JOURNAL FOR THE SCIENTIFIC STUDY OF RELIGION

RELTRAD2: Refining the State of the Art of Religious Classification by Reconsidering the Categorization of Nondenominational Respondents

JOSH GAGHAN
Duke Global Health Institute
Duke University

DAVID EAGLE
Department of Sociology, Duke Global Health Institute
Duke University

RELTRAD is a major religious taxonomy used by a large number of researchers. Although criticisms have been raised about its utility, improving the algorithm to capture contemporary religious dynamics is important given its widespread use. The present RELTRAD taxonomy classifies more religiously active nondenominational respondents as Conservative Protestants and codes the remainder as missing data. A growing number of Americans indicate they are either nondenominational or only Christian or Protestant, which means using RELTRAD in its existing form codes a nonrandom and increasingly large number of respondents with a missing value for religious affiliation (growing from 2 percent to 5 percent of the US General Social Survey (GSS) sample between 2000 and 2018). Using a machine learning algorithm to predict the likely religious tradition of nondenominational respondents, we demonstrate the shortcomings of this approach and introduce a new coding scheme, RELTRAD2, which classifies nondenominational respondents who report a Black racial identity as Black Protestant, non-Black respondents who never attend religious services as Mainline Protestant, and the remainder as Conservative Protestant. Code to derive RELTRAD2 from the GSS is provided.

Keywords: *religion, religious classification, RELTRAD.*

INTRODUCTION

Social scientists have long recognized that religious affiliation is associated with a myriad of consequential outcomes. The salience of religion, coupled with the sheer number of denominations and religious traditions in the United States, has produced considerable demand for grouping people into analytically useful religious categories. Classification is always a balancing act. Religious taxonomies must simultaneously capture the distinctiveness and broad similarities of various religious traditions. Since its inception in 2000, RELTRAD has received over 1400 citations and become the dominant way to classify respondents' religious affiliation in social scientific research in the United States (Steensland et al. 2000).[1] Based on their self-reported religion or denomination, RELTRAD classifies respondents into one of seven religious affiliations: Black Protestant, Conservative Protestant, Mainline Protestant, Roman Catholic, Jewish, unaffiliated, and "other".

Correspondence should be addressed to Josh Gaghan, Duke Global Health Institute, Duke University, 310 Trent Drive, Durham, NC 27710, USA. E-mail: joshua.gaghan@duke.edu

[1]Search performed on Web of Science on December 1, 2023.

Despite its prominence, RELTRAD has not been immune to criticism. For example, researchers have debated the merits and drawbacks of including Black Protestantism as a separate category. On the one hand, aggregating all of the Black Protestant traditions into a single group conceals the religious diversity of Black religious groups in the United States (Shelton and Cobb 2017). On the other hand, this approach acknowledges how a legacy of discrimination and inequality in American Protestantism has produced a distinctive Black Protestant tradition (Shelton 2018). Others have pointed out that many churches within Mainline Protestant denominations have a strong Conservative Protestant character (McKinney and Finke 2002). Finally, some researchers have critiqued RELTRAD for not leaving any middle ground between Conservative and Mainline Protestantism; for conflating religious affiliation and practice; and for its exclusion of a nonrandom sample of respondents (Burge and Djupe 2021; Hackett et al. 2018; Lehman and Sherkat 2018).

Nevertheless, RELTRAD has remained a dominant approach despite the criticism levelled against it. Analyses comparing RELTRAD to other classification methods have suggested that its use is defensible or even preferable to the alternatives (Hackett et al. 2018; Shelton 2018). While developing a novel religious taxonomy that consistently outperforms RELTRAD and is widely adopted by researchers may be feasible, this is not the focus of this project. Instead, we recognize many researchers will continue to use RELTRAD for the foreseeable future. Our goal is to offer a straightforward improvement to RELTRAD that researchers can easily implement.

RELTRAD, like all religious taxonomies, struggles with classifying respondents who state they are either nondenominational or only Christian or Protestant. These respondents are likely heterogeneous. Steensland et al. (2000) argued the group of people who identify only as Christian or Protestant tend to be more nominally attached to their religion, whereas those who state they are "nondenominational" are a subset of Conservative Protestantism (below, we will refer to members of both of these groups as *nondenoms* or nondenominational respondents). Because the US General Social Survey (GSS, Davern et al. 2024) places nondenominational respondents into the same religious category as those who are only Christian or Protestant, Steensland and colleagues recommended differentiating them using an attendance cutoff: respondents who reported attending religious services at least monthly should be classified as Conservative Protestant and those reporting less than monthly attendance as missing data on this variable (Steensland, Woodberry, and Park 2018; Steensland et al. 2000).

In recent years, the Steensland et al. classification method of *nondenoms* has generated increased criticism. The percentage of respondents in the United States affiliated with no denomination in the GSS has nearly doubled since the inception of RELTRAD, growing from 6 percent in 1998 to 11 percent in 2018. Applying the attendance cutoff to the 2018 General Social Survey leaves 5 percent of all respondents unclassified, resulting in a considerable amount of nonrandom missing data (Burge and Djupe 2021; Hackett et al. 2018). In addition, the assumption that church-attending *nondenoms* are Conservative Protestants is questionable. Some nondenominational megachurches are part of the Black Protestant tradition (Thumma and Travis 2007), and some nondenominational respondents may simply be unaware of their denominational affiliation (Lehman and Sherkat 2018). There is no compelling reason to assume that Conservative Protestantism is the only possible category in which *nondenoms* fit.

These issues have spawned a range of alternative suggestions. The first comes from the authors of RELTRAD, who recognize the problem of excluding the growing population of low-attending *nondenoms*. They suggest researchers consider placing nondenominational respondents—along with others who are denominationally ambiguous—in a "nominal religion" category if they report not believing in God or life after death and rarely, if ever, praying or attending religious services (Woodberry et al. 2012). Despite leaving fewer respondents unclassified, this technique presents a significant problem. Only less religious *nondenoms* are eligible to be classified as nominals. It has long been known that people who affiliate with a specific religious tradition may also rarely attend religious services or pray and do not believe in God or an afterlife (Vernon 1968). Although we could apply the nominal religion category to affiliates of all religious traditions, this would conflate religious affiliation with beliefs and practices, which the RELTRAD taxonomy seeks to avoid.

More recently, Burge and Djupe (2021) explored a range of options for dealing with the issue of unclassified respondents and recommended placing all nondenominational respondents into a separate category. This assumes that nondenominationalism is a cohesive tradition distinct from Black, Conservative, and Mainline Protestantism. The benefits of this approach are that it minimizes the share of respondents who are left unclassified and illustrates the rise of nondenominationalism. As they considered alternative methods of classifying *nondenoms*, they dismissed the possibility of using religious service attendance or any other variable to distinguish individuals in this group, citing the awkwardness of differentiating the religious traditions of two people in the same church based on their attendance. However, researchers have suggested that, instead of forming a coherent tradition, nondenominationalism is instead an amalgamation of individuals with different religious affiliations (Dougherty, Johnson, and Polson 2007; Lehman and Sherkat 2018; Steensland et al. 2000; Woodberry et al. 2012). If this is true, we should judiciously place *nondenoms* into their respective traditions instead of assigning them to their own group. Although the idea of differently classifying two individuals in the same pew by their church attendance or some other variable may feel awkward, these differentiations may be necessary when applied to the wider population of *nondenoms* to accurately account for the heterogeneity of this group. If nondenominational respondents as a group are truly a mixture of individuals with different religious affiliations, any system will produce some degree of misclassification. Classifying *nondenoms* separately does not obviate this issue. Instead, we should accept that some level of misclassification is inevitable and consider which method introduces the least bias. Otherwise, separately classifying nondenominational respondents will unnecessarily alter the size and composition of other religious groups and potentially obscure our ability to track trends in religious affiliation as the number of *nondenoms* in the general population grows.

Finally, others have noted that *nondenoms* do not fit neatly within Conservative Protestantism (Shelton 2018). Dougherty, Johnston, and Polson (2007) adjust for this by differentiating Black and non-Black *nondenoms* while eliminating the attendance cutoff by placing the former in Black Protestantism and the latter in Conservative Protestantism. Unlike the approaches described earlier, this method uniquely deals with the heterogeneity of nondenominational respondents by classifying them into separate religious categories.

To date, the proposed modifications of these alternate classification schemes suffer from two major shortcomings. First, while excluding a nonrandom proportion of nondenominational respondents (as RELTRAD does) may produce bias (for example, by making Conservative Protestants appear more religious as a group), categorizing this heterogenous group is likely to result in substantial misclassification. None of the proposed alternative approaches to handling *nondenoms* delineate the potential bias these schemes may introduce.

Second, researchers have given little attention to developing empirically driven methods of separating *nondenoms* into their likely religious traditions. While Burge and Djupe (2021) considered and decided against classifying low-attending *nondenoms* as religious nones or Mainline Protestants and Dougherty, Johnson, and Polson (2007) divided them into Conservative and Black Protestants, researchers have only used respondents' reported race and attendance at religious services to classify them into religious families. We argue that a better approach is using a broader range of variables to consider the classification of *nondenoms*. Although this method entails using a range of variables distinct from religious affiliation to predict one's religious tradition, it is not conflating affiliation with other components of individual's religious identity (e.g., beliefs and practices). Instead, we recognize that religious affiliation correlates in predictable ways with other aspects of identity. This is a well-established practice. Researchers frequently measure the performance of religious classification methods by comparing how well they predict the distinct beliefs and practices of adherents of different religious traditions (Shelton 2018; Steensland et al. 2000). The difference here is that we are leveraging this insight to predict the religious classification of *nondenoms*, instead of the reverse.

In this study, we outline a simple, empirically validated method of classifying *nondenoms*. Our goal is to recommend a straightforward method that both limits the amount of missing data and best preserves the broader associations between a person's reported religious tradition and other social factors. To do this, we test the following algorithms, which are either already used to categorize nondenominational respondents or are simple extensions thereof:

OG: Steensland et al.'s (2000) original model where *nondenoms* who report attending religious services at least monthly are coded as Conservative Protestant. Everyone else is assigned as missing a RELTRAD value.
CP: Every *nondenom* is classified as Conservative Protestant. This mirrors the original model, but eliminates the attendance cutoff.
CP-BP: Dougherty, Johnston, and Polson's (2007) method of coding Black *nondenoms* as Black Protestant and everyone else (including those who do not report their race) as Conservative Protestant.
BP-CP-MP: Black *nondenoms* are coded as Black Protestant. Non-Black *nondenoms* (including those who do not report their race) who never attend religious services are coded as Mainline Protestant. All other respondents are coded as Conservative Protestant (including those who do not report their race or frequency of attendance). The basis for this coding scheme comes from our analysis of the church's nondenominational respondents attend described later in this article.

We measure the performance of each of these methods by using a machine-learning classification algorithm to determine the most likely religious affiliation of *nondenoms*. Although imputing the religious classification of nondenominational respondents using a machine learning algorithm may be preferable to using one of the simpler methods described above, we recognize that few researchers are likely to do this. For this reason, we test each of the simple coding schemes by relying on the predicted religious tradition of *nondenoms* to examine how each may bias both the size and composition of each religious tradition. Although our solution to the problem of classifying *nondenoms* is not perfect, it offers the best trade-off of simplicity and accuracy.

Data And Methods

We explore the religious composition of nondenominational respondents using a pair of methods and datasets to answer a basic question: Among participants with a clearly defined religious affiliation, whom do *nondenoms* most closely resemble? To answer this, we used cross-sectional data from the US GSS, a nationally representative survey of adults in the United States that NORC at the University of Chicago has conducted since 1972 (Davern et al. 2024). We limited our sample to responses from 1972 to 2018 since changes to the survey methodology in 2021 may have impacted data quality (Schnabel, Bock, and Hout 2022). We constructed the RELTRAD variable according to the method outlined by Steensland et al. (2000), except we retain a separate category for *nondenoms*. We then filtered the data to only include people who are either *nondenoms* or in religious traditions in which we could plausibly classify these respondents: Black, Conservative, and Mainline Protestantism.

Even though some nondenominational respondents may closely resemble the religiously unaffiliated, we adopt a straightforward definition of a religious none. Namely, they are individuals who, when asked, indicate they are not affiliated with a religion (Vernon 1968). Because all nondenominational respondents were given the option to identify as unaffiliated but explicitly chose not to, we should not, by definition, classify them as religious nones. In reality, the boundary between identifying as having no religious affiliation versus having one is fuzzy (Voas 2009). We do not question the respondent's decision to identify with a religious tradition. Although some portion of nondenominational respondents may sit on the threshold between being religious affiliated and unaffiliated and closely resemble nones (Hout 2017; Lim, MacGregor, and Putnam 2010), we take their stated religious affiliation at face value (insofar as they are explicitly not identifying as unaffiliated) and place them into the most plausible religious tradition.

To examine which of the Protestant groups the nondenominational respondents most closely resemble, we used Extreme Gradient Boosting (XGBoost), a machine learning algorithm widely used for classification (Chen and Guestrin 2016). We chose XGBoost over alternative classification algorithms because of its ability to maximize out-of-sample predictive accuracy using a broad set of independent variables while mitigating overfitting (Bentéjac, Csörgő, and Martínez-Muñoz 2021). XGBoost is a supervised machine learning algorithm used to classify observations from a dataset into their most likely category. Using a specified set of predictors, the algorithm builds a series of weak-learning decision trees to determine the likely classification of an observation (Friedman 2002). Since its development in 2016, XGBoost has been applied to a wide range of scientific problems, from diagnosing Chronic Kidney Disease (Ogunleye and Wang 2020) and epilepsy (Torlay et al. 2017) to detecting cyberattacks (Chen et al. 2018). For our analysis, we began by using the XGBoost model using the XGboost package in R to predict RELTRAD based on all the other variables in the GSS, excluding variables that are definitionally connected to any of the RELTRAD categories (i.e., RELIG, DENOM, OTHER, FUND, RELIGID, and RELID1) (Chen et al. 2023).[2] To train the model (i.e., fit the model to the data), we dropped all *nondenoms* and randomly sampled 80 percent of the remaining cases to use as a training set to develop model parameters. We used the other 20 percent as a test set to check the accuracy of model predictions. To avoid under-and overfitting, we used Recursive Feature Elimination to narrow the dataset by testing what number of features would optimize the model (Guyon et al. 2002). This entailed, after fitting the XGBoost algorithm to the dataset containing all of the variables in the GSS, analyzing the variables the model used by the level of influence they had on the predictions (i.e., the improvement in accuracy a feature provides weighted by the number of observations in the nodes where the feature is used). We then systematically eliminated the least important variables to determine that training the algorithm on the 121 most important variables generated the most accurate model. Finally, we applied the algorithm to the *nondenom* respondents to predict their RELTRAD classification (i.e., whether respondents were Black Protestants, Conservative Protestants, or Mainline Protestants). To be clear, this model is not predicting a respondent's "true" religious affiliation. Rather, it is estimating what category (i.e., religious affiliation) of respondents an unknown observation most closely matches based on how other social characteristics tend to correlate with religious affiliation. The predicted religious tradition of these respondents serves as a proxy for their RELTRAD classification. The goal of this procedure is not to determine the "true" religious affiliation of a given individual, but rather to model the religious characteristics of a population, given how religion and other social categories correlate.

To develop our final candidate classification algorithm, we analyzed the religious traditions of the congregations nondenominational respondents attend. To do this, we used linked data from the National Congregations Study (NCS) and the General Social Survey (Chaves et al. 2020; Chaves et al. 2020). The NCS generates a nationally representative sample of religious congregations by asking GSS respondents who report attending religious services in 1998, 2006, 2012, and 2018 to name the congregation they attend. The NCS-GSS data link respondents' individual GSS responses with the responses of their religious congregation. The NCS asks a key informant

[2]As each of these variables include information that directly corresponds to at least one RELTRAD category (e.g., all respondents who are affiliated with the Southern Baptist Church in DENOM would be classified as conservative Protestant), including these variables in the analysis would train the algorithm to make perfect predictions of denominational respondents using these variables while providing limited, if any, information on how to classify nondenominational respondents.

from the named religious congregation (generally the senior staff leader) to identify the denominational tradition of their congregation. NCS researchers then used this information to place the congregation into one of five religious traditions that correspond with the RELTRAD categories.

We used the NCS because it identifies the religious tradition of the congregations nondenominational respondents attend. Although an individual's religious tradition does not necessarily correspond with the affiliation of their congregation, they are strongly correlated, enabling us to use the church tradition of nondenominational respondents as a proxy for their religious classification. To determine how to predict the religious classification of *nondenoms* using a simple set of variables, we created a multinomial logit model with the religious tradition of the congregation nondenominational respondents report attending as the dependent variable. The independent variables for this model include a continuous variable for the frequency of attending religious services, a binary variable indicating whether someone identifies as Black, and an interaction between these two variables. We selected these variables because they are already used to construct RELTRAD. Finally, we analyzed the model's predictions to assess the probability that nondenominational respondents primarily attend Black, Conservative, or Mainline Protestant congregations. We did this to determine how to create a simple classification scheme that most accurately reflects the religious tradition of the congregations nondenoms attend by race and their frequency of attendance.

The remainder of our analysis focuses on evaluating these methods for categorizing nondenominational respondents. We begin by comparing the assignments of the simple classification schemes to our XGBoost predictions. The methods that match our predictions made using GSS data at higher rates perform better according to this metric. Unfortunately, match rates alone cannot tell us which alternative method will be the most accurate in modeling the population. Different mismatches produce varying levels and types of bias. For example, one method may tend to misclassify Black Protestants who closely resemble Conservative Protestants as Conservative Protestant while another might erroneously categorize liberal Mainline Protestants as Conservative Protestant. In this case, it is possible that the latter method would produce more bias even if it boasts a higher match rate. This is because a smaller group of respondents who are wrongly placed into a group they do not resemble can introduce more bias than larger group of respondents who are erroneously placed into a group they closely resemble. Moreover, a method may generate a lower match rate, but predict the "correct" proportion of *nondenoms* in each religious group. For example, a system that classifies a number of Conservative Protestants as Mainline Protestant and an equal number as the reverse will maintain the accurate size of each religious group despite having yielding mismatches. For this reason, we cannot use match rates alone to measure bias. Methods that tend to incorrectly place individuals into groups they nearly resemble and maintain an accurate proportion of *nondenoms* in each religious group may accurately represent each RELTRAD classification in the aggregate, even with a comparatively low match rate.

For this reason, we need to consider how researchers use RELTRAD in practice to test the accuracy of each alternative classification method. Researchers use RELTRAD as a dependent variable (Bengtson et al. 2018; Skirbekk, Kaufmann, and Goujon 2010), an independent variable (Baunach 2012; Lim and Putnam 2010; Uecker, Regnerus, and Vaaler 2007), or as a category to capture broad trends in religious affiliation (Chaves 2017; Thiessen and Wilkins-Laflamme 2020). A simple method of classifying *nondenoms* should produce minimal bias when using RELTRAD in any of these ways. This requires that a classification scheme accurately captures the size and composition of each religious tradition, as the former is needed when using RELTRAD to capture broad trends in religious affiliation, the latter when using it as an independent variable, and both when using RELTRAD as a dependent variable. To test each method's bias when using each of these three methods, we can compare their estimates to those our XGBoost predictions produce, since these classifications represent our best approximation of the "true" composition of *nondenoms*. We can consequently assume that the methods that consistently produce estimates closer to those of the XGBoost predictions are less biased. Because each method's bias may differ depending on whether RELTRAD is used as a dependent variable, an independent variable, or as a category to capture broad trends in religious affiliation, we assess the bias of each alternative algorithm when using RELTRAD to examine the size and composition of each religious group.

We measure each classification scheme's ability to measure the size of each religious tradition by comparing the proportion of respondents each method predicts are in each RELTRAD group by year to the XGBoost predictions. We do this by calculating the root-mean-square error (RMSE) between probability of an individual affiliating with each religious group by year and the XGBoost predictions (Chai and Draxler 2014). This metric measures the difference between each algorithm's predictions and the "true" proportion of respondents in each religious group by year with the following formula:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}, \tag{1}$$

where $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are each classification method's predicted probabilities, $y_1, y_2, \ldots, y_n$ are each of the XGBoost predicted probabilities, and n is the total number of observations. A lower RMSE indicates that a classification method more accurately estimates the proportion of individuals in each religious tradition. Finally, we bootstrapped the data with 1000 iterations to calculate the confidence intervals for each of these estimates (Hesterberg 2011). This accounts for the sampling error, but not the error introduced by our XGBoost predictions. Because this algorithm is not perfectly accurate and, even if it were, the predicted religious tradition of nondenominational respondents is not necessarily identical to their actual religious tradition, the uncertainty of our estimates should be higher than confidence intervals indicate.

To measure each method's accuracy when measuring the composition of each RELTRAD group, we calculated the cross-model marginal effect differences of using each alternative algorithm—instead of the XGBoost predictions—to predict a range of religious, social, and political views and identities (Mize, Doan, and Long 2019). This entailed running a series of logit models with the following dependent variables:

Belief in God: This is a binary variable (GOD) indicating whether one believes in the existence of God without doubts.
Biblical Literalism: This is a binary variable (BIBLE) indicating whether one believes the Bible is the inspired word of God to be taken literally.
Born Again: This is a binary variable (REBORN) indicating whether one has had a "born again" experience.
Abortion for Any Reason: This is a binary variable (ABANY) indicating whether a pregnant woman should be able to obtain a legal abortion for any reason.
Same-Sex Relations: This is a binary variable (HOMOSEX) indicating whether one believes sexual relations between two adults of the same sex are not wrong at all.
Democrat: This is a binary variable (PARTYID) indicating whether one is a Democrat (including those who are "independent, close to Democrat").

We fit a series of logit models predicting each of these dependent variables on the basis of a person's religious affiliation (i.e., three binary variables indicating whether a person is Black, Conservative, or Mainline Protestant) measured by each classification technique (e.g., the XGBoost predictions and the OG algorithms). Because we are interested in measuring the degree of bias each method introduces in practice, we included all denominational respondents in these analyses. Although the composition of these denominational respondents does not vary across classification methods, their inclusion shrinks the estimates to reflect the bias each method introduces when analyzing the general population, not just *nondenoms*. We applied weights (WTSSALL) and added the following controls to each model: survey year (as a categorical variable); their family's inflation-adjusted income (REALINC); a continuous variable indicating their highest degree earned (DEGREE); a binary variable indicating whether the respondent's sex is female; a binary variable indicating whether the interview was conducted in the South; a binary variable indicating whether the respondent identifies as Black; and age. We then ran a linear model with the same independent variables to predict the following dependent variable:

Religious Service Attendance: This is a standardized continuous variable (ATTEND) indicating how frequently one attends religious services from "never" to "several times a week".

Our outcome of interest was the cross-model marginal effect differences of using each of the simple classification methods in lieu of the XGBoost predictions. This entailed identifying the effect of affiliating with a given religious group from each model and calculating the coefficient differences between the models using each alternative algorithm and the XGBoost predictions for every dependent variable. For example, in the models using Conservative Protestant affiliation to predict biblical literalism, the cross-model marginal effect differences are the effect of affiliating with Conservative Protestantism using simple classification methods (e.g., the BP-CP algorithm) minus the effect using the XGBoost predictions. Because the coefficients using the XGBoost predictions serve as proxies for the "true" estimates, these cross-model differences enable us to measure the degree and direction of bias each classification scheme introduces. Simply put, classification methods with less bias will tend to produce smaller absolute cross-model marginal effect differences. Finally, we accounted for the cross-model covariances when calculating our confidence intervals of the marginal effects by using the bootstrap method with 1000 iterations (Mize, Doan, and Long 2019), though this method will underestimate the uncertainty of these models to some degree due to its inability to account for the error inherent to incorporating imperfect predictions into the models.

Results

The accuracy of the XGBoost model when applied to the test data was 89 percent (95 percent CI [87.8 percent to 89.3 percent]), meaning the model correctly predicted the religious tradition of 89 percent of the cases in the test dataset ($k = 0.82$). The balanced accuracy for each group, which is the arithmetic mean of specificity and sensitivity, ranged from 0.89 for Conservative and Mainline Protestants to 0.97 for Black Protestants, suggesting that model accurately predicted when survey respondents did and did not affiliate with each religious group. This high level of accuracy suggests that, to the extent that the correct classification of a nondenominational respondent corresponds with their true religious tradition, the additional error the XGBoost predictions introduce into the RMSE estimates and cross-model marginal effect differences is small, especially given how this uncertainty only affects *nondenoms*, which comprise 6 percent of the observations in our data.

Using the parameter estimates from the training model, we calculated the most probable religious family of each respondent in the GSS. These estimates allow us to explore how the composition of *nondenoms* has changed since the 1972. Figure 1 shows the proportion of *nondenoms* the XGBoost model predicts are Black, Conservative, or Mainline Protestant by year. The model indicates that the proportion of *nondenoms* it projects to be Mainline Protestant has dropped considerably since 1972, shrinking from 72 percent of *nondenoms* in 1972 to 24 percent in 2018. By contrast, the percentages of Black and Conservative Protestants have risen considerably during this timeframe. The majority of *nondenoms* have most closely resembled Conservative Protestants since the 1980s. It is important to note that, even though the percentage of *nondenoms* who resemble Mainline Protestants has declined since 1972, the number of nondenominational Mainline Protestants in the general population has remained relatively steady, fluctuating between one

Figure 1
The proportion of nondenominational respondents and 95% confidence intervals by year that the XGBoost algorithm predicts are Black, Conservative, or Mainline Protestant [Color figure can be viewed at wileyonlinelibrary.com]

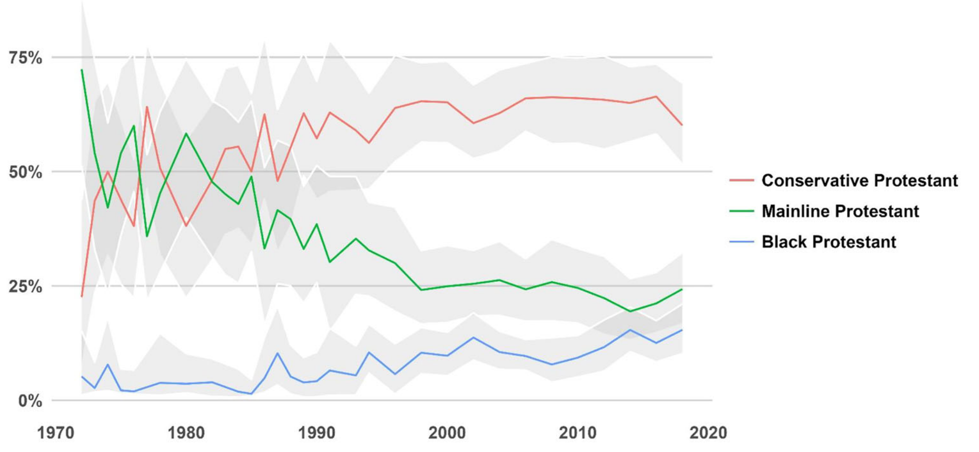
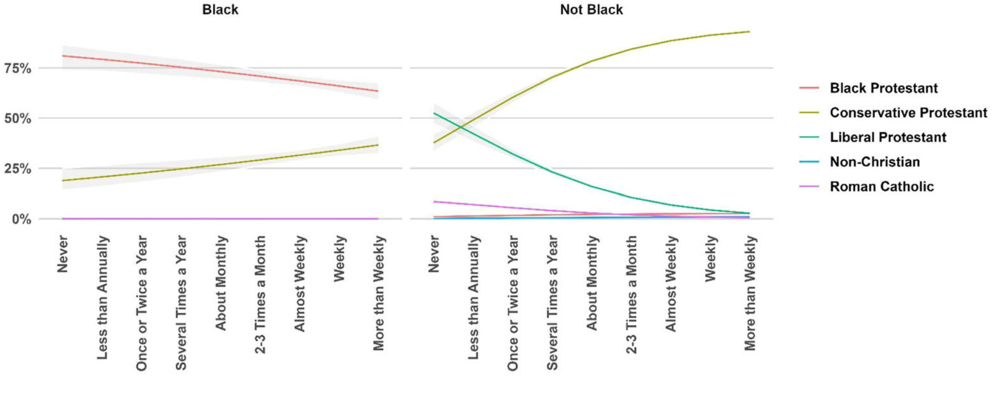


Figure 2
The probability that a nondenominational respondent attends a congregation of a particular religious tradition by race and frequency of religious service attendance [Color figure can be viewed at wileyonlinelibrary.com]



and three percent over the past five decades. Meanwhile, the percentage of Mainline Protestants who are nondenominational has increased from 5 percent in 1972 to 20 percent in 2018. The reason why the share of *nondenoms* resembling Mainline Protestants has fallen is because steeper rise of Black and Conservative Protestants identifying as nondenominational.

The NCS-GSS revealed that nondenominational respondents are distinctly unlikely to attend a nondenominational congregation. Among all of the nondenominational NCS-GSS respondents (which is limited to those who attend religious services at least annually), only 44 percent attend a nondenominational congregation.[3] By contrast, the proportion of Christian respondents in other religious groups who attend denominational congregations in their respective traditions ranges from 72 percent of Conservative Protestants to 91 percent of Roman Catholics.

Figure 2 shows results from our multinomial logistic regression predicting the probability that

[3]Only 6% of NCS-GSS respondents who do not identify as *nondenoms* attend a nondenominational congregation.

Table 1: The rate each classification method matches the XGBoost predictions of the religious traditions of nondenominational respondents. Rates in parentheses count unclassified respondents as unmatched

Classification Method	Match Rate	Match Rate 95% CI	Percent Unclassified
OG	68.1% (34.4%)	66.0%–70.1% (32.0%–34.9%)	51%
CP	61.3%	59.8%–62.8%	0%
BP-CP	69.4%	67.9%–70.8%	0%
BP-CP-MP	65.8%	64.3%–67.3%	0%

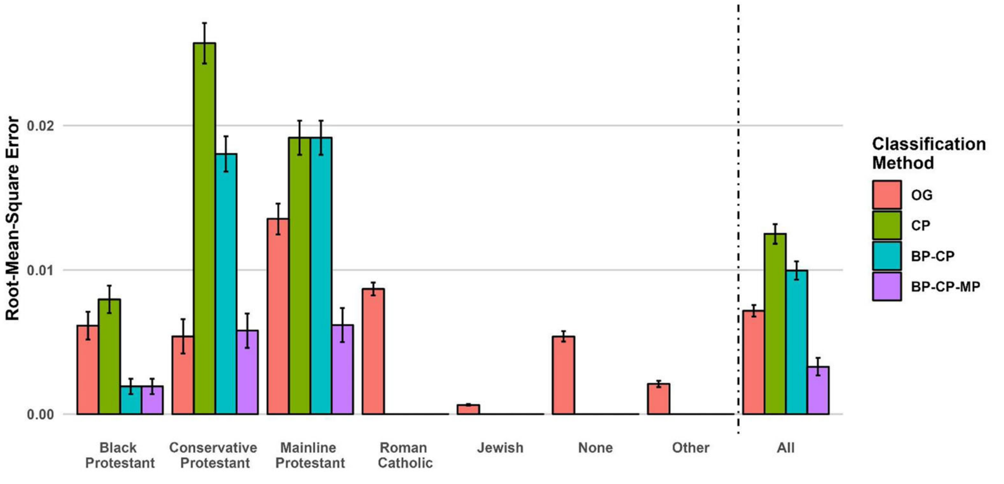
a nondenominational respondent reported attending a congregation of a particular religious tradition. Because the NCS-GSS only includes respondents who report attending religious services at least annually, we extrapolate the trends to predict the religious traditions of the congregations *nondenoms* attend who attend services less than annually. The majority of Black *nondenoms* are likely to attend a Black Protestant church regardless of how often they attend religious services. The frequency of religious service attendance is more salient when we consider the congregations non-Black *nondenoms* attend. Non-Black nondenominational respondents who attend religious services more frequently are more likely to attend a Conservative Protestant church. As this group attends religious service less often, it becomes increasingly likely that the congregation they attend is Mainline Protestant. The projected tipping point where non-Black *nondenoms* are more likely to attend a Mainline Protestant church than a Conservative Protestant one is between those who attend less than annually and those who never attend at all. Finally, it is worth mentioning that, no matter their race or attendance frequency, the probability that a nondenominational respondent attends a non-Protestant religious congregation is minimal, further validating our decision to restrict our plausible classifications of *nondenoms* to Protestant groups.

We should note that using a model to project where someone who never attends religious services attends a congregation of a particular tradition is paradoxical. In almost every case, nondenominational respondents who never attend religious services do not belong to a congregation of a certain tradition. At the same time, neither are these respondents religiously unaffiliated. The traditions of the congregations respondents attend merely serve as stand-ins for the religious tradition certain *nondenoms* are affiliated with. Because non-Black *nondenoms* are more likely to attend a Mainline Protestant church the less frequently they attend religious services, we are inferring those who never attend religious services may be more likely to affiliate with Mainline Protestantism than those who do. Using these projections, we identified our final candidate classification method wherein we code Black *nondenoms* as Black Protestant, non-Black *nondenoms* (including those who do not report their race) who never attend religious services as Mainline Protestant, and all other respondents as Conservative Protestant (including those who do not report their race or frequency of attendance).

Table 1 shows the match rates between the predictions of the XGBoost algorithm and the alternative classification methods and the percentage of nondenominational respondents each method leaves unclassified. The OG and BP-CP algorithms performed the best according to this metric, matching the XGBoost predictions approximately 68 percent to 69 percent of the time. However, the OG algorithm did this, in part, by leaving half of nondenominational respondents uncategorized. If we consider unclassified respondents unmatched, the OG algorithm's match rate plummets to 34 percent while the match rates of the others remain unchanged. Finally, the match rates of the CP and BP-CP-MP methods were two to eight points lower than the other two classification schemes.

Although match rates are an indication of the quality of an alternative classification method, they alone cannot tell us which of these produces the least bias. We also need to consider how

Figure 3
Root-mean-square error and 95% confidence intervals between the probability algorithm estimates for an individual affiliating with each religious group by year and the XGBoost predictions [Color figure can be viewed at wileyonlinelibrary.com]

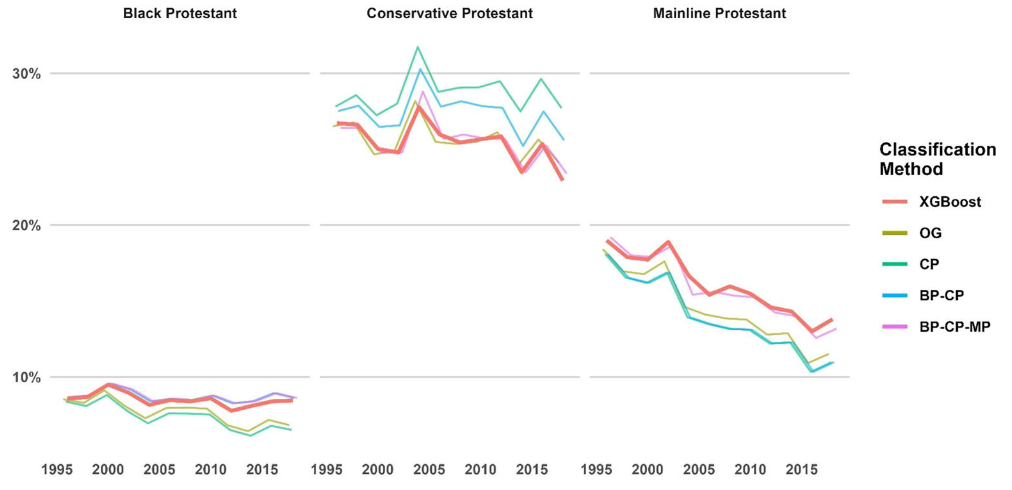


accurately each of these methods captures the size and composition of each religious group. For the former, we measured each method's bias by calculating the RMSE between the probability each algorithm estimates for an individual affiliating with each religious group by year and the XGBoost predictions. Figure 3 shows the RMSE for each method by RELTRAD group. Overall, the predictions of the BP-CP-MP algorithm consistently produce a low RMSE, meaning this method consistently makes nearly identical estimates to the XGBoost predictions. The OG algorithm comes in a distant second. It performed well in estimating the proportion of nondenominational Conservative Protestants each year, but resulted in a higher RMSE by underestimating the proportion of *nondenoms* who resemble Black and Mainline Protestants. Moreover, by leaving half of nondenominational respondents unclassified, it consistently overestimated the size of non-Protestant religious groups. The BP-CP algorithm comes in third, despite boasting the highest match rate, by overestimating the proportion of *nondenoms* who resemble Conservative Protestants and failing to classify any *nondenoms* as Mainline Protestant. Finally, the CP method fared the worst by grossly overestimating the proportion of *nondenoms* who resemble Conservative Protestants.

Applying the BP-CP-MP method to track the religious composition of the United States over time reinforces its accuracy. Figure 4 shows how each classification scheme measures the proportion of Americans in each of the Protestant religious groups since 1996—the year when the proportion of *nondenoms* in the United States began to consistently increase. Besides its one-point overestimate of the proportion of Black Protestants in the United States in recent years and slight underestimate of the proportion of Mainline Protestants, the BP-CP-MP algorithm's estimates are, again, nearly identical to the XGBoost predictions. The OG algorithm, by contrast, consistently underestimates the proportion of Black and Mainline Protestants by failing to place any *nondenoms* into these groups. Likewise, the shortcomings of the CP and BP-CP methods described above are apparent here. Adopting the BP-CP-MP algorithm instead of the alternatives is a straightforward way of sharpening researchers' understanding of these broad trends in religious affiliation.

In order to examine how accurately each method captures the composition of each religious group, we consider the cross-model marginal effect differences between the estimates of models incorporating the XGBoost predictions and those using each alternative classification method. These differences are displayed in Figure 5 and capture both the magnitude and direction of the

Figure 4
The proportion of people in different religious groups estimated by each of the nondenominational classification schemes in the US GSS, 1996–2018 [Color figure can be viewed at wileyonlinelibrary.com]



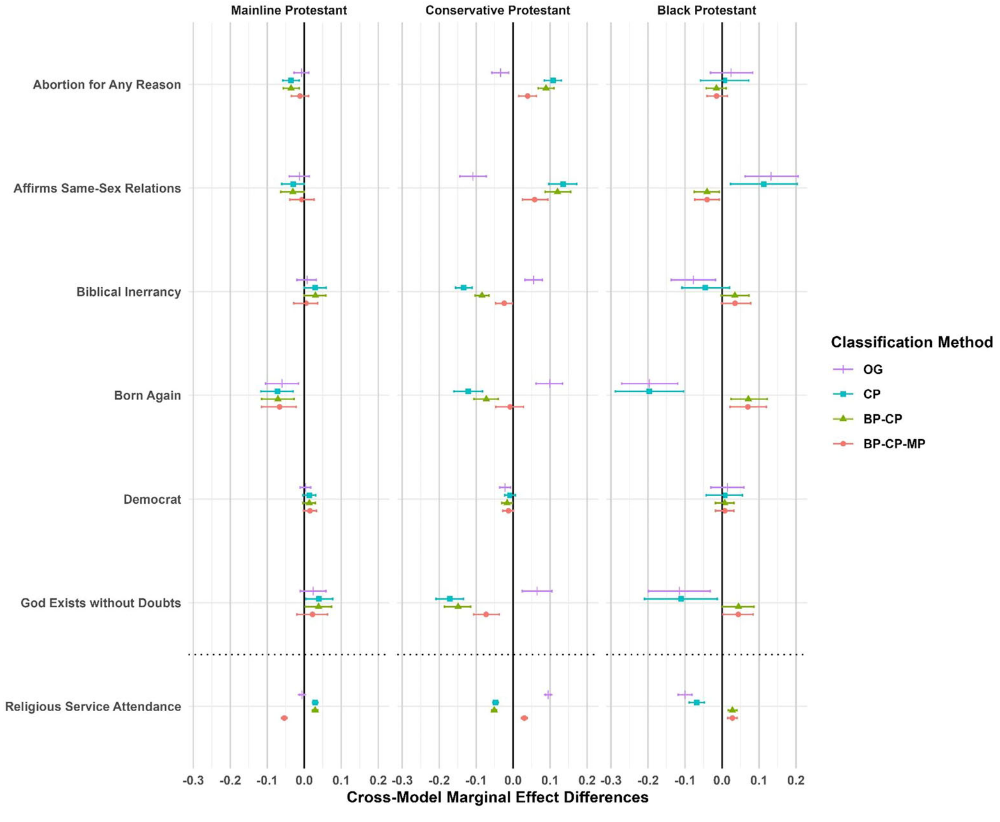
bias each algorithm introduces for each dependent variable and religious group. When we look at how each classification method biases the composition of Black Protestants, we can see that the BP-CP and BP-CP-MP algorithms, which categorize Black Protestants identically, boast lower average absolute cross-model marginal effect difference among the logit models (0.035) than the CP (0.08) and OG (0.093) algorithms. The direction of each algorithm's bias depends on whether it classifies Black *nondenoms* as Black Protestant. The methods that do (i.e., BP-CP and BP-CP-MP) have a slight conservative bias, while those that do not (i.e., OG and CP) have a subtle liberal bias.

When we consider the extent to which each method biases the composition of Mainline Protestants, all the methods perform similarly, with the OG and BP-CP-MP algorithms standing slightly above the rest. Their average absolute cross-model marginal effect difference among the logit models are 0.019 and 0.021, respectively—compared to the average of 0.037 of the other methods. Nevertheless, the overall bias of these alternative algorithms when considering the composition of Mainline Protestants is minimal. The exception to this is that each method significantly underestimates the proportion of Mainline Protestants who report having had a "born again" experience. The direction of each method's bias is less clear here. Each method produces bias in the same direction for each item. However, it varies whether this bias is liberal or conservative.

The accuracy of each classification diverges when we look at how well each measures the composition of Conservative Protestants. Here, we find that BP-CP-MP clearly performs better than the alternatives, with an average absolute cross-model marginal effect difference among the logit models for Conservative Protestants of 0.036—compared to the average of 0.064–0.11 for the other methods. This algorithm deviates minimally the XGBoost estimates while not consistently exhibiting an ideological bias in one direction. The BP-CP and CP algorithms come in a distant second and third, respectively. Both methods have a consistent liberal bias due to their classification of *nondenoms* who never attend religious services as Conservative Protestant. Finally, the OG algorithm comes in last, exhibiting a consistently strong conservative bias on all measurements due to its exclusion of low-attending *nondenoms*.

Finally, we can consider how each algorithm biases our measurement of religious service attendance. By excluding low-attending *nondenoms* and not classifying any as Mainline Protestant, the OG algorithm measures the attendance of Mainline Protestants almost perfectly. The CP

Figure 5
This shows the cross-model marginal effect differences and 95% confidence intervals between using XGBoost predictions for religious categorizations and each of the alternative classification techniques. The columns display how these classification differences impact the measurement of each Protestant religious group, and the rows illustrate how this varies by the outcome we are measuring [Color figure can be viewed at wileyonlinelibrary.com]



and BP-CP methods, by contrast, slightly overreport the religious service attendance of Mainline Protestants while the BP-CP-MP method, by only classifying never-attending *nondenoms* as Mainline Protestant, underreports this. Interestingly, the BP-CP-MP algorithm has the least bias when measuring the attendance of Conservative Protestants. Although it slightly overreports this by not classifying never-attending *nondenoms* as Conservative Protestant, doing so would lead to underreporting, as evidenced by bias of the CP and BP-CP algorithms. However, the OG algorithm's monthly attendance cutoff goes too far, leading to a considerably stronger overreporting of Conservative Protestant religious service attendance. Finally, our results suggest that classifying Black *nondenoms* as Black Protestant most accurately captures the religious service attendance of Black Protestants. The failure to do so leads to the considerable underreporting of Black Protestant religious service attendance, implying that nondenominational Black Protestants attend services more frequently than denominational Black Protestants.

If we take the average absolute cross-model marginal effect difference among these linear models, the BP-CP and BP-CP-MP algorithms perform the best at 0.036 and 0.037 respectively. The CP method comes in third at 0.049, with the OG algorithm coming in a distant last at 0.067. Altogether, this suggests that using attendance to construct part of RELTRAD does not necessarily lead to considerable bias when measuring this variable. Although the OG algorithm biases religious service attendance more than the others, a major reason for is its exclusion of nondenominational respondents attending religious services less than monthly. Without this cutoff, the BP-CP-MP accurately captures the religious service attendance frequency of each of these religious groups.

In sum, these findings suggest that the BP-CP-MP method is considerably less biased than the alternatives. Its success measuring Conservative Protestants stems from its unique ability to avoid ideological bias by classifying all *nondenoms*, except Black respondents and individuals who do not attend religious services, into this group. It captures the size of each religious group and composition of Conservative Protestants more accurately than the others, while providing similarly minimal levels of bias when measuring the composition of Black and Mainline Protestants.

Discussion

Given its consistent success in minimizing bias by accurately capturing the size and composition of each religious group, we recommend that researchers use the BP-CP-MP algorithm to classify nondenominational respondents. That is, analysts should classify Black *nondenoms* as Black Protestant; those who report never attending religious services as Mainline Protestant; and those who do attend (or do not report their attendance at all) as Conservative Protestant. Although this classification system does not match our XGBoost predictions at the highest rate, it is the best way of dealing with the heterogeneity of nondenominational respondents. It does this by ensuring the overall size and composition of each Protestant group are consistent with the XGBoost model's predictions. For the sake of simplicity, we term this new method RELTRAD2.

By proposing RELTRAD2, we are contending that nondenominational respondents are religiously heterogeneous. Several of our findings support this. No more than two-thirds *nondenoms* have resembled adherents of any particular religious tradition on a consistent basis; the religious composition of nondenominational respondents have shifted considerably since the inception of the GSS; and the majority of *nondenoms* who report attending religious services at least annually are not a part of a nondenominational congregation. Altogether, this evidence counters Burge and Djupe's (2021) idea nondenominational respondents comprise a cohesive religious tradition distinct from Black, Conservative, and Mainline Protestantism. Instead, we need to concede that *nondenoms* are an amalgamation of people who do not identify as unaffiliated or with a denomination for a number of reasons. Although many in this population may hold evangelical views and attend independent churches as Steensland et al. (2000) predicted, others may attend an independent Black Protestant church (Thumma and Travis 2007), not be attached to any denomination in particular, or simply not know the name of their denomination (Lehman and Sherkat 2018). Although we used theory to develop a list of plausible religious traditions for classifying nondenominational respondents and simple methods of disentangling these subgroups, determining which method minimizes bias while classifying these respondents had to be an empirical enterprise.

We can attribute the success of RELTRAD2 to several factors. First, the majority of Black nondenominational respondents resemble Black Protestants more than any other religious group. Classifying all these respondents as Black Protestant is a simple way of more accurately categorizing this population. Nevertheless, RELTRAD2 does this imperfectly. Because our XGBoost model predicts that not every Black nondenominational respondent is Black Protestant, RELTRAD2 overestimates the size of this group. Unfortunately, there is no straightforward method of remedying this. Even among the Black nondenominational respondents least likely to resemble Black Protestants (i.e., those who attend services more than weekly), we predict the large majority are Black Protestant. Consequently, applying an attendance cutoff wherein we classify certain Black respondents who attend religious services frequently as Conservative Protestant will not improve the accuracy of this method. The best we can do—so long as we want a simple method—is to classify all Black nondenominational respondents as Black Protestant.

Next, RELTRAD2 effectively distinguishes Conservative and Mainline Protestants. It does this by acknowledging that most *nondenoms* resemble Conservative Protestants. This is increasingly true as we consider *nondenoms* who attend religious services more frequently. RELTRAD2, therefore, classifies non-Black *nondenoms* who never attend religious services as Mainline Protestant and the rest as Conservative Protestant. Although, our XGBoost algorithm predicts that most of these never-attending *nondenoms* resemble Conservative Protestants more than Mainline Protestants, the benefits of this approach outweigh this drawback. Splitting the remainder of *nondenoms* using this attendance cutoff accurately tracks the proportion who resemble Conservative and Mainline Protestants over time. Moreover, never-attending *nondenoms* resemble Mainline Protestants more than Conservative Protestants in the aggregate, even though our XGBoost predictions classify the majority as the latter. This is because this group of *nondenoms* holds a range of beliefs and identities, wherein the average respondent looks like a Mainline Protestant despite the median resembling an evangelical. As a result, RELTRAD2 consistently outperforms the other algorithms. Although it incorporates religious services attendance into the classification of certain nondenominational respondents, it does this without meaningfully narrowing or widening the religiosity gap between Conservative and Mainline Protestants. Instead, this method avoids the OG algorithm's conservative bias and the CP and BP-CP methods' liberal bias toward evangelicals while yielding a minimal amount of liberal bias among Mainline Protestants.

The RELTRAD2 method is not perfect. It trades simplicity for accuracy. When working with a large, representative dataset, imputing the religious affiliation of *nondenoms* using an advanced algorithm like XGBoost is preferable. Sophisticated imputation techniques tend to alleviate the issue of relying on other variables to make RELTRAD classifications. Unlike the simple classification methods, no one variable plays a determinative role when predicting religious affiliations, limiting the extent to which these imputed classifications may confound statistical analyses that incorporate both RELTRAD and the other variables used to construct it.

However, categorizations made with RELTRAD2 do a good job approximating the results from more sophisticated imputation methods. It is a simple approach that researchers can readily apply with simple modifications to existing code. Although it relies upon race and religious service attendance to classify nondenominational respondents, this is nothing new. RELTRAD has always depended upon these variables to classify certain respondents. In our analyses, which control for race, RELTRAD2 consistently outperforms the other algorithms. Moreover, it accurately captures the religious service attendance of Black, Conservative, and Mainline Protestants, never misestimating their attendance frequency by more than 0.054 standard deviations. Although researchers should still be mindful of RELTRAD2's use of these two variables and the direction of bias this method introduces, this issue is not unique to RELTRAD2. Any simple categorization scheme will inevitably misclassify certain nondenominational respondents and thereby confound statistical analyses. With RELTRAD2, our analysis shows that, despite its use of race and religious service attendance, the influence of this bias is minimal.

Conclusion

In recent years, RELTRAD has accumulated a growing share of criticism over its handling of nondenominational respondents. Our analyses suggest that its current method of classifying nondenominational respondents leads it to underestimate the proportion of Black and Mainline Protestants and introduce significant amount of conservative bias in its measurement of evangelicals. As a result, RELTRAD may add a considerable amount of bias when used as an independent or dependent variable. To reduce bias and improve the art of religious classification, we introduce a new method for categorizing *nondenoms*, which we term RELTRAD2. RELTRAD2 follows the standard RELTRAD code, except when classifying *nondenoms*. Instead of classifying all nondenominational respondents who attend religious services at least monthly as Conservative Protestant and leaving the remainder uncategorized, RELTRAD2 classifies all Black *nondenoms* as Black Protestant, all non-Black *nondenoms* (including those who do not provide their race) who never attend religious services as Mainline Protestant, and the rest (including those who do not report their frequency of attendance) as Conservative Protestant. We include both R and Stata code needed to generate RELTRAD2 from the GSS in an Open Science Foundation (Eagle and Gaghan 2024) repository.

The major limitation of this approach is that it relies upon data from the United States. It is not known how well this approach will work in other countries. Moreover, because RELTRAD2 is parsing a religiously heterogeneous group, it is possible that in the coming decades the ideal classification of nondenominational respondents will alter due to the changing composition of nondenominational respondents. Although RELTRAD2 is unlikely to fundamentally reshape our understanding of many aspects of religion in the United States, it markedly improves one of the most popular ways researchers use to categorize the religious affiliation of Americans. It does this by using an empirically validated method to classify *nondenoms*, who represent a rapidly growing segment of the U.S. population, into the religious families they most closely resemble. Although a consensus may one day be reached on using an alternate religious taxonomy, our best option at the moment is to improve the current state of the art of religious classification.

DATA AVAILABILITY STATEMENT

The data underlying this article will be shared on reasonable request to the corresponding author.

Conflict of Interest Statement

The authors report no conflicts of interest on this project.

Funding

This project was funded by a grant from the Rural Church Area of The Duke Endowment.

References

Baunach, Dawn M. 2012. Changing same-sex marriage attitudes in America from 1988 through 2010. *Public Opinion Quarterly* 76(2):364–78. https://doi.org/10.1093/poq/nfs022.

Bengtson, Vern L., R. David Hayward, Phil Zuckerman, and Merril Silverstein. 2018. Bringing up nones: Intergenerational influences and cohort trends. *Journal for the Scientific Study of Religion* 57(2):258–75. https://doi.org/10.1111/jssr.12511.

Bentéjac, Candice, Anna Csörgő, and Gonzalo Martínez-Muñoz. 2021. A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review* 54(3):1937–67. https://doi.org/10.1007/s10462-020-09896-5.

Burge, Ryan P., and Paul Djupe. 2021. Falling through the cracks: Dealing with the problem of the unclassifieds in RELTRAD. *Review of Religious Research* 63(3):411–33. https://doi.org/10.1007/s13644-020-00441-y.

Chai, Tianfeng, and Roland R. Draxler. 2014. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* 7(3):1247–50. https://doi.org/10.5194/gmd-7-1247-2014.

Chaves, Mark. 2017. *American religion: Contemporary trends* (Second edition). Princeton, NJ: Princeton University Press.

Chaves, Mark, Shawna Anderson, Alison Eagle, and Mary Hawkins. 2020. *National Congregations Study. Cumulative data file and codebook*. Durham, NC: Department of Sociology, Duke University.

Chaves, Mark, Mary Hawkins, Anna Holleman, and Joseph Roso. 2020. Introducing the fourth wave of the National Congregations Study. *Journal for the Scientific Study of Religion* 59(4):646–50. https://doi.org/10.1111/jssr.12684.

Chen, Tianqi, and Carlos Guestrin. 2016. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–94. San Francisco, CA: ACM. https://doi.org/10.1145/2939672.2939785.

Chen, Tianqi, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, Mu Li, Junyuan Xie, Min Lin, Yifeng Geng, Yutian Li, Jiaming Yuan, 2023. Xgboost: Extreme gradient boosting. https://cran.r-project.org/web/packages/xgboost/index.html.

Chen, Zhuo, Fu Jiang, Yijun Cheng, Xin Gu, Weirong Liu, and Jun Peng. 2018. XGBoost classifier for DDoS attack detection and analysis in SDN-based cloud. In *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 251–56. Shanghai, China. https://doi.org/10.1109/BigComp.2018.00044.

Davern, Michael, Rene Bautista, Jeremy Freese, Pamela Herd, and Stephen L. Morgan. 2024. General Social Survey 1972-2024. [Machine-readable data file]. Principal Investigator, Michael Davern; Co-Principal Investigators, Rene Bautista, Jeremy Freese, Pamela Herd, and Stephen L. Morgan. Sponsored by National Science Foundation. NORC ed. Chicago: NORC, 2024: NORC at the University of Chicago [producer and distributor]. Data accessed from the GSS Data Explorer. website at gssdataexplorer.norc.org.

Dougherty, Kevin D., Byron R. Johnson, and Edward C. Polson. 2007. Recovering the lost: Remeasuring U.S. religious affiliation. *Journal for the Scientific Study of Religion* 46(4):483–99. https://doi.org/10.1111/j.1468-5906.2007.00373.x.

Eagle, David, and Joshua Gaghan. 2024. Refining Religious Classification Project. *OSF* March 25. https://doi.org/10.17605/OSF.IO/EQ24N.

Friedman, Jerome H. 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis, Nonlinear Methods and Data Mining* 38(4):367–78. https://doi.org/10.1016/S0167-9473(01)00065-2.

Guyon, Isabelle, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Machine learning* 46(1):389–422. https://doi.org/10.1023/A:1012487302797.

Hackett, Conrad, Philip Schwadel, Gregory A. Smith, Elizabeth Podrebarac Sciupac, and Claire Gecewicz. 2018. Choose the method for aggregating religious identities that is most appropriate for your research. *Journal for the Scientific Study of Religion* 57(4):807–16. https://doi.org/10.1111/jssr.12555.

Hesterberg, Tim. 2011. Bootstrap. *WIREs Computational Statistics* 3(6):497–526. https://doi.org/10.1002/wics.182.

Hout, Michael. 2017. Religious ambivalence, liminality, and the increase of no religious preference in the United States, 2006–2014. *Journal for the Scientific Study of Religion* 56(1):52–63. https://doi.org/10.1111/jssr.12314.

Lehman, Derek, and Darren E. Sherkat. 2018. Measuring religious identification in the United States. *Journal for the Scientific Study of Religion* 57(4):779–94. https://doi.org/10.1111/jssr.12543.

Lim, Chaeyoon, Carol Ann MacGregor, and Robert D. Putnam. 2010. Secular and liminal: Discovering heterogeneity among religious nones. *Journal for the Scientific Study of Religion* 49(4):596–618.

Lim, Chaeyoon, and Robert D. Putnam. 2010. Religion, social networks, and life satisfaction. *American Sociological Review* 75(6):914–33. https://doi.org/10.1177/0003122410386686.

McKinney, Jennifer, and Roger Finke. 2002. Reviving the mainline: An overview of clergy support for evangelical renewal movements. *Journal for the Scientific Study of Religion* 41(4):771–83. https://doi.org/10.1111/1468-5906.t01-1-00153.

Mize, Trenton D., Long Doan, and J. Scott Long. 2019. A general framework for comparing predictions and marginal effects across models. *Sociological Methodology* 49(1):152–89. https://doi.org/10.1177/0081175019852763.

Ogunleye, Adeola, and Qing-Guo Wang. 2020. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 17(6):2131–40. https://doi.org/10.1109/TCBB.2019.2911071.

Schnabel, Landon, Sean Bock, and Mike Hout. 2022. Switch to web-based surveys during Covid-19 pandemic left out the most religious, creating a false impression of rapid religious decline. https://doi.org/10.31235/osf.io/g3cnx.

Shelton, Jason E. 2018. Is RELTRAD still the gold standard? *Journal for the Scientific Study of Religion* 57(4):817–26. https://doi.org/10.1111/jssr.12557.

Shelton, Jason E., and Ryon J. Cobb. 2017. Black RELTRAD: Measuring religious diversity and commonality among African Americans. *Journal for the Scientific Study of Religion* 56(4):737–64. https://doi.org/10.1111/jssr.12378.

Skirbekk, Vegard, Eric Kaufmann, and Anne Goujon. 2010. Secularism, fundamentalism, or Catholicism? The religious composition of the United States to 2043. *Journal for the Scientific Study of Religion* 49(2):293–310. https://doi.org/10.1111/j.1468-5906.2010.01510.x.

Steensland, Brian, Lynn D. Robinson, W. Bradford Wilcox, Jerry Z. Park, Mark D. Regnerus, and Robert D. Woodberry. 2000. The measure of American religion: Toward improving the state of the art. *Social Forces* 79(1):291–318. https://doi.org/10.1093/sf/79.1.291.

Steensland, Brian, Robert D. Woodberry, and Jerry Z. Park. 2018. Structure, placement, and the quest for unidimensional purity in typologies of American denominations. *Journal for the Scientific Study of Religion* 57(4):800–806. https://doi.org/10.1111/jssr.12559.

Thiessen, Joel, and Sarah Wilkins-Laflamme. 2020. *None of the above: Nonreligious identity in the US and Canada*. New York: NYU Press.

Thumma, Scott, and Dave Travis. 2007. *Beyond Megachurch Myths: What We Can Learn from America's Largest Churches*. Hoboken, NJ: John Wiley & Sons.

Torlay, L., M. Perrone-Bertolotti, E. Thomas, and M. Baciu. 2017. Machine learning–XGBoost analysis of language networks to classify patients with epilepsy. *Brain Informatics* 4(3):159–69. https://doi.org/10.1007/s40708-017-0065-7.

Uecker, J. E., M. D. Regnerus, and M. L. Vaaler. 2007. Losing my religion: The social sources of religious decline in early adulthood. *Social Forces* 85(4):1667–92. https://doi.org/10.1353/sof.2007.0083.

Vernon, Glenn M. 1968. The religious 'nones': A neglected category. *Journal for the Scientific Study of Religion* 7(2):219–29. https://doi.org/10.2307/1384629.

Voas, David. 2009. The rise and fall of fuzzy fidelity in Europe. *European Sociological Review* 25(2):155–68. https://doi.org/10.1093/esr/jcn044.

Woodberry, Robert D., Jerry Z. Park, Lyman A. Kellstedt, Mark D. Regnerus, and Brian Steensland. 2012. The measure of American religious traditions: Theoretical and measurement considerations. *Social Forces* 91(1):65–73. https://doi.org/10.1093/sf/sos121.