty -1.904e+00  5.670e-01  -3.358 0.000796 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.536 on 2472 degrees of freedom
## (31 observations deleted due to missingness)
## Multiple R-squared:  0.4751, Adjusted R-squared:  0.473
## F-statistic: 223.8 on 10 and 2472 DF, p-value: < 2.2e-16
```

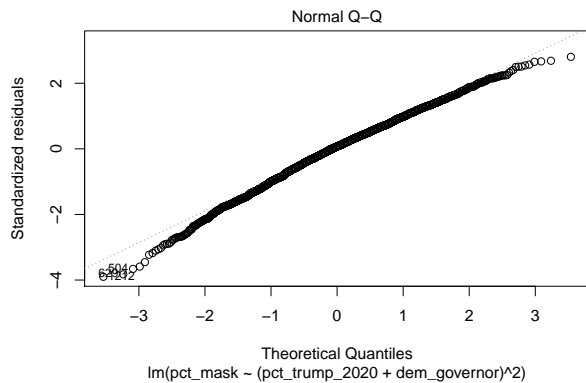
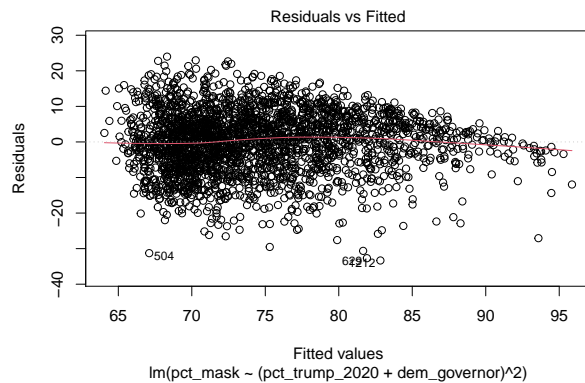
Interaction Model Assumptions Checks:



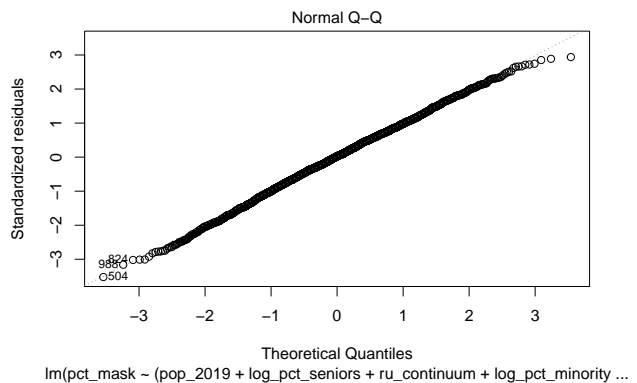
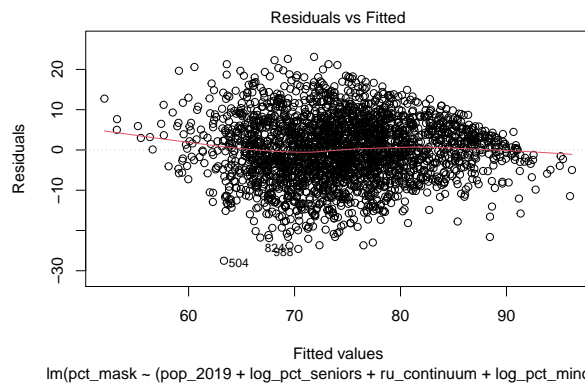
Selected Model Assumption Checks:



Political Model Assumption Checks:



Apolitical Model Assumption Checks:



Section 2: Mask Mandate Models and Diagnostics

Here is the summary and output for our logistic model (called `logit_1`)

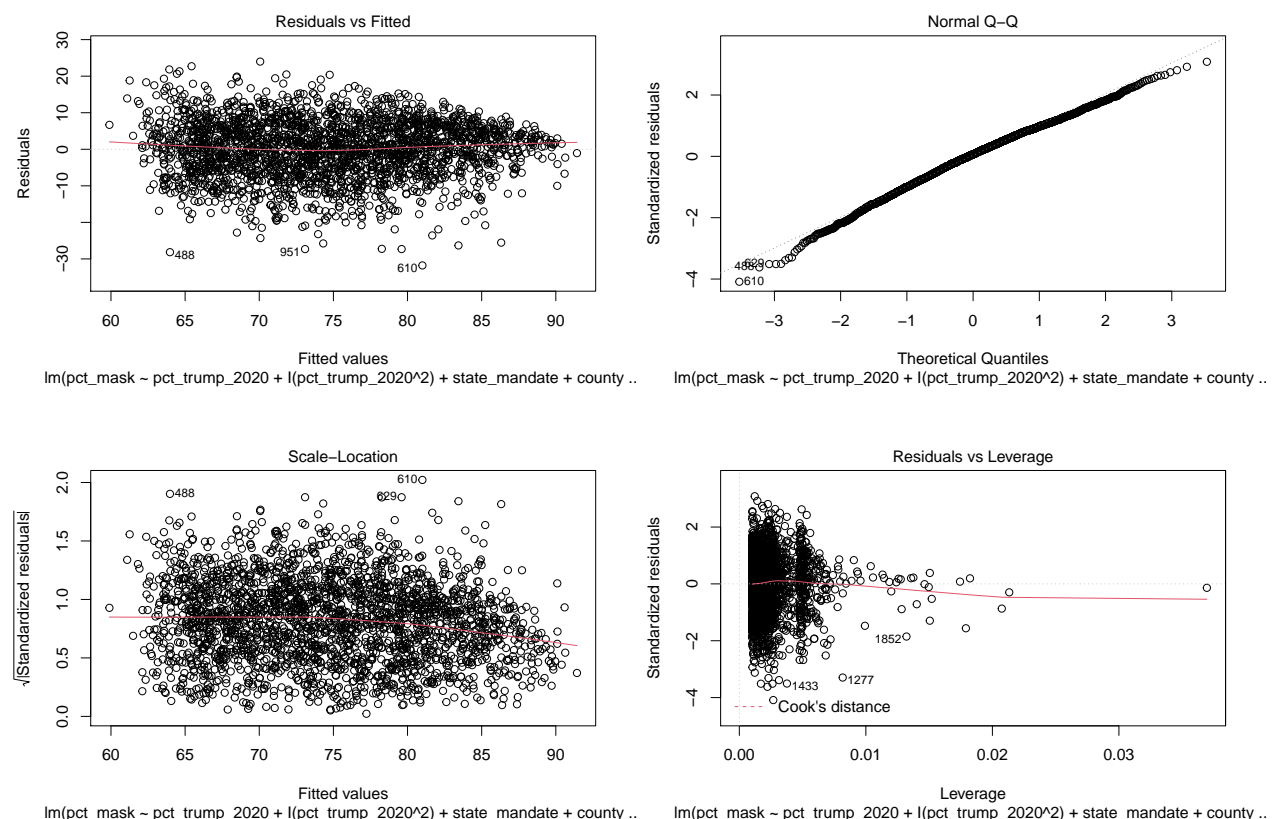
```
##
## Call:
## glm(formula = county_mandate ~ pct_trump_2020, family = "binomial",
##      data = data_train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0133  -1.0387  -0.8004   1.1614   1.7020
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.437969   0.189008  12.90  <2e-16 ***
## pct_trump_2020 -0.039170   0.002823 -13.87  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 3432.9  on 2480  degrees of freedom
```

```
## Residual deviance: 3214.4 on 2479 degrees of freedom
## (33 observations deleted due to missingness)
## AIC: 3218.4
##
## Number of Fisher Scoring iterations: 4
```

Here is the formula and output for our final model to predict mask wearing from state and county mask mandates with an interaction term (called `lm_trump_state_interact`):

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	82.8495	1.7211	48.14	0.0000
pct_trump_2020	-0.0864	0.0596	-1.45	0.1475
I(pct_trump_2020^2)	-0.0016	0.0005	-3.12	0.0018
state_mandate	7.4699	0.5883	12.70	0.0000
county_mandate	6.3295	0.4351	14.55	0.0000
state_mandate:county_mandate	-4.8829	0.7464	-6.54	0.0000

Next, we needed to check the assumptions for this model; diagnostic plots are shown below:



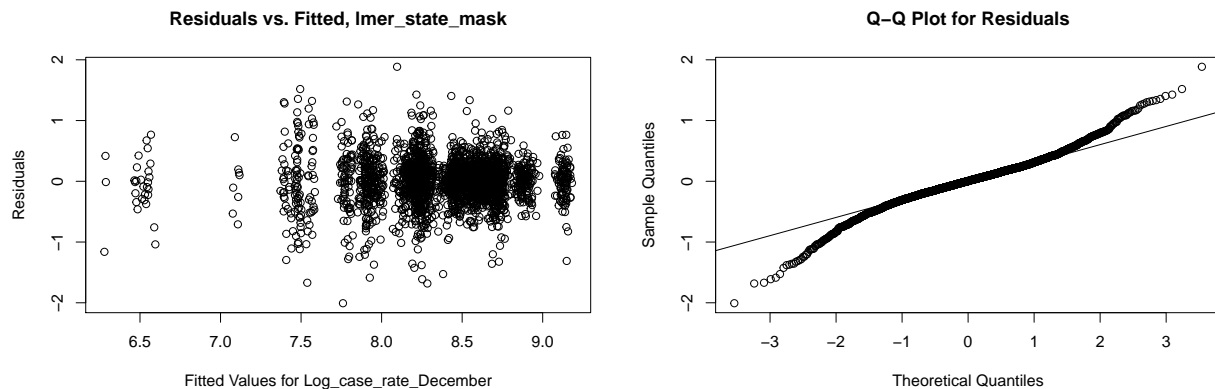
Overall, diagnostics look fairly good, particularly linearity and constant variance. There is some slight deviation in the qqplot in the tails, but nothing too worrisome. We do have an outlier in the x direction, as shown on the leverage vs. standardized residual plot, but it does not have a large residual and is well within the Cook's distance lines, suggesting it is not too influential.

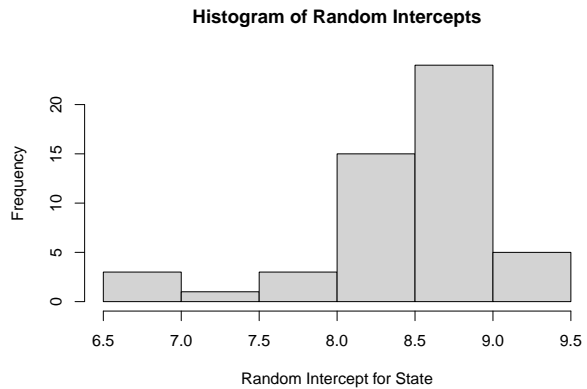
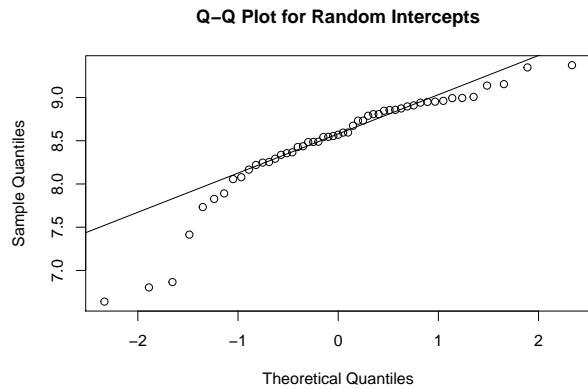
Section 3: December Case Prediction Assumption Checking and Models

Here is the summary and output for our final mixed effects model, `lmer_state_mask`:

```
## Linear mixed model fit by maximum likelihood ['lmerMod']
## Formula: log_case_rate_december ~ pct_mask + (1 | state)
## Data: data_train
##
##      AIC      BIC    logLik deviance df.resid
## 2498.8  2522.2 -1245.4  2490.8    2507
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -5.2847 -0.5209  0.0077  0.5433  4.9627
##
## Random effects:
## Groups   Name      Variance Std.Dev.
## state    (Intercept) 0.3526   0.5938
## Residual              0.1443   0.3798
## Number of obs: 2511, groups: state, 51
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  8.487461   0.121356  69.939
## pct_mask     -0.003855   0.001134  -3.399
##
## Correlation of Fixed Effects:
##          (Intr)
## pct_mask -0.720
```

And here are the diagnostic plots for the mixed model. Note that in addition to the standard assumptions of linearity, normality of residuals, and constant variance, we must also check the assumption that the random intercepts for states are normally distributed.





Linearity seems to be a reasonable assumption, as points are clustered randomly around 0 (note: the visible vertical clusters appear because the magnitude of the random effect of state is much bigger than that of the fixed effect of `pct_mask`, thus ensuring that the points for a state fall more or less in a line). There does not appear to be evidence of heteroskedasticity. Normality of residuals looks fairly good; they deviate a bit more than expected at the tails, but do so in a symmetric way. The qqplot for the distribution of the random effects for state is slightly concerning in that it is left skewed (you can also see this in the histogram); more than anything else, this suggests that a log transformation on case rates may have been an over-correct. While this will not bias our estimates, it could lead to unrealistic standard errors for the random effect of `state`, which we were wary of when interpreting this model.