Journal of Religion and Health (2019) 58:271–288
https://doi.org/10.1007/s10943-018-0737-5

ORIGINAL PAPER

A Data Mining and Data Visualization Approach to Examine the Interrelationships Between Life Satisfaction, Secularization and Religiosity

Chong Ho Yu[1] · Hailey Trier[2] · Maryann Slama[3]

Published online: 22 November 2018

Abstract

Previous studies have suggested a positive relationship between secularization and happiness, as well as a geographical, cultural and development pattern, primarily based on data gathered in Europe and the USA. To gain a more holistic view on the interrelationships between secularization, religiosity and subjective perception of well-being, this study utilized the Wave 6 archival data set (2010–2014) of World Values Survey, which contains 74,042 observations from 60 countries. In this study, the rating of satisfaction with life was treated as the dependent variable. Four secular value indices and 12 variables related to religiosity were extracted from the data set for predictive analysis. Data mining tools, such as the partition tree and bootstrap forest approaches, suggested that only secular values were influential. Specifically, secular values had a negative relationship with satisfaction. In addition, hierarchical clustering based on secularization and satisfaction did not suggest a meaningful pattern. For example, the dendrogram showed that South Korea, Lebanon, Estonia and Algeria were grouped together. This implies that secularization and satisfaction could vary from country to country, regardless of their geographical location, culture and development status. Specifically, countries that are similar in terms of satisfaction and secularization are not necessarily culturally similar or geographically clustered. This discourages the idea that these factors (e.g., location, development status, culture) play a role in mediating the relationship between secularization and religiosity. By directly contradicting previous work showing a pattern based on Europe and the USA, this finding challenges the existing understanding of the relationship between these factors. By expanding the scope of study to the whole world, the current analysis suggested that the existing view regarding the positive relationship between secularization and well-being might be oversimplified.

✉ Chong Ho Yu
cyu@apu.edu; chonghoyu@gmail.com

Hailey Trier
trier.hailey@gmail.com

Maryann Slama
mslama12@apu.edu

[1] Department of Psychology, Azusa Pacific University, Azusa, CA, USA

[2] Department of Psychology, Oxford University, Oxford, England

[3] Program of Industrial and Organizational Psychology, St. Mary's University, Halifax, Canada

Keywords Secularization · Religiosity · Well-being · Data mining · Big data analytics · Happiness · Mental health

Introduction

Several scholars have argued that socioeconomic (objective) well-being and secularization are positively correlated. According to Paul (2014a, b), religious America has a lower level of well-being than secular Europe in social well-being indices such as incarceration rate, divorce rate, homicide rate and abortion rate. Furthermore, qualitative studies by Zuckerman (2008, 2012) indicated that both subjective well-being and objective well-being are tied to secular values. To demonstrate this, Zuckerman (2008, 2012, 2014, 2015) argued that the excellent educational systems, strong economies, well-supported arts, universal health care and just social policies in Sweden and Denmark have created a healthy environment in which citizens do not feel any need for religious comfort. He found that secularists and people who abandoned religion are psychologically healthy, life-affirming, courageous, smart, deeply moral and open-minded.

However, a recent study (Yu et al. 2017) with the archival data sets found in United Nations Human Development, Gallap's Global Wellbeing Index and WVS indicated that the preceding blanket statement might be misleading. In some cases, secularization or lack of religiosity is seemingly linked to better quality of life (e.g., disbelief and inequality-adjusted human development index), while in other cases, the relationship is reversed (e.g., skepticism and adolescent birth rate). In most situations, there is no association at all. In addition, many past studies narrowly focused on comparing secular Europe and relatively religious America. To gain a more holistic view on the interrelationships between secularization, religiosity and subjective perceptions of well-being, the present study utilized the Wave 6 archival data set (2010–2014) of World Values Survey (WVS), which contains 74,042 observations from 60 countries. Further, because the relationships between objective (socioeconomic) well-being, religiosity and secularization have been previously examined by Yu et al. (2017), this study focused on subjective well-being only. In this study, the subjective rating of satisfaction with life was treated as the dependent variable. It is important to point out that satisfaction with life is one of many aspects of psychological well-being (Ryff and Keyes 1995; Ryff 2014). The authors by no means equated psychological well-being with life satisfaction only, although this analysis focused on life satisfaction because it was necessary to limit the scope to a single dependent variable. Five secular value indices and twelve variables related to religiosity were extracted from the data set for predictive analysis. With such a large sample size and multiple independent variables, conventional statistical procedures would face the insurmountable problem of excessive statistical power and the threat of multi-collinearity. As a remedy, data mining tools, including recursive partition trees and the bootstrap forest, and data visualization tools, such as quantile smoothing, were employed to analyze this large-scale data set.

Literature Review

There is a growing body of literature concerning the relationship among secularization, religiosity and individual well-being. Currently, however, there is no consensus on the definitions of secularization and religiosity among researchers. Very often their meanings depend on the geographical location under study, the authors' positions on religion and

methods of data collection (e.g., Berger 2014; Berger et al. 2008; Bruce 2002; Eberstadt 2013; Norris and Inglehart 2012; Pollack 2008; Stark 1999; Wilson 1969, 1976a, b, 1982, 1994; Zuckerman 2008, 2012, 2014, 2015). Nonetheless, certain common threads can be found in the literature. Increasingly, religiosity is not necessarily tied to attending religious services or participating in religious activities; rather, it can be simply defined as a belief in the existence of a supernatural being or cosmic power and in one's dependence on the providence of such a being or power (Bowland et al. 2012; Hill and Pargament 2008). Conversely, secularism is to distance oneself from this belief. Simply put, in most cases religiosity and secularization are in a seesaw relationship. Often when an author argues for the positive impact of secularization on well-being, the argument implies or bluntly declares that religiosity results in adverse effects (e.g., Paul 2014a, b; Zuckerman 2008, 2012, 2014, 2015). The present authors are well aware of the diversity across different cultures and religions. Some religions are more individualistic (e.g., Protestantism), whereas some are more collective (e.g., Catholicism). Some emphasize subjective experience (e.g., Pentecostalism), while some value intellectual inquiry (e.g., Buddhism). However, the manner in which specific rituals, traditions and theologies of different religions influence happiness is not the focus of this study. Rather, the focal interest is whether the belief in a supreme power can be observed to correlate with a self-reported happier life. Hence, religiosity, in this study, is defined in an ecumenical sense (belief in a supreme being or a cosmic power). By the same token, there are also different versions of secularism. Passive secularism (e.g., Europeans) is indifferent but not hostile to religion, whereas militant secularism (e.g., New Atheism) devotes tremendous effort to refuting religion (de Waal 2014). Again, in a cross-cultural study using worldwide data, it is impractical to put this distinction into account.

Subjective well-being is also a difficult construct to define. Different organizations, such as United Nations Development Programme (2015), Gallup (2014) and Happy Planet Index (HPI) (New Economic Foundation 2015), release their rankings of well-being by country based on different criteria. Interestingly, very often the ranking reports of various organizations are vastly different. For example, the UN ranks secular European nations on top of the human development list, whereas HPI rates developing countries, such as Costa Rica, Vietnam, Colombia, Belize and Cuba, as the happiest nations in the world. This may be attributed to the fact that in HPI one of the elements of a happy life is preservation of the natural environment (ecological footprint), and as a result, less industrialized developing countries received higher scores than their developed counterparts because of lower carbon emissions. The mixed results of this type of ranking can be partly attributed to the diverse meanings of well-being. Hence, instead of considering various objective well-being indices (e.g., gross national income, ecological footprint and gender equality), this article narrowly focuses on the perception of well-being (i.e., life satisfaction). It does not mean that the controversy stops here once and for all; nonetheless, it makes the current review and data analysis more manageable.

A number of measures of well-being have been examined for the present study, from objective observations to subjective self-reports of personal well-being; however, both positive and negative correlations are demonstrated in the literature. Lynch (2012) conducted a study on the associations between self-esteem, perceived stress, general health differences and religiosity, in which 100 Irish undergraduate students completed four questionnaires each: the Francis Scale of Attitude Towards Christianity (Francis and Stubbs 1987), the Rosenburg Self-Esteem Scale (Rosenburg 1965), the Perceived Stress Scale (Cohen et al. 1983) and the General Health Questionnaire (Goldberg 1992). It was found that there was a positive correlation between secularization and perceived stress

levels. In addition to this finding, it was found that students who were identified as less religious showed a decrease in psychological health through measures from the General Health Questionnaire (Goldberg 1992).

Congruently, Hayward et al. (2016) analyzed data from a survey on religion and health in American adults to determine group differences—both physical and psychological—between religious group members and non-religiously affiliated individuals (atheist, agnostic and other). Of interest to the current review were the measures of subjective well-being: positive affect, composite happiness (e.g., "I consider myself [not a very happy person…a very happy person]"), life satisfaction (e.g., "The conditions of my life are excellent") and self-esteem (e.g., "I feel I am a person of worth…"). Hayward et al. (2016) discovered that religiously affiliated individuals and those with no preference tended to report higher measures of psychological well-being than atheists and agnostics, including higher mean happiness and mean self-esteem. Further, religiously affiliated individuals had higher mean levels of life satisfaction than agnostics. This study was notable because it included a wide range of measures for well-being that measured multiple aspects of the subjective life experience.

Zullig et al. (2006) studied relationships between self-perceived health, perceived spirituality, religiosity and life satisfaction among a sample of American college students. Results showed that perceived spirituality and life satisfaction were mediated by self-perceived health, and that perceived religiosity and life satisfaction were practically mediated by self-perceived health. In interpreting these results, Zullig et al. (2006) highlighted the theoretical connection between self-perceived health and life satisfaction; i.e., perceiving that one is in good physical and mental health likely leads to increased life satisfaction. They also mentioned asset theory, which posits that positive material, social and personal characteristics can be used as tools to achieve goals and fulfill personal needs; Zullig et al. (2006) argued that religion and spirituality are assets that help students fulfill personal goals and needs. These theories contributed to the overall hypothesis that college students who describe themselves as either spiritual or religious are more likely to perceive that they are in better health, leading to a conclusion of greater life satisfaction in general.

For a perspective highlighting the influence of culture and context, Sinnewe et al. (2015) studied the relationship between religiosity and life satisfaction in both East Germany and West Germany. The authors stratified data between East and West due to the residual effects of political, social and economic division in the twentieth century, and found that the two regions yielded different conclusions: West German respondents displayed a positive correlation between religious service attendance and individual life satisfaction, while East German respondents yielded no such association. The study also investigated social networks as a possible variable and found that social networks only partially mediate religious service attendance, indicating that religiosity exerts a substantial direct influence on life satisfaction throughout Germany. These data show that attending religious services regularly may impact personal life satisfaction regardless of the closeness or security of one's social groups. In addition, this study provides evidence for the influence of cultural, historical and social factors in the interaction between religiosity and life satisfaction. The interesting interaction between geography and the religiosity/life satisfaction correlation in this study encouraged the authors of the present study to look into the correlation between similar variables on a global scale.

In weighing the results of these studies, it is important to consider other factors that may influence the relationship between the variables. In particular, a number of studies indicate that social contexts often mediate the relationship between religiosity and personal well-

being: Religious acceptance, family orientation, cultural context and level of societal development all play a role.

Galen (2015) contended that positive correlations between religion and well-being are due not to beliefs but to social engagement and embeddedness in supportive groups. The author pointed out that many studies examining the benefits of religious belief analyze measures of religiosity that incorporate the social context or do not pertain specifically to beliefs, such as attendance of religious services or commitment to religious practices. Galen (2015) also raised important considerations such as the disconnection between life meaning and overall happiness; it is possible that one be unconvinced that life has an underlying meaning and yet fully satisfied with life in general. From that article, it is evident that underlying factors alternative to religious beliefs must be scrutinized in order to understand the relationships between secularization and subjective well-being.

Green and Elliott (2010) analyzed data from the 2006 data set of the General Social Survey (GSS) in their study of religion, physical health and psychological well-being. The GSS consisted of 4510 adult participants from throughout the USA. Information about health and happiness was collected through self-report on a scale of one to five. Questions about religious beliefs included topics such as religious identification, liberality of beliefs, beliefs about the consequences of sin, strength of beliefs and frequency of religious activities. What Green and Elliott (2010) concluded was that people who held beliefs that were fundamental to their religion were happier with regard to how they identified with their religion than those who took a more liberal stance. No significance was seen in the relationship between religious affiliation and physical health or happiness; however, interestingly, what was found to make a significantly positive difference on both health and happiness was the extent to which a person identified with the doctrine of their religion and felt acceptance within the community. Practice in religious events and rituals had no benefit unless the participant found personal meaning in these events.

In another study on individual life satisfaction and secularization, Sabatier et al. (2011) studied relationships between religiosity and life satisfaction in youth (M=15.76 years of age) from four countries: France, Germany, Poland and the USA. Their study examined not just religiosity and life satisfaction but also the mediating roles of family orientation and country context of religiosity to form a more detailed picture of the factors influencing individual life satisfaction. Family orientation in particular was a variable of interest because the lives of adolescents are typically highly impacted by the connectedness of the family; researchers considered the family relationship value (value placed on harmony, obligations and support within the family) and family interdependence in evaluating the role of family orientation. Results showed that American and Polish adolescents reported a higher importance of religion than their French and German counterparts, and that family orientation moderated the relationship between subjective religiosity and life satisfaction in all countries. These relationships bear further scrutiny in order to fully understand the interrelationships between factors affecting individual life satisfaction.

Li and Bond (2010) examined the type of cultural, historical and social factors that were acknowledged by Sinnewe et al. (2015) in the above geographically stratified study, and studied the factors' relationships to individual life satisfaction, personal secularization and societal secularization. As a measure of societal development, these authors used the human development index (HDI) as a variable and measured the extent to which it mediated secularism and life satisfaction in four waves of the World Value Survey. Li and Bond (2010) found that prior to the 1990s, secular values inversely correlated with life satisfaction in all societies; after the 1990s, however, societies that were high-HDI showed

a positive correlation between secular values and life satisfaction, while low-HDI societies retained an inverse correlation.

However, the Li and Bond's (2010) hypothesis must be considered with limitations. Bonini (2008) discovered that large-scale factors such as region, national wealth and human development accounted for just 19% of the variance in individual well-being measures among global respondents to the WVS; in contrast, the author attributed 81% of variance to individual characteristics and error. This analysis demonstrates that although cultural fit and environmental factors may impact individual life satisfaction, individual factors are a variable that demand consideration and are not totally mediated by societal influence. More importantly, the primary data analytical methods of Li and Bon's study are correlation and regression, which are sensitive to sample size and subject to overpowering. Specifically, one of the data sources of Li and Bond's study (2010) is World Value Survey, and the sample sizes are over 40,000 across Wave 1–4 of the survey. When correlation and regression are run with such a large sample size, the statistical power level is 0.9999, resulting in a high risk of false findings.

In summary, the relationship between religiosity/secularization and subjective well-being is not necessarily dependent on an individual's objective well-being and varies across cultural and societal contexts. Factors that may mediate the relationship between religiosity and subjective well-being are self-perceived health, family orientation, country context of religiosity, historical and geographical influences, and level of societal development. However, the impact of larger social contexts is less certain than more immediate social contexts and the influence of individual factors. Both positive and negative correlations can be found in the literature, and the inconclusive nature of this review necessitates the present analysis.

Methodology

Data Source and Variables

The data used in this study were sourced from WVS (World Values Survey Association 2014). WVS is developed and conducted by a global network of social scientists with its headquarters in Stockholm, Sweden. The project started in 1981, and since then, the organization has administered six waves of survey. In this study, the most recent data set (Wave 6, 2010–2014) was used.

The survey has 384 variables. Table 1 lists those used in this study. There are several measurement scales regarding subjective well-being in WVS, such as "state of health" (subjective), "feeling of happiness" and "satisfaction with life." However, in the question regarding subjective perception of health, the survey item does not specify whether it is referring to physical health or mental health, although most respondents likely interpreted it as physical health. In addition, four options were provided to the participants ("Very good," "Good," "Fair" and "Poor"), but balanced wordings are missing from these categories (e.g., "Very poor"). Due to its ambiguity and imbalanced options, this item was dropped from this study. "Feeling of happiness" is a 4-point Likert scale, whereas the options of "satisfaction with life" range from 1 to 10. Some authors argue that too many rating-scale categories give researchers a false sense of precision. Lincare (2015) gave this example: "On a scale of 1–10, rate the cleanliness of your hotel room." Lincare argued that no one can distinguish 10 levels of cleanliness. Although this might be true to rate external objects or events, such as cleanliness, in self-reflection people are usually able to identify

Table 1 List of variables extracted from World Values Survey

Question number/label and variable/question	Type	Scale range or options
Dependent variable		1–10
Q23: Satisfaction with your life (All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole?)	Continuous	
Independent variables indicating secular values		
SACSECVAL: Overall secular values (composite score)	Continuous	0–1
Secular value: defiance (in country-level sample only)	Continuous	0–1
Secular value: disbelief (in country-level sample only)	Continuous	0–1
Secular value: relativism (in country-level sample only)	Continuous	0–1
Secular value: skepticism (in country-level sample only)	Continuous	0–1
Independent variables indicating religiosity		
Q25: Active/inactive membership of Church or religious organization? (Now I am going to read off a list of voluntary organizations. For each organization, could you tell me whether you are an active member, an inactive member or not a member of that type of organization?)	Nominal	1 Active member 2 Inactive member 3 Not a member
Q145: How often do you attend religious services? (Apart from weddings and funerals, about how often do you attend religious services these days?)	Ordinal	1>once per week 2 Once a week 3 Once a month 4 Only on special holy days 5 Once a year 6 Less often 7 Never, practically never
Q146: How often do you pray? (Apart from weddings and funerals, about how often do you pray?)	Ordinal	1 Several times a day 2 Once a day 3 Several times each week 4 Only when attending religious services 5 Only on special holy days 6 Once a year 7 Less often than once a year 8 Never, practically never
Q147: Are you a religious person? (Independently of whether you attend religious services or not, would you say you are a religious person?)	Nominal	Religious person Not a religious person An atheist
Q148: Believe in God (do you believe in God?)	Nominal	Yes/no
Q149: Believe in hell (do you believe in hell?)	Nominal	Yes/no
Q150: What is the meaning of religion? (Which one of the following statements do you agree most?)	Nominal	1 To follow religious norms and ceremonies 2 To do good to other people
Q151: What is the meaning of religion? (Which of the following statements do you agree most?	Nominal	1 To make sense of life after death 2 To make sense of life in this world

Table 1 continued

Question number/label and variable/question	Type	Scale range or options
Q152: How important is God in your life? (Please use this scale to indicate. 10 means "very important" and 1 means "not at all important.")	Ordinal	1–10
Q9: Important in life: Religion (For each of the following, indicate how important it is in your life. Would you say it is religion?)	Ordinal	1–5
Q108: Confidence in churches (I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: Is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all? The churches)	Ordinal	1–5
Q144: Do you belong to a religion or religious denomination? If yes, which one?	Nominal	Yes/no

their position in a 10-point psychological scale. For example, the Perceived Stress Scale (PSS) used by American Psychological Association (APA) (Cohen et al. 1983) and the visual analog scale (VAS) used in psychophysical and medical research (Bond and Fox 2015) are both 10-point scales. Because a wider scale provides a wider distribution in the data, "satisfaction with life" instead of "feeling of happiness" was adopted as the dependent variable.

Although it is a common practice to treat a single Likert scale as ordinal, the researcher team did not adopt this approach because a 10-level categorical variable would render an excessively complicated model. It is not uncommon for analysts to force categorical-scaled data into continuous for modeling convenience. For example, a Pareto chart, which indicates the ranking of various causes of failure or problems, is commonly used in statistical quality control, but in this approach the categorical data are treated as continuous (e.g., 4 categories make up the top 85% of product flaws) (Grzegorzewski 2012). Besides modeling convenience, there is a theoretical justification for converting ordinal data into continuous. In item response theory and Rasch modeling, it is assumed that the response to the item is conditional on the subject's location on a latent trait, which is a continuum. Nonetheless, in a partial credit model even though different response categories are discrete, the analyst treats them as continuous in order to compute the step function based upon maximum likelihood (Falk and Cai 2016). By the same token, although a binary variable (1/0, yes/no) is also discrete in essence, a logistic function assigns probabilities to the positions between 1 and 0 as if it is a continuum.

The overall secular value is a composite score resulting from measurements of defiance, disbelief, relativism and skepticism. In other words, secularization is defined as embracing these four attitudes. Specifically, "defiance" measures a national culture's distance from "sacred" sources in the sphere of patrimonial authority. Likewise, "disbelief" is a measure of the distance from "sacred" sources in the domain of religious authority. "Skepticism" is concerned with the cultural distance from order institutions, whereas "relativism" is about how far people in a nation move away from the normative authority. These four secular values are theoretically and empirically studied for their cross-cultural reliability and validity (Welzel 2013), and only these four dimensions were adopted into WVS. The authors are well aware that there are other definitions of secularization; nonetheless, it is the conviction of the authors that the approach taken here is less ambiguous than others. For instance, in an earlier article Welzel equated secular values and rational values

(Inglehart and Welzel 2010). However, this approach is problematic because rationality is not exclusively observed in secular societies; religious people can be rational, too. More importantly, the 2013 Welzel's framework has been used in research on secularization and yielded fruitful results (e.g., Kistler et al. 2017; Welzel and Delhey 2015).

Although the breakdown scores are accessible in an online analytical tool of the WVS website at the country level, the downloadable database has one composite secular value score only. For the analysis with all individual observations, the composite secular index was the primary unit of analysis, whereas the subscale scores were utilized for the country-level analyses.

There are many variables related to religion in WVS, but some of them seem to measure degrees of religious narrow-mindedness, such as "Would not like to have people of a different religion as neighbors," "Whenever science and religion conflict, religion is always right" and "The only acceptable religion is my religion". Being religious is not necessarily being narrow-minded, and thus, this study used only variables related to religion that have a neutral tone, such as "How important is God in your life?" and "How often do you attend religious services?"

Because the data source is archival data, the researchers had no control of the instrument. In this survey, many items are standalone instead of forming a factor (construct). The authors are well aware of this limitation; nevertheless, some of the standalone items in WVS are identical to those in validated scales. For example, the question about subjective well-being (satisfaction with life) resembles the item "I am satisfied with my life" in the Satisfaction with Life Scale (SWLS; (Diener et al. 1985)). The question "How often do you attend religious services" is similar to "How often do you attend church or other religious meetings" in the Duke University Religion Index (DURE; (Koenig et al. 1997)).

In addition, in WVS there are two variables related to education, namely, "highest educational level attained" and "education (country specific)." In the former, there is only one category for university education; in the latter, the US sample breaks down undergraduate and graduate degrees. Furthermore, the variables "confidence in churches" and "denomination" are applicable only to nations where Christianity is popular and therefore were used in the US sample only.

Data Analysis

With such a huge sample size, any trivial effect would be misidentified as significant by conventional parametric tests. In addition, because there are many independent variables, the validity of regression analysis is threatened by multi-collinearity. As a remedy, the recursive partition tree (Breiman et al. 1984; Speybroeck 2012; Vanitha and Niraimathi 2013; Yu 2010), which is not subject to statistical power and immune to parametric assumptions, was employed to analyze this large data set. Furthermore, the partition tree approach tends to yield a simpler model than its regression-based counterpart. It can resolve multi-collinearity and avoid overfitting by preventing redundant variables from entering into the model.

Variable selection by data partition is based on the LogWorth statistics: $-\log10(p \text{ value})$. The partition tree examines each independent variable to identify the ones that can decisively split the sample with reference to the dependent variable. If the input is continuous (e.g., overall secular values), every value of the variable could be a potential split point. If the input is categorical (e.g., Do you believe in God?), then the average value of the outcome variable is taken for each level of the predictor variable. Afterward two

subgroups demarcated by the split point are generated and a 2×2 crosstab table is formed. Next, Pearson Chi-square is computed and the LogWorth statistics inverse the p value of the Chi-square. Unlike the p value that is tied to certain alpha levels ($p<0.1$, 0.05 or 0.01), there is no cutoff in LogWorth. The JMP software (SAS Institute 2016) automatically checks all possible split points of all predictors in order to maximize the LogWorth statistics (Klimberg and McCullough 2013).

For verification, the bootstrap forest approach was used to examine whether replications by randomly resampling from the data set would yield the same conclusion. As the name implies, a bootstrap forest is a collection of many partition trees. The result of a single analysis could be capitalization on chance; to alleviate this problem, the bootstrap forest algorithm reruns the same analysis by randomly selecting observations from the same data. The converged result of a bootstrap forest can be viewed as a vote count of the importance of all the potential predictors (Breiman 2001, 2004).

The authors realized that oversampling and undersampling occurred in transnational data sets. To counteract this problem, it is advisable that either sampling weights are put into the analysis or multi-level modeling is used. However, in this data set the sampling weights are available to the secular variables only. Although the data mining techniques described above do not include sampling weights or multi-level analysis, there are other ways to alleviate the problem. To avoid bias in standard errors and overfitting, in each partition tree about 30% of the data were held back for cross-validation. Additionally, there are two other remedies for the overfitting problem: Use small trees and increase the number of trees (Fadaia 2013). Both approaches are intuitive. Increasing the size of the forest (the number of trees) creates more opportunities for machine learning and thus an overfitted result that capitalized on chance can be avoided. Furthermore, when the size of each tree is minimized, it means variables that have trivial effects on the dependent variable cannot be considered.

To further detect data patterns, regression analysis and hierarchical clustering were applied to the summarized data set by country. There are several reasons for transitioning from individual-level data to country-level data. One rationale is data reduction for clarity. With too many observations, data visualization is difficult due to overplotting. Second, when the data were reduced to 60 observations (nations) only, regression analysis was feasible. As mentioned before, using a regression model with 74.042 observations would result in excessive statistical power and probably type I error. In regression modeling the sampling weights of secular values were also taken into account. Lastly, the US sample (n =2232) was analyzed by the partition tree approach because some variables were specific to the US sample only, and its result was compared against the overall result yielded from the entire data set. Additionally, data visualization, such as quantile rendering, was employed to verify this relationship. The authors realized that summary-level analysis is subject to the ecological fallacy, which entails the misperception that relationships found in groups are also true for individuals (Freedman 1999). If a major discrepancy is found between the individual-level and summary-level analyses, probably there is an ecological fallacy. Nevertheless, if the two results are in alignment, then the conclusion is further verified by this triangulation.

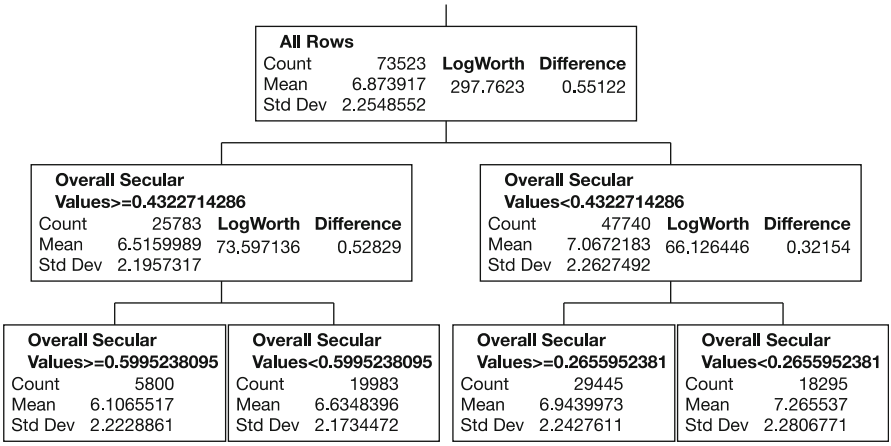

Fig. 1 Recursive partition tree result

Results

Figure 1 shows the partition tree result using 73,523 observations (519 missing values). Although 17 independent variables were examined by the partition tree algorithm, only one variable was considered important: secular values. Even when the analysis tried to push the program to further partition the data, the same variable "secular values" kept recurring, meaning that no more variables could be added to enhance the model. Figure 1 includes the redundant nodes to illustrate that the model had been saturated with one predictor. Only the first partition needs to be interpreted. The partition tree showed that if the secular value score was equal to or more than 0.432, then the expected mean of satisfaction was 6.51. If this value is less than 0.432, then the satisfaction mean score was 7.07.

A bootstrap forest was generated in order to verify whether secular values were the only predictor that possessed the highest predictive and explanatory power. In each tree, about 70% of the data set was selected for training and the remaining 30% was reversed for validation. Initially 500 trees were requested, but the algorithm stopped at 156 due to saturation. (Additional trees would not yield additional information.) The sum of squares (SS) and the portion of contribution to the model of the variables were rank-ordered and are given in Table 2. The SS of secular values is 4169.57, and obviously, it substantively outweighs the second variable (active/inactive membership: church or religious organization, SS=1013.64) and all other variables. In terms of portion, overall secular index contributes 44.28%. However, the portions of all other variables range from 0.0040 to 0.1076 only.

Next, the entire data set was summarized by country (n=60) for pattern recognition. The overall analysis with all individual observations by the partition tree approach suggested that only secular values were the most important predictor of satisfaction. Therefore, the country-level analysis focused on this predictor only. In the country-level data set, there are only 60 observations. When the data of the subscales of secular values, defiance, disbelief, relativism and skepticism were placed into the scatterplot matrix, four outliers were identified: China, India, Japan and Qatar. Nonetheless, the scatterplots showed that the data pattern met the requirement of linearity. After excluding these outliers, it was found that only disbelief and defiance had a significantly positive correlation (r=0.76, p<0.0001). In

Table 2 Bootstrap forest result showing the sum of squares and the portion of contribution

Variable	SS	Portion
Overall secular values	4169.57	0.4428
Active/inactive membership: church or religious organization	1013.64	0.1076
How often do you attend religious services	988.66	0.1050
How often to you pray	704.16	0.0748
How important is God in your life	669.41	0.0711
Meaning of religion: to make sense of life after death versus to make sense of life in this world	494.75	0.0525
Religion is important	399.86	0.0425
Confidence in church	353.67	0.0376
Meaning of religion: to follow religious norms and ceremonies versus to do good to other people	226.56	0.0241
Believe in hell	191.03	0.0203
Religious person	167.39	0.0178
Believe in God	37.94	0.0040

other words, each subscale could potentially contribute unique explanatory power to satisfaction. Table 3 shows the descriptive statistics, the p values and the confidence intervals of the correlation coefficients.

Multiple regression modeling did not yield any significant predictor of satisfaction (see Table 4). Interestingly, when the composite score of secular values with sampling weights was used for regressing against satisfaction, a different picture emerged.

The scatterplot (see Fig. 2a) of the mean value of satisfaction and the mean value of overall secular values shows that Kuwait is a bivariate outlier. After the outlier was removed, the scatterplot shows an associational pattern between satisfaction and secularization in a negative way even though the p value is not significant ($b=-3.25$, SE=1.96, $t=-1.65$, $p=0.1045$) (see Fig. 2b).

Paul (2014a) put European countries into one cluster due to their commonalities in secularization and objective well-being indices. However, when satisfaction and secularization were used as input for hierarchical clustering, the picture was no longer clear-cut (Fig. 3). In a dendrogram, similar nodes are next to each other based on their common attributes. For example, at the top of Fig. 3 Algeria and Estonia have similar response

Table 3 Correlation and confidence intervals of the subscales of secular values

Variable pair	Pearson' r	Lower 95%	Upper 95%	p
Disbelief and defiance	0.7614	0.6201	0.8549	**<.0001**
Relativism and defiance	0.1716	−0.1008	0.4200	0.2148
Relativism and disbelief	−0.0787	−0.3370	0.1906	0.5678
Skepticism and defiance	−0.0358	−0.3007	0.2342	0.7971
Skepticism and disbelief	−0.0554	−0.3161	0.2130	0.6878
Skepticism and relativism	0.0471	−0.2210	0.3085	0.7328

Bold p value indicates significance at the alpha level of .01

Table 4 Regression analysis results

Variable	Slope	SE	*t* ratio	*p*
Defiance	−0.8182	1.8650	−0.44	0.6628
Disbelief	1.2799	0.9469	1.35	0.1827
Relativism	−0.5363	0.9492	−0.57	0.5746
Skepticism	0.3183	1.1790	0.27	0.7883

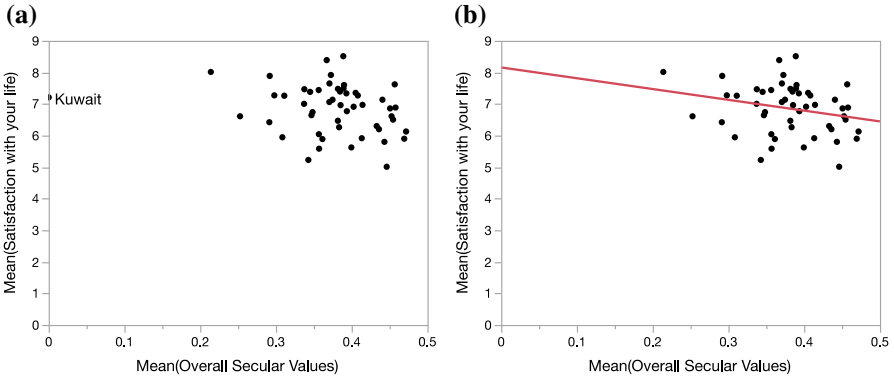

Fig. 2 **a** Scatterplot showing Kuwait as a bivariate outlier. **b** Scatterplot without Kuwait showing a negative slope

patterns to satisfaction and secularization, and therefore, these countries are paired into a group. By the same token, South Korea and Lebanon form another group. These two clusters could be joined to form a bigger group in the next level. Hierarchical clustering is so named because clusters are nested and organized as a hierarchical tree. Interestingly, in this hierarchical tree, European nations were scattered all over the dendrogram; the USA, Trinidad/Tobago and Australia had close proximity; and South Korea, Lebanon, Estonia and Algeria were grouped together. This implies that secularization and satisfaction could vary from country to country regardless of geographical location, culture or development status.

Interestingly, the partition tree result of the US data alone (n=2216, 16 missing values) is similar to that of the whole world data (see Fig. 4). As mentioned in "Data Source and Variables" section, some variables were reserved for the US sample only. Even though additional variables were used in this analysis, the US sample model was saturated with just one variable: secular values. Like the overall result with all nations, secular values were also inversely correlated with satisfaction in the USA. Specifically, if the secular value score is less than 0.403, then the average satisfaction is 7.758. If the score is equal to or less than 0.403, then the mean of satisfaction is 6.858.

Conclusion

Previous studies have suggested a positive relationship between secularization and well-being, as well as a geographical, cultural and development pattern primarily based on data gathered in Europe and the USA. Further review of the literature showed that the relationship between religiosity/secularization and subjective well-being is not necessarily

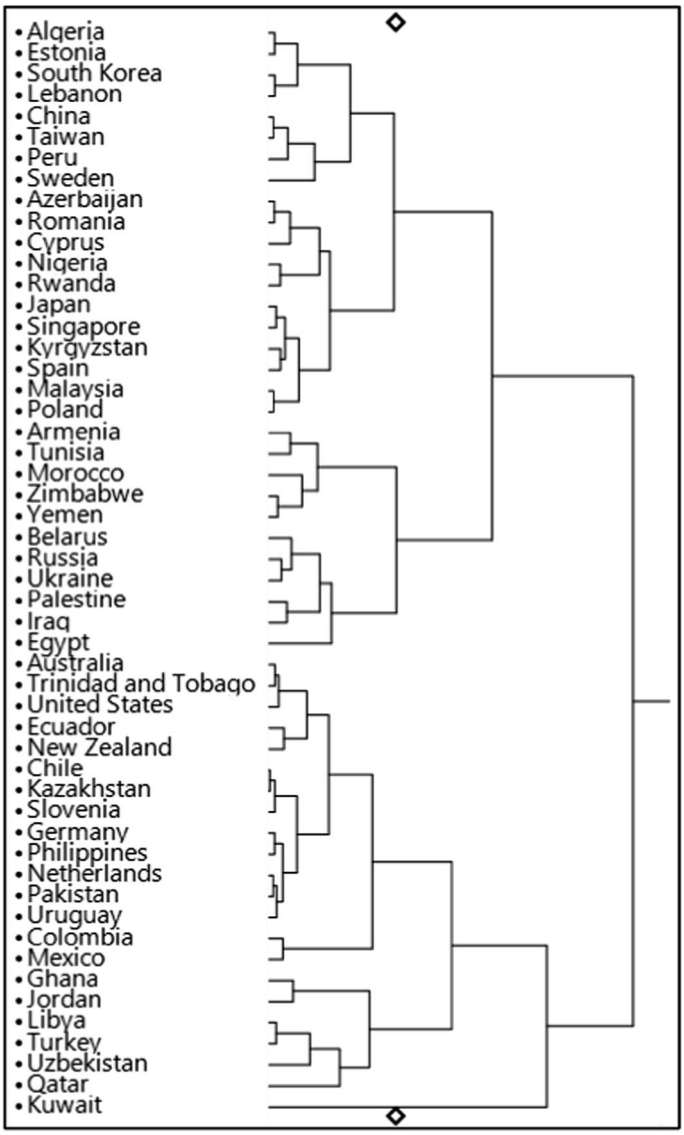

Fig. 3 Dendrogram of hierarchical clustering of nations by satisfaction and secularization

correlated with an individual's objective well-being, and that it varies according to cultural, social and individual factors. The literature review preceding the present study demonstrates that both positive and negative correlations have been found, making necessary the insight of the present, global analysis. By expanding the scope of analysis to the whole world, the current study suggested that existing perspectives might be oversimplified.

The findings that secular values were inversely correlated with satisfaction in the USA, as well as other findings in this study, are revealing. On a worldwide scale, the WVS data showed that there was an inverse relationship between subjective satisfaction and secularization in both individual-level and summary-level analyses. Frankly speaking, the

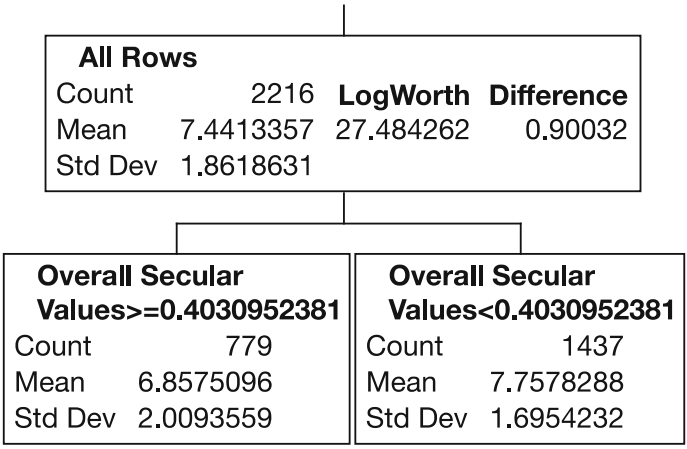

Fig. 4 Recursive partition tree of the US sample

current narrative of secularization might be overly simplistic. Welzel (2013) attributed modernization to secularization and in particular proposed that tolerating diverse views and practicing freedoms enhance life quality and well-being in modernized and pluralistic societies. Hence, people shift the moral foundation from religious, sacred authority to the secular one by embracing openness, resulting in a growing sense of security. However, using the archival data based on Welzel's framework, these authors became skeptical of the alleged association between secularization and well-being, at least the aspect of life satisfaction.

Perhaps Alexis de Tocqueville (1840/2002) could shed light on the inverse relationship between secular values and life satisfaction. In the seminal classic Democracy in America, Tocqueville observed that unlike Europe, the US de-emphasized aristocracy. Traditionally Americans believe that personal efforts, rather than nobility or the position in the social class, is the key to success and life satisfaction. The notions of equality and personal freedom granted by God are intertwined with democratic mores. In other words, democratic values and religious values go hand in hand. This may support a theory that discontent is expected when the fabric of the American culture is destabilized by secularization. Nonetheless, this conjecture awaits further investigation.

There are limitations to the present study. First, the scale of the outcome variable presented a challenge. Using a non-interval outcome variable violates the parametric assumption, but using ordinal logistic regression analysis with a 10-point scale produced an unstable model. Thus, the OLS regression models' results should be interpreted with this limitation in mind. Second, the OLS regression models were computed with summarized data, while the original data contain both normal and non-normal distributions. The authors considered log transformation, but the result would be more difficult to interpret. Hence, OLS regression models were retained, but the results should be interpreted with caution. Further, self-reported satisfaction, needless to say, is subject to several potential measurement errors. The participants might not be truthful, and the answer highly depends on the moment they responded to the survey. For example, the answers could be tied to whether one got a job promotion or lost the house due to foreclosure. More importantly, without qualitative data (e.g., open-ended answers) we could not tell why people were satisfied/dissatisfied or the nature of their happiness. One could be delusional but still very

satisfied with life (Bénézech 2011), and thus, some critics (e.g., Coyne 2015) charged that religious people are living under false ideas. Hence, the results of this analysis are simply fact-finding in nature and by no means imply any endorsement of religiosity or downplay of the value of secularization. Nevertheless, compared with classical procedures such as regression modeling, this data mining and data visualization approach can yield a more stable and optimal model. Further investigation on this topic utilizing data mining and data visualization is recommended.

Compliance with Ethical Standards

Conflict of interest The authors did not receive any funding for this project, and thus, we declare that we have no conflict of interest with any party.

Ethical Approval This article is based on archival data and does not contain any studies with human participants performed by any of the authors.

References

Bénézech, M. (2011). Pathological happiness: On pleasant or joyous hallucinatory states and delusions. *Annales MÕdico Psychologiques, 169,* 209–214.

Berger, P. (2014). *The many altars of modernity: Toward a paradigm for religion in a pluralist age*. Berlin: De Gruyter Mouton.

Berger, P., Davie, G., & Fokas, E. (2008). *Religious America, secular Europe? A theme and variations*. Burlington, VT: Ashgate.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York, NY: Routledge.

Bonini, A. N. (2008). Cross-national variation in individual life satisfaction: Effects of national wealth, human development, and environmental conditions. *Social Indicators Research, 87,* 223–236. https://doi.org/10.1007/s11205-007-9167-6.

Bowland, S., Edmond, T., & Fallot, R. D. (2012). Evaluation of a spiritually focused intervention with older trauma survivors. *Social Work, 57*(1), 73–82. https://doi.org/10.1093/sw/swr001.

Breiman, L. (2001). Random forests. *Machine Learning, 45,* 5–32.

Breiman, L. (2004). *Consistency for a simple model of random forests* (Technical report 670), UC Berkeley. https://www.stat.berkeley.edu/ ~ breiman/RandomForests/consistencyRFA.pdf. Accessed 26 Sept 2018.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth International Group.

Bruce, S. (2002). *God is dead: Secularization in the West*. New York, NY: Blackwell.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24,* 385–396.

Coyne, J. (2015). *Faith versus fact: Why science and religion are incompatible*. New York, NY: Viking.

de Tocqueville, A. (1840/2002). *Democracy in America*. Chicano, IL: University of Chicago Press.

de Waal, F. (2014). *The bonobo and the atheist*. New York, NY: W. W. Norton & Company.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75.

Eberstadt, M. (2013). *How the West really lost God: A new theory of secularization*. West Conshohocken, PA: Templeton Press.

Fadaia, B. (2013). Random forest on grouped data. *Cross validation*. https://stats.stackexchange.com/questions/19803/random-forest-on-grouped-data. Accessed 25 Sept 2018.

Falk, C. F., & Cai, L. (2016). Maximum marginal likelihood estimation of a monotonic polynomial generalized partial credit model with applications to multiple group analysis. *Psychometrika, 81,* 434–460. https://doi.org/10.1007/s11336-014-9428-7.

Francis, L., & Stubbs, M. (1987). Measuring attitudes towards Christianity: From childhood into adulthood. *Personality and Individual Differences, 8,* 741–743.

Freedman, D. (1999). *Ecological inference and the ecological fallacy* (Technical report 549), UC Berkeley. http://statistics.berkeley.edu/sites/default/files/tech-reports/549.pdf. Accessed 27 Sept 2018.

Galen, L. (2015). Atheism, wellbeing, and the wager: Why not believing in God (with others) is good for you. *Science, Religion and Culture, 2*(3), 54–69.
Gallup, Inc. (2014). *Gallup global wellbeing*. https://news.gallup.com/poll/126965/gallup-global-wellbeing.aspx. Accessed 15 Dec 2016.
Goldberg, D. (1992). *General Health Questionnaire (GHQ-12)*. Windsor: NFER-NELSON.
Green, M., & Elliott, M. (2010). Religion, health, and psychological well-being. *Journal of Religion and Health, 49,* 149–163.
Grzegorzewski, P. (2012). Statistical tests for comparing Pareto charts. *Communications in Numerical Analysis, 2012,* 1–12. https://doi.org/10.5899/2012/cna-00110.
Hayward, R. D., Krause, N., Ironson, G., Hill, P. C., & Emmons, R. (2016). Health and well-being among the non-religious: Atheists, agnostics, and no preference compared with religious group members. *Journal of Religion and Health, 55*(3), 1024–1037.
Hill, P. C., & Pargament, K. I. (2008). Advances in the conceptualization and measurement of religion and spirituality: Implications for physical and mental health research. *Psychology of Religion and Spirituality, S(1),* 3–17. https://doi.org/10.1037/1941-1022.S.1.3.
Inglehart, R., & Welzel, C. (2010). Changing mass priorities: The link between modernization and democracy. *Perspectives on Politics, 8,* 551–567. https://doi.org/10.1017/S1537592710001258.
Kistler, D., Thöni, C., & Welzel, C. (2017). Survey response and observed behavior: Emancipative and secular values predict prosocial behaviors. *Journal of Cross-Cultural Psychology, 48,* 461–489.
Klimberg, R., & McCullough, B. D. (2013). *Fundamentals of predictive analytics with JMP*. Cary, NC: SAS Institute.
Koenig, H. G., Meador, K. G., & Parkerson, G. (1997). Religion index for psychiatric research. *American Journal of Psychiatry, 154,* 885–886.
Li, L. M. W., & Bond, M. H. (2010). Does individual secularism promote life satisfaction? The moderating role of societal development. *Social Indicators Research, 99,* 443–453. https://doi.org/10.1007/s11205-010-9591.
Lincare, M. (2015). The sin of false precision: Too many rating-scale categories. *Rasch Measurement Transactions, 28*(2), 1463.
Lynch, J. (2012). *Religiosity: Investigating its affect on self-esteem, perceived stress, and general health in undergraduate students* (Bachelor's Thesis, Dublin Business School, Ireland). https://esource.dbs.ie/handle/10788/433. Accessed 15 Oct 2017.
New Economic Foundation. (2015). *Happy Planet Index*. http://www.happyplanetindex.org/.
Norris, P., & Inglehart, R. (2012). *Sacred and secular: Religion and politics worldwide*. Cambridge: Cambridge University Press.
Paul, G. (2014a). The health of nations. *Skeptic, 19*(3), 10–16.
Paul, S. (2014b). *The science of religion*. http://www.gspaulscienceofreligion.com/. Accessed 15 Dec 2016.
Pollack, D. (2008). Religious change in Europe: Theoretical considerations and empirical findings. *Social Compass, 55*(2), 168–186.
Rosenburg, M. (1965). *Society and the adolescent self image*. Princeton, NJ: Princeton University Press.
Ryff, C. D. (2014). Psychological well-being revisited: Advances in science and practice. *Psychotherapy and Psychosomatics, 83,* 10–28. https://doi.org/10.1159/000353263.
Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69,* 719–727. https://doi.org/10.1037/0022-3514.69.4.719.
Sabatier, C., Mayer, B., Friedlmeier, M., Lubiewska, K., & Trommsdorff, G. (2011). Religiosity, family orientation, and life satisfaction of adolescents in four countries. *Journal of Cross-Cultural Psychology, 42,* 1375–1393. https://doi.org/10.1177/0022022111412343.
SAS Institute. (2016). *JMP 13 [Computer Software]*. Cary, NC: SAS Institute.
Sinnewe, E., Kortt, M. A., & Dollery, B. (2015). Religion and life satisfaction: Evidence from Germany. *Social Indicators Research, 123,* 837–855. https://doi.org/10.1007/s11205-014-0763-y.
Speybroeck, N. (2012). Classification and regression trees. *International Journal of Public Health, 57,* 243–246.
Stark, R. (1999). Secularization, R.I.P. *Society of Religion, 60,* 249–273.
United Nations Development Programme. (2015). *Human Development Index*. http://hdr.undp.org/en/content/human-development-index-hdi. Accessed 15 Dec 2016.
Vanitha, A., & Niraimathi, S. (2013). Study on decision tree: Competent data classification. *International Journal of Computer Science and Mobile Computing, 2,* 365–370.
Welzel, C. (2013). *Freedom rising: Human empowerment and the quest for emancipation*. New York, NY: Cambridge University Press.
Welzel, C., & Delhey, J. (2015). Generalizing trust: The benign force of emancipation. *Journal of Cross-Cultural Psychology, 46,* 875–896.

Wilson, B. (1969). *Religion in secular society*. Harmondsworth: Pelican.
Wilson, B. (1976a). *Contemporary transformations of religion*. Oxford: Oxford University Press.
Wilson, B. (1976b). Aspects of secularization in the West. *Japanese Journal of Religious Studies, 3,* 259–276.
Wilson, B. (1982). *Religion in sociological perspective*. Oxford: Oxford University Press.
Wilson, B. (1994). If religion declines, will the social order survive? *Onze Alma Mater, 48,* 85–96.
World Values Survey Association. (2014). *World values survey*. http://www.worldvaluessurvey.org/wvs.jsp. Accessed 15 Dec 2016.
Yu, C. H. (2010). Exploratory data analysis in the context of data mining and resampling. *International Journal of Psychological Research, 3*(1), 9–22.
Yu, C. H., Reimer, D., Lee, A., Snijder, J., & Lee, H. S. (2017). A triangulated and exploratory study of the relationships between secularization, religiosity, and social wellbeing. *Social Indicators Research, 131,* 1103–1119. https://doi.org/10.1007/s11205-016-1290-9.
Zuckerman, P. (2008). *Society without God*. New York, NY: New York University Press.
Zuckerman, P. (2012). *Faith no more: Why people reject religion*. New York, NY: Oxford University Press.
Zuckerman, P. (2014). *Living the secular life: New answers to old questions*. New York, NY: Penguin Press.
Zuckerman, P. (2015). *Living the secular life: New answers to old questions*. Lecture delivered at Skeptics Society, Pasadena, CA.
Zullig, K. J., Ward, R. M., & Horn, T. (2006). The association between perceived spirituality, religiosity, and life satisfaction: The mediating role of self-rated health. *Social Indicators Research, 79,* 255–274. https://doi.org/10.1007/s11205-005-4127-5.