Individual Characteristics Predicting Anti-Immigrant Attitudes Among Americans

Virginia Choi (vavkc@umd.edu), *University of Maryland*

Joseph Mandwee (jmandwee@mail.uni-mannheim.de), *University of Mannheim*

Yusang Zhang (yusazhan@mail.uni-mannheim.de), *University of Mannheim*

17/5/2020

Introduction

Anti-immigrant attitudes among individuals in western nations play a fundamental role in shaping both government policy towards migrants, as well as influencing the way in which these outsiders are received into the host society. Once arrived, migrants face the extraordinary challenges of integration into their new countries, and under the constant threat of racism, discrimination, and xenophobia by locals. Recently, much research has been devoted to understanding these factors in western nations. There have also been many studies that attempt to theorize the reasons for, or motivations behind, negative attitudes towards immigrants. For example, competitive threat and prejudice towards racial minority groups in Europe have been linked to an increase in negative or xenophobic attitudes towards all other immigrants in Europe as well, and the effect of this aforementioned racial prejudice has been shown to increase in relation to the size of the minority population (Gorodzeisky & Semyonov, 2016). This is one indicator that anti-immigrant attitudes have less to do with the identity of the group entering and much more to do with the personal characteristics of the individuals that comprise the local population of these western nations. Furthermore, some research evidence suggests that Big Five personality traits influence anti-immigrant attitudes as well (Gallego & Pardos-Prado, 2013), directly pointing to individual characteristics in the formation and exhibition of anti-immigrant sentiments among members of western society.

Although social scientists do have a somewhat concrete understanding of prejudice and discrimination towards outsiders living within western societies, there is a much weaker and less developed body of literature examining what characteristics are associated with individuals who hold anti-immigrant or xenophobic attitudes, and very few pieces of research have attempted to assess personal, individual characteristics (i.e., values, beliefs, socioeconomic status, etc.…) that impact these attitudes. Nevertheless, there are some studies that point to Social Dominance Orientation (SDO) and authoritarianism in addressing anti-immigrant sentiments. Other research points to socioeconomic factors and political leanings as the influencers of these views. These findings are briefly summarized in the literature review below. The following study then attempts to provide further insights into personal characteristics that predict anti-immigrant attitudes among individuals in the United States, a western nation that is often noted by social scientists for its exceptional and historical openness to immigration and integration. To do so, this research will utilize information from the General Social Survey, which contains tens of thousands of points of data and hundreds of individual variables. We will run the gathered data through a machine learning model to find individual characteristics that meet our criteria in determining which ones play a role in predicting anti-immigrant attitudes. Lastly, we will convey our findings in the graphics and results section that follow and conclude with a brief discussion, limitations to our findings, and recommendations for further research.

Literature Review

According to Social Dominance Theory, prejudice is a type of legitimizing myth that justifies discrimination against low-status individuals in order to maintain a group-based hierarchy and to ensure the dominance of high-status members of a society, thereby maintaining one's relative position within that hierarchy. Across the world, there is a strong link between individuals' Social Dominance Orientation (SDO), prejudice, and discrimination against immigrants and other groups (Küpper, Wolf, & Zick, 2010). In Europe, lower-income individuals were found to be "more prone" to a higher Social Dominance Orientation, to hold weaker diversity beliefs, and stronger anti-immigrant prejudices; individuals with a higher Social Dominance Orientation were also more likely to discriminate against immigrants and hold stronger anti-immigrant attitudes overall (Küpper et al., 2010). Therefore, an individual's Social Dominance Orientation, beliefs on diversity, prejudices, and socioeconomic status all appear to be highly imperative personal characteristics that influence one's anti-immigrant attitudes.

A correlational link also exists between Social Dominance Orientation and right-wing authoritarianism, both predicting support for anti-immigrant, ethnic persecution (Thomsen, Green, & Sidanius, 2008), but for different reasons. Migrants' refusal to assimilate into the culture of the dominant group provokes the right-wing authoritarian aggression because it violates conformity in the ingroup (a.k.a. ingroup conformity hypothesis), while aggression from individuals with a high Social Dominance Orientation is more so associated with migrants who do actually assimilate because this blurs the lines and boundaries between ranks in the hierarchy (a.k.a. status boundary enforcement hypothesis); the research data supports both of the hypotheses (Thomsen et al., 2008). These results provide further evidence that individual factors lead to different reasons for outgroup prejudice and discrimination, while both Social Dominance Orientation and right-wing authoritarianism influence the ethnic persecution of immigrants, and therefore, anti-immigrant attitudes at its base.

Moreover, Oyamot Jr., Fisher, Deason, & Borgida (2012) studied the influence of social norms and humanitarian values on authoritarians' attitudes toward immigrants utilizing a survey experiment across positive, negative, and mixed social norm conditions. In a previous study, the authors found correlational evidence that "clear social norms for attitudes toward an outgroup (favorable or unfavorable) influence the authoritarianism–attitude relationship in the direction of the norm, and in the absence of clear social norms, endorsement of humanitarian–egalitarian values attenuate the intolerant tendencies of authoritarians" (Oyamot Jr. et al., 2012). The authors found that authoritarianism was negatively associated with sentiments towards immigrants under the negative norm condition (in which study participants were told that Americans generally had negative opinions towards immigrants); under the positive norm condition (participants were told that Americans generally had positive opinions towards immigrants), intolerance among authoritarians was reduced; under the mixed social norm scenario, authoritarians' attitudes depended on their humanitarian values: humanitarian authoritarians held positive attitudes while non-humanitarian authoritarians held the "most-negative" attitudes (Oyamot Jr. et al., 2012). Therefore, authoritarianism not only depends on context in the formulation of anti-immigrant attitudes, but humanitarian values as a personal characteristic also play an immense role in the display of positive or negative attitudes towards immigration.

Aside from Social Dominance Orientation, authoritarianism, and humanitarian values, multiple socioeconomic, occupational, political, and geographical factors also influence anti-immigrant attitudes among westerners. As one example, Belgians with contextual characteristics like lower occupational status and unemployment were found to be more likely to vote for the anti-immigrant party in Flanders (Rink, Phalet, & Swyngedouw, 2009). These findings draw a direct link between anti-immigrant attitudes and voting behaviors, which in turn reflect political stance. The authors of this study also predicted that the anti-immigrant party would receive more popularity and more votes with increasing immigration numbers and a higher unemployment rate, giving support to the Group Threat Theory's concept of perceived collective threat mentioned in the study (Rink et al., 2009), the results of which confirmed a curvilinear relationship between migrant population size and anti-immigrant voting. The effect of an increased immigrant population on anti-immigrant voting was also significantly reduced among more educated voters (Rink et al., 2009), suggesting at least a partial role for one's level of education in anti-immigrant attitudes or behaviors that are eventually reflected at the ballot box. While right-wing populism and anti-immigrant parties continue to gain popularity across Europe and the United States, one may further argue that the factors found to be associated with anti-immigrant voting in this Belgian study likewise reflect the voting behaviors and attitudes of anti-immigrant voters in other western nations, from the election of Donald Trump in the U.S. to the rise of the *Alternative für Deutschland* in Germany. The data gathered in another European study conducted in Germany, Austria, and Switzerland indicated that "increasing and visible diversity" was connected to negative/anti-immigrant attitudes among natives of the political right in particular (Karreth, Singh, & Stojek, 2015), yet again demonstrating the role of political beliefs on overall anti-immigrant sentiments. Using data from the General Social Survey, Garcia & Davidson (2013) also found that political ideology determined attitudes towards immigrants among rural Americans, while political affiliation determined attitudes among urban Americans; race, education level, and beliefs about immigration were significant predictors of attitudes among Americans in both rural and urban regions, while rural Americans were more likely to pronounce opposition towards immigrants overall, indicating more negative/anti-immigrant attitudes than urban Americans (Garcia & Davidson, 2013). This study demonstrates political, geographical, and even cultural differences between Americans in the formulation and projection of anti-immigrant attitudes. Rural vs. urban may seem like a basic difference in location or region, but is in fact a cultural difference in identity at a time when beliefs and sentiments are becoming ever more tribal in nature. The few factors considered in this study using the General Social Survey were used as a means to compare and contrast urban and rural Americans in their attitudes towards immigration. Nevertheless, there are many more factors that could potentially be discovered and studied in further exploration of anti-immigrant sentiments among Americans as a whole. We will attempt to pinpoint and discuss these factors in our current study.

Methods

Data Preparation. The found data used for the present study was collected from the General Social Survey (GSS), a project out of the University of Chicago's non-partisan and objective research organization, also referred to as NORC (Smith et al., 2019). The GSS has been gathering data on U.S. societal trends since 1972, annually up to 1994, then bi-annually since 1994 in even numbered years. Their sampling methodology utilizes an area probability design, which randomly selects U.S. households from diverse geographic areas (e.g., rural, urban, and, suburban), wherein select respondents per household participate in face-to-face interviews by NORC-trained staff.

For this project, we began our investigation with the cumulative version of the GSS. This data consisted of 64,814 observations and 6,108 variables ranging in time from 1972 to 2018. It is important to note that for every year of the GSS, the same questionnaire is never repeated. Many items are cut, and new items are added per wave of the GSS. Thus, in our choice of our primary outcome variable that could speak to American attitudes on immigration, we primarily considered selecting an item that allowed us to keep as many observations or cases as possible. To do this, we sought a variable that had been repeatedly used for many waves of the GSS. While about nine different GSS variables on the topic of immigration had been repeated for two to three years, only one variable had been included in seven different waves. Labelled *letin1a*, the variable referred to the question of whether people wished the number of immigrants in the U.S. at the time should be increased or reduced. This question was asked through a five-point Likert scale that was transformed into a binary variable. Responses that included "remain the same as it is," "increased a little," and "increased a little," were coded as "0" to represent neutral-to-pro immigrant views. The responses for "reduced a little" and "reduced a lot" were coded as "1" to represent anti-immigration views. We reduced the cumulative GSS down to the observations that provided a response to this immigration variable, leaving us with a sample of 11,773 cases. Since accuracy is sensitive to imbalanced classes, we also checked to make sure that the numbers found in both classes were overly uneven and found this was not the case.

Next, we dropped the variables that did not have at least 80 percent of responses from the remaining sample, which left us with 188 variables. Other final sample considerations were made, for instance, participants deemed to have had poor understanding of the interview questions were removed. Interviewer characteristics were also removed in order to ensure our inputs are focused on respondent characteristics. Separate survey items that asked about the same theme of questions (e.g., race, background, ethnicity, language spoken) were all consolidated into one cohesive indicator, while keeping the original variables. This aggregation step allowed us to have variables with fewer missing responses, since such responses are less interpretable. Variables that seemed to have no variance in their answers among the remaining sample such as "Are you the head of the household?" were dropped. All remaining variables went through a cleaning process. An example of this cleaning process is changing answers of "don't know" to their own factor level. Variables with a scale that surpassed 15 options deemed uneven or arbitrary were reset to represent a more explanatory and interpretable scale of factor levels. Although the final number of cases and features were not as high as anticipated, our final features included a diverse range of variables from the candidate the respondents voted for in the presidential election of 2016 to their astrological signs. The ability to have such a unique mix of variables on a cross-section of thousands of Americans remains a strength of this project's choice of data.

Models. An 80/20 split of our data was conducted to create a training set and test set. The goal of the project was to test several exploratory models, including the use of machine learning techniques, in order to derive a final parsimonious generalized linear model (GLM) that could accurately predict anti-immigration attitudes within a U.S. sample. Each model was assembled using R version 3.6.2, beginning with our theory-driven model.

In our theory driven GLM (GLM-1), 10 features were chosen to approximate the predictors of anti-immigration attitudes that has been found in the literature. We included the following demographic-related characteristics of the respondents that have been previously linked to anti-immigration views: race and ethnicity, education level, financial situation, political leaning and political affiliation, region of the country in which they reside, and whether their town of residence is rural or urban. The feature on being a Clinton versus Trump supporter in the 2016 election was included due to the vast academic and media coverage on how Trump support is linked to support for stricter anti-immigration policies. As a proxy for authoritarianism, we included the GSS variable that asked respondents' preference for using spanking as a means to discipline their child. Lastly, to represent social dominance orientation (SDO), we used two GSS variables—one that was on whether the respondent believed African Americans were more disadvantaged in American society due to an inborn inability to learn, and a second item on how hardworking the respondent believed Whites in America to be. Both of these items, according to the project's researchers, appeared to meet the central premises of the SDO construct.

We simulated four models using different supervised learning techniques that were both tuned, then evaluated through the training function provided by the *caret* package in R (Kuhn, 2008). Integrating machine learning methods into this project was a vital means of exploring many features at a large scale at once, particularly within complex data such as ours that included mostly categorical variables. Machine learning algorithms are capable of more precisely assessing the non-parametric relationships found between many features, providing more flexibility than running a GLM on a select number of factors that explores only linear trends. Additionally, big data typically pools together an enormous volume of information per observation, which can help us to identify trends, but machine learning methods allow us to use big data to accurately predict specific outcomes specifically in the test data. For our classification objectives, we used the ensemble learning techniques of random forest and eXtreme gradient boosting. For the random forest model, the *ranger* package was used via R (Wright & Ziegler, 2015) and the *xgboost* package was used for eXtreme gradient boosting (Chen & Guestrin, 2016).

Our first machine learning model was a random forest model (RF-1), a method which has been shown to deliver strong predictive performance results in many classification research problems, including many disciplines in the social sciences. The model is built through blending together an ensemble, or "forest," of decision trees. In order to deal with the variance problem that is found with one single decision tree, random forest uses a technique called "bagging" to randomly sample several subsets of the training data before arriving at a result from aggregating together its trees. For the RFM, an internal evaluation method using 5-fold cross-validation was applied with the split rule to minimize the Gini impurity. *Mtry* was set to the square root and log of total feature numbers. Alternatively, we also ran three different models (i.e., XGB-1, XGB-2, and XGB-3) that used random forest eXtreme gradient boosting ("XGBoost"). This is another ensemble learning method that adds weights to predictions with high error in its resampling of the training data. In comparison, random forest uses a random re-sampling method. XGBoost's "gradient" element further reduced error in subsequent iterations of its algorithm by incorporating a loss function, ensuring the final model is not over-fitting the data. For all our tree-based models, the *train* function in *caret* trained over several parameter options with optimal parameters selected through the metric of the highest receiver operating characteristic (ROC) estimates. For all our XGB models, we allowed the algorithm to randomly search for its own tuning parameters. However, XGB-1 and XGB-2 used 2-fold cross-validation due to computational time constraints, but we doubled the tuning length for XGB-2. In XGB-3, we used 5-fold cross-validation and a tune length of three as the middle ground between XGB-1 and XGB-2. In addition, we ran another logistic regression (GLM-2), which was implemented using boosting, that incorporated the full set of GSS features in our data. Estimation of the variable importance across all our tree-based models was done through the parameter of permutation accuracy. Permutation accuracy involves randomly removing an input feature, then re-running another model to compute performance change—if the performance decreases in this iteration, this removed feature is considered of great explanatory importance. This feature importance information was derived through the *iml* package (Molnar et al., 2018). Through the comparison of all these models, we hoped to detect the most important features that could inform our final GLM that can accurately predict anti-immigration attitudes. The top five features from the tree-based models were selected for our final MLM-informed GLM (GLM-3).

Results

By combining the top five features of variable importance found among our tree-based models, we included the following GSS inputs for GLM-3: whether or not the respondents' parents was born in the U.S. (*parborn*), whether or not the respondent was born in the U.S. (*born*), whether or not the respondent believed African-Americans were socio-economically disadvantaged in the U.S. due to lack of willpower (*racdif4*), how in favor the respondent would be if a black person married a close relative (*marblk*), the respondent's views on the U.S. Supreme Court decision that states should not require public schools to have their students read the Bible (*prayer*), the respondent's views on capital punishment for murder (*cappun*), and if the respondent agree or disagreed that courts were not being harsh enough on crime (*courts*). Table 1 provides a summary of the top ranked features in models RF-1, XGB-1, XGB-2, and XGB-3.

Rank	RF-1	XGB-1	XGB-2	XGB-3
1	parborn	racdif4	marblk	parborn
2	racdif4	parborn	parborn	racdif4
3	marblk	marblk	racdif4	prayer
4	born	courts	prayer	marblk
5	prayer	prayer	cappun	cappun

Table 1. Top five features per tree-based model.

The reason our final model is ultimately the GLM one, instead of any of the tree-based models, is due to its easier interpretability. It is far more translatable to provide parting information on predictors of anti-immigration views to organizations and the public that may seek straightforward recommendations from our findings. On the other hand, a random forest model's multitude of splits and combinations are difficult to exact in terms of easily actionable guidance.

The results of our GLM-3 showed several significant predictors of anti-immigration views (see Table 2). Anti-immigration attitudes were significantly predicted by participants who were any less than strongly in favor of a close relative marrying a black person, who disapproved of not requiring the Bible as required reading in public schools, and who felt courts needed to be harsher with criminals. Views against immigrants were less significantly probable when participants reported their mother was their only parent born in the U.S., when both of their parents were not born in the U.S., they themselves were not born in the U.S., when the respondent did not believe African-Americans lacked willpower, and when they opposed the death penalty for a murder.

Feature	Variable Description	B	SE	z	p
(intercept)		-0.43	0.09	-4.55	< 0.001***
Parborn -X1	Mother only born in U.S.	-0.32	0.14	-2.34	0.02*
Parborn -X2	Father only born in U.S.	-0.22	0.15	-1.50	0.14
Parborn -X3	Mom born in U.S., father-DK	-0.34	0.48	-0.72	0.47
Parborn -X4	Mom not born in U.S., father-DK	-1.97	1.10	-1.79	0.07
Parborn -X5	Dad born in U.S., mother-DK	-0.19	1.07	-0.18	0.86
Parborn -X6	Dad not born in U.S., mother-DK	0.94	1.24	0.75	0.45
Parborn -X7	DK for both	-0.97	1.15	-0.84	0.40
Parborn -X8	Neither parent born in U.S.	-0.86	0.11	-7.79	< 0.001***

Parborn -DK	No Answer/ DK	0.01	0.56	0.02	0.98
racdif4 -X2	African-American SES differences NOT due to lack of willpower	-0.51	0.05	-10.41	< 0.001***
racdif4 -DK	No Answer/ DK	-0.26	0.12	-2.15	0.03*
Marblk -X2	Favor marriage	0.46	0.08	5.71	< 0.001***
Marblk -X3	Neither favor nor oppose	0.23	0.06	3.84	< 0.001***
Marblk -X4	Oppose	0.83	0.09	9.06	< 0.001***
Marblk -X5	Strongly oppose	1.42	0.11	13.13	< 0.001***
Marblk -DK	No Answer/ DK	0.52	0.34	1.53	0.13
Prayer -X2	Disapprove of no Bible reading in public schools	0.48	0.05	9.78	< 0.001***
Prayer -DK	No Answer/ DK	0.21	0.12	1.73	0.08
Cappun -X2	Oppose capital punishment	-0.45	0.05	-8.61	< 0.001***
Cappun -DK	No Answer/ DK	-0.27	0.10	-2.60	< 0.01**
Born -X2	Not born in the U.S.	-0.54	0.13	-4.24	< 0.001***
Born -DK	No Answer/ DK	-10.87	137.66	-0.10	0.94
Courts -X2	Courts not harsh enough dealing with criminals	0.45	0.07	6.16	< 0.001***
Courts -X3	Courts are about right in dealing with criminals	0.12	0.09	1.43	0.15
Courts -DK	No Answer/ DK	0.10	0.11	0.92	0.36

Table 2. Results of GLM-3.

We next tested how well all our models performed, i.e., correctly predicted anti-immigration views in our test set. A primary focus here was to see if our final model, the MLM-informed GLM, performed better than our first theory-driven GLM. In total, seven models were compared: the theory-based GLM (GLM-1), random forest model (RF-1), XGB-1, XGB-2, XGB-3, a GLM with all variables (GLM-2), and finally, our MLM-based GLM using select features (GLM-3). As a performance metric, ROC_AUC curves were used, derived through the *pROC* package (Robin et al., 2011). Table 3 presents all the area under the curve (AUC) results. When comparing all seven models, it is clear that XGB-3 with the greatest number of tuning parameters had the highest AUC of 0.75. The poorest performing model was our first theory based GLM-1, followed by the generalized linear model with all the GSS features retained in our data that incorporated boosting (GLM-2). However, there was no huge AUC difference found among our models with scores only varying by 0.01 increments. A markedly sizable difference was found between our theory-based GLM-1 (*AUC* = 0.69) and our final MLM-based GLM-3 (*AUC* = 0.72).

Additionally, it is worth noting that the GLM-3 shares only a 0.03-difference from our best performing model XGB-3.

Conclusion and Discussion

Many previous studies have examined attitudes towards immigrants in relation to the increasing number of immigrants. As Western Europe and the U.S. have historically been popular destinations for immigrants, our review of the literature covered many natives' characteristics and preferences as it influences their perspective on immigrants. This included their country of origin, their socio-economic status, and even their language skill levels. Our study further investigated the individual characteristics that predict anti-immigrant attitudes among Americans. The characteristics highlighted in previous studies included many domains including people's values, beliefs, education levels, political leanings, their social dominance orientation (SDO), authoritarianism, their geographic location, etc. According to Küpper, Wolf and Zick (2010), individuals with a lower income were far more likely to hold a higher SDO, and they were more likely to have strong negative beliefs about immigrants. Individuals with a higher SDO tended to hold more negative attitudes towards immigrants overall. Another study also found that authoritarianism is a key predictor in individuals' anti-immigrant attitudes (Oyamot Jr., et. al., 2012). Oyamot and colleagues found that authoritarianism was negatively associated with positive sentiments towards immigrants under a negative norm condition—where participants were told that other Americans generally had negative opinions towards immigrants, versus a positive norm condition where participants were told that Americans

Model Name	AUC
GLM-1	0.6907
GLM-3	0.7156
GLM-2	0.7323
RF-1	0.7446
XGB-1	0.7487
XGB-2	0.7498
XGB-3	0.7538

Table 3. AUC results.

generally had positive opinions towards immigrants. Under a positive norm manipulation, intolerance among authoritarian participants was reduced. There also exists a correlation between SDO and right-wing authoritarianism, which is also able to predict anti-immigrant attitudes. On the one hand, people with lower occupational status and unemployment were more likely to hold anti-immigrant attitudes (Rink, Phalet, & Swyngedouw, 2009). On the other hand, the study found natives' negative attitudes towards immigrants increased with increasing and visible diversity of immigrants (Singh & Stojek, 2015). At the same time, anti-immigrant attitudes decreased as individuals' education levels increased (Rink et. al., 2009). In investigating American attitudes towards immigrants, the authors of a recent study found variance in these views based on individuals' different geographic regions, i.e., urban versus rural areas (Garcia & Davidson, 2013). For both urban and rural areas, the study found that race, education level, and beliefs about immigration were also significant predictors of anti-immigrant attitudes.

In the present study, we tested several models to examine the most significant variables predicting anti-immigrant attitudes. First, we built a theory-driven model, which includes variables we mentioned in the literature review. We tested individuals' race and ethnicity, education level, political party affiliation, geographical region, their beliefs about those who are of a disadvantaged status and so on. We then developed a second model that included all individual characteristics, comprising of more than a hundred variables. We also utilized several machine learning techniques in our follow-up models, including random forest. In our next step, we chose the top 20 significant variables in all the tree-based models we built. After analyzing this data, we found nine variables (personal characteristics) of significance to anti-immigrant attitudes. People whose mother (only) was born in the U.S. were 32.3% more likely to hold a positive attitude towards immigrants. When neither of one's parents were born in the U.S., he or she was 86.5% more likely to hold a positive attitude towards immigrants. People born outside the U.S. were 54.1% more likely to hold a positive attitude towards immigrants. Those who believed that African Americans were of a disadvantaged status not due to a lack of willpower (e.g., representing a more humanitarian view), they were 50.8% more likely to hold a positive attitude towards immigrants. People who supported a looser attitude on capital punishment for murder was 45.4% more likely to not hold a negative attitude towards immigrants. Individuals who strongly opposed interracial marriage (social dominance) tended to hold a negative/anti-immigrant attitude. People who approved of requiring the reading of the Bible in public school were 47.5% more likely to hold a negative attitude towards immigrants. Finally, people who believed that "courts dealing with criminals are not harsh enough" were 44.9% more likely to hold negative attitudes towards immigrants.

We attempted to build a model connected to the previous research and theories, but the predictors in our final model did not match the same predictors conceived of in our initial model. For instance, we did not find region to be a significantly important predictor of anti-immigrant attitudes in our tree-based models. Political leaning in our models was not a top predictor nor was diversity of ethnic groups in one's neighborhood. Nevertheless, we found that the final variables that we did use remained consistent with the theory found in our literature review.

In a real-world situation where researchers need to understand anti-immigration attitudes expediently based on a briefer survey, the GSS's 600-plus items is an impractical number of variables to test. Having access to all these variables, however, allowed us to examine many possible influences of anti-immigrant attitudes to arrive confidently at the most optimal variables. Many of the variables selected in our final model are subtle and indirectly related to anti-immigration views. Thus, this finding indicates that direct questions on immigration views could actually lead to dishonest or less valid responses. Differences between responses and people's true attitudes may be contaminated from social desirability bias if survey respondents are unwilling to admit their actual preferences. Through our study, we may be able to solve this issue by asking one of our top explanatory variables instead that are more indirectly connected to anti-immigration views. Our machine learning methods was able to pinpoint several of these possible items, among numerous possible predictors, that can best explain anti-immigration views over researcher-selected, seemingly face-valid items.

Thus far, we have been able to determine individuals' characteristics such as birthplace and parental background, attitudes towards African-Americans, attitudes towards marriage with a black person, individuals' attitudes towards the death penalty for murder, their attitudes of requiring Bible in public schools, and their attitudes towards punishment towards criminals as significant predictors in our models and data analysis. However, we did not reveal the causal effects of the negative attitudes; that is, our research does not determine any links of causation between the characteristics and attitudes, which is one of our many limitations. We also did not consider interviewer characteristics in this project, but perhaps in a subsequent iteration, it is possible that we could have found this to be an important factor in people's willingness to provide valid response. Characteristics such as the interviewers' level of experience and their race may have been important determinants that could inform future data collection. The present study, however, wanted to improve upon our knowledge of individual characteristics that predict anti-immigration views, not necessarily optimize on interview techniques. We also ideally would be able to test our model on future test sets such as the 2020 GSS, which has not yet been published, to see if our model performance remains similarly accurate and stable when tested upon future attitudes.

There are many questions that remain, despite our knowledge of these top explanatory variables for anti-immigration views. We may assume that people whose parents were not born in the U.S may have been discriminated against, and thus they are more tolerant to immigrants and teach their children to empathize with minority groups. They may also be less likely to believe that immigrants are a threat to their cultural environment since they are not totally natives themselves. Those who believed that African Americans were in a disadvantaged position not due to willpower may hold a positive attitude towards the group, rather than just attribute their disadvantaged status to lack of willpower. Similarly, they might also hold a positive opinion toward immigrants and the inevitable challenges they experience. But these are only assumptions, and there were certainly some variables that were even more difficult to connect with anti-immigrant attitudes. The connection of desiring Bibles to be read in public schools, but religion not rising as a top predictor, makes these variables' connection to anti-immigration views less clear beyond a tenuous connection to religiosity.

The increasing number of immigrants around the world has led to severe conflicts between Americans and immigrants, and these conflicts are often dynamic and complex in ways that survey statistics alone cannot portray. Nevertheless, in this study we attempted to determine which individual characteristics most predict anti-immigrant attitudes. Although we were not able to personally match factors to the literature, per say, we still confirmed many of their theoretical basis. Meanwhile, unexpected variables determined important in our study to affect individuals' attitudes are noteworthy to future researchers in terms of must-have survey items to include in assessments of immigrant views. Lastly, we found several causal effects that mitigate anti-immigrant attitudes and therefore may inform policies and efforts to reduce the negative attitudes directed toward immigrants, which has been a fundamental goal of this research.

References

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). ACM.

Gallego, A., & Pardos-Prado, S. (2013). The Big Five personality traits and attitudes towards immigrants. *Journal of Ethnic and Migration Studies*, *40*(1), 79-99.

Garcia, C. E., & Davidson, T. (2013). Are rural people more anti-immigrant than urban people? A comparison of attitudes toward immigration in the US. *Journal of Rural Social Sciences*, 80.

Gorodzeisky, A., & Semyonov, M. (2016). Not only competitive threat but also racial prejudice: Sources of anti-immigrant attitudes in European societies. *International Journal of Public Opinion Research*, *28*(3), 331-354.

Johannes Karreth, Shane P. Singh & Szymon M. Stojek (2015) Explaining Attitudes toward Immigration: The Role of Regional Context and Individual Predispositions, *West European Politics*, 38:6, 1174-1202, DOI: 10.1080/01402382.2015.1022039

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28(5), 1 - 26. doi:http://dx.doi.org/10.18637/jss.v028.i05

Küpper, B., Wolf, C., & Zick, A. (2010). Social status and anti-immigrant attitudes in Europe: An examination from the perspective of social dominance theory. *International Journal of Conflict and Violence (IJCV)*, *4*(2), 205-219.

Molnar C, Bischl B, Casalicchio G (2018). "iml: An R package for Interpretable Machine Learning." JOSS, 3(26), 786. doi: 10.21105/joss.00786, http://joss.theoj.org/papers/10.21105/joss.00786.

Oyamot Jr, C. M., Fisher, E. L., Deason, G., & Borgida, E. (2012). Attitudes toward immigrants: The interactive role of the authoritarian predisposition, social norms, and humanitarian values. *Journal of Experimental Social Psychology*, *48*(1), 97-105.

Rink, N., Phalet, K., & Swyngedouw, M. (2009). The effects of immigrant population size, unemployment, and individual characteristics on voting for the Vlaams Blok in Flanders 1991–1999. *European sociological review*, *25*(4), 411-424.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics*, *12*(1), 77.

Smith, T. W., Davern, M., Freese, J., & Morgan, S. L. (2019). General social surveys, 1972–2018. *Retrieved October*, *20*, 2019.

Thomsen, L., Green, E. G., & Sidanius, J. (2008). We will hunt them down: How social dominance orientation and right-wing authoritarianism fuel ethnic persecution of immigrants in fundamentally different ways. *Journal of Experimental Social Psychology*, *44*(6), 1455-1464.

Wright, M. N., & Ziegler, A. (2015). ranger: A fast implementation of random forests for high dimensional data in C++ and R. *arXiv preprint arXiv:1508.04409*.