Religious discrimination in American public schooling: A large-scale audit experiment to assess bias among PK-12 principals

April 20, 2018

To what extent does religious discrimination occur in public schools across the United States? We examine this question by conducting a randomized audit study of prospective school enrollment. We emailed the principals of more than 45,000 PK-12 schools in 33 U.S. states and asked for a meeting, randomly assigning the religious affiliation/non-affiliation of the family as well as a signal for the intensity of their beliefs. In line with our theoretical expectations, we found substantial evidence that public school principals discriminate against Muslims and atheists but also against Protestants and Catholics who inquire about the compatibility of the school with their religious beliefs or request religious accommodation.

To what extent does religious discrimination occur in public schools across the United States?[1] This is an important question not only because such discrimination is illegal but also because the role of religion in American society is changing.[2] The United States stands out among advanced democracies not only for its relatively high level of religiosity but also for its religious diversity. In striking contrast to the citizens of many other wealthy democracies, an overwhelming majority of Americans continues to profess religious belief. At the same time, in recent decades the religious landscape has been dramatically reshaped as mainline groups have declined, religious diversity has increased, and a rapidly growing share of Americans identify themselves as nonbelievers (Baker and Smith 2015; Putnam and Campbell 2010; Sherkat 2014).

These changes raise questions concerning the toleration of religious differences, social integration, and the ability of public officials to observe the civil rights of diverse American families. Public education is a particularly sensitive and important arena in which to observe the consequences of expanding religious diversity. Because students of different religious worldviews are equally entitled to services and obliged to interact, it is a vital domain for the exploration of questions of discrimination (Justice and Mcleod 2016; Reese 2011).

The most prominent and controversial markers of the changing American religious landscape are the growing shares of Americans identifying as Muslims or nonbelievers. The Pew Research Center (Pew 2015) found that the share of the U.S. adult population identifying as Muslim doubled between 2007 and 2014 (albeit from only 0.4% to 0.9%), mostly as a result of immigration (Read 2008). Although the share of Americans that profess Islam as their religion is small, this group has become a highly visible and controversial minority (Peek 2010). As the median age of adult American Muslims is only 33 years (the U.S. median is 46 years), the enrollment of Muslim students in public schools can be expected to continue to increase substantially.

In a reversal of the long-term trend which had previously "churched" American society (Finke and Stark 1992), the share of American nonbelievers has increased dramatically over

the last two decades (Sherkat 2014). In 2014, Pew (2015) found that religious "nones" (a category that includes the unaffiliated, atheists, and agnostics) comprise about 23% of the adult U.S. population, a share that has increased from 16% since 2007. The sharp increase in "nones" is the result not only of disaffiliation from many Christian denominations but also the wholesale avoidance of religion by younger people—fully 35% of Millennials report no religious preference. Of course, not all "nones" are non-believers (Ammerman 2013; Hout and Fischer 2002; Marler and Kirk 2002), but the share of American adults that identifies as atheist (the corresponding share of agnostics is in parentheses) has likewise nearly doubled from 1.6% (2.4%) to 3.1% (4.0%) during the 2007 to 2014 period. Because the median age of adult atheists is just 34 years (Pew 2015), the public school enrollment of students from households of committed non-believers can also be expected to increase in coming decades.

Although religious discrimination is a rising concern, most of the existing research on diversity in American public education focuses on ethnic and racial disparities in performance and enrollment (see, e.g., Banks and McGee Banks 2010; Kao and Thompson 2003; Roscigno 1998, 2000; Steele and Aronson 2005; Johnson, Crosnoe, and Elder 2001; Warikoo et al. 2016). Even though students of all religious creeds enjoy formal protection from harassment and equal rights to school admission, in practice, the religious liberties of students may be frequently violated (Lippy 2006; Peek 2010). There are many reported instances of religious discrimination (Cragun et al. 2012; Hammer et al. 2012; Manning 2015). In 2004, for example, the Department of Justice sued a school for prohibiting a Muslim girl from wearing a headscarf (Hearn and United States v. Muskogee Public School District). In 2007, a high school student was kicked off the women's basketball team for refusing to take part in the Lord's Prayer (AHA, N.D.). In 2012, a high school student was subject to harassment after asking that a prayer banner be removed from a place of prominent display within a public school (Ahlquist v. Cranston).

Existing survey research suggests that hostile attitudes toward newcomer religious minorities and non-believers are widespread in American society (Cragun et al. 2013; Edgell,

Gerteis, and Hartmann 2006). For example, experimental research on religious biases in hiring indicates that religious discrimination exists in the labor market (Wallace, Wright, and Hyde 2014; Wright et al. 2015). What is missing from the literature is research on religious discrimination in public schooling. We simply do not know whether reported cases of discrimination represent exceptional incidents or merely the tip of the iceberg. Given the lack of systematic, rigorous evidence it is impossible to gauge the extent of religious bias in American schools.

In order to address this gap in the literature, we conducted a large-scale audit experiment of prospective school enrollment in which we emailed the principals of more than 45,000 public PK-12 schools in 33 U.S. states. In our emails, we claimed to be a parent who is considering enrolling his or her child in that school and asked for a meeting with the principal. We randomly assigned the religious affiliation/non-affiliation of the family (no information given, Protestant, Catholic, Muslim, or atheist) and the intensity of the religious signal (low [identification], medium [compatibility inquiry], high [accommodation request]). We then observed whether principals replied to our email. Compared to baseline emails, we found substantively high levels of discrimination against Protestant and Catholic parents who inquired about the compatibility of the school with their beliefs or asked for accommodation of their beliefs. We also found that Muslim and atheist parents are discriminated against for merely revealing their beliefs in the signature part of their emails. Finally, discrimination against Muslim and atheist parents increases dramatically if these parents inquire about the compatibility of the school with their beliefs or ask for religious accommodation, with discrimination against atheists significantly larger than discrimination against Muslims.

I. Previous experimental research on discrimination

Over the last decade, social scientists have advanced the study of discrimination by using field experiments to address the well-known limitations of surveys and observational studies in demonstrating bias (Bertrand and Duflo 2016).[3] Audit experiments have examined if bias occurs in response to group-based identification on the basis of race, ethnicity, gender, or

sexual orientation (e.g., Blommaert, Coenders, and van Tubergen 2014; Butler 2014; Butler and Broockman 2011; Gaddis 2015; Neumark 2012; Pager and Shepherd 2008; Pedulla 2016). Experimental studies are also beginning to examine religious discrimination. For example, Wright et al. (2013) and Wallace, Wright, and Hyde (2014) found that U.S. job applicants expressing a religious identity were less likely than those who did not to receive a response from a potential employer, with minorities such as Muslims and atheists suffering the greatest bias and evangelical Christians and Jews suffering little or no discernible bias. The U.S. does not appear to be unique in this regard. In France, not only do Muslims have lower incomes than matched Christian households but a Muslim job candidate is about 2.5 times less likely to receive a job interview callback than a racially similar Christian counterpart (Adida, Laitin, and Valfort 2010). In Greece, Drydakis (2010) found that résumés from religious minority groups were less likely to receive interview offers than those from members of the established Orthodox Church.

Whereas experimental research on bias in market situations is widespread, research on bias among public officials is much more limited. A small but rapidly growing literature examines bias among *elected* public officials (e.g., Butler and Broockman 2011; Butler 2014). This strand of research, however, has largely ignored the question of whether *unelected* public officials exhibit bias. Moreover, studies involving both elected and appointed public officials primarily focus on partisan and racial discrimination (e.g., Einstein and Glick 2017; White, Nathan, and Faller 2015). To our knowledge, no experimental research exists on religious biases in either elected or unelected public officials.

II. Sources of religious discrimination in public education

Broadly speaking, over the past sixty years there have been two opposing trends affecting the relationship between religion and public education. Secularizing activists and federal courts have drawn a sharper line between religion and public education. At the same time, such efforts have been met by push-back, especially from conservative religious groups and politicians. As local public officials entrusted with educational management, principals are pulled into different directions by these opposing trends (Justice and Mcleod 2016;

Reese 2011). We propose three theoretical perspectives on the sources of potential religious discrimination by PK-12 principals, focusing on the influence of secularism, Judeo-Christian nationalism, and civil religion.

A. Secularism as a basis for anti-religious bias

Secularization theory proposes that modernization propels the decline of religion at the level of institutions, attitudes, and beliefs. Modernization provides individuals with moral autonomy, opportunity, and personal security, initiating a "culture shift" away from religious traditions to post-traditional values and lifestyles (Norris and Inglehart 2004). At the same time, education and social diversity erode religion's plausibility, intensity, and authority, leading to a widely shared preference that religion be left to the private realm (Berger 1967; Bruce 2001).

Although the U.S. was long considered an exception to the secularization thesis, recent developments suggest that secularization processes are unfolding and may be accelerating (Putnam and Campbell 2010; Sherkat 2014; Voas and Chaves 2016). An important factor driving contemporary secularization is political conflict over values and religious issues, which is leading increasing shares of moderates and liberals to eschew organized religion and to regard it as damaging and divisive (Baker and Smith 2015; Hout and Fischer 2014; Manning 2015). At the institutional level, secularizing processes have been evident in American public education for decades. Widespread challenges to public school curricula and alleged violations of constitutional norms surrounding church-state separation are persistent features of American education.

Clashes over the secular nature of the public school system, the limits of religious accommodation, and state support for religious activities are nothing new (Alexander and Alexander 2012; Justice and Mcleod 2016). Historically, Protestant churches exerted substantial influence on public schooling. Religious discrimination toward minority faith communities and newcomer religions was commonplace (Reese 2011). Smith (2003) has shown how "secularizing activists" intent on challenging Protestant custodianship altered the institutions of public life including the schools. Even though campaigns to extend church-state

separation faced substantial resistance from religious conservatives, they were remarkably successful in the domain of public schooling. Despite popular religiosity, public education secularized more rapidly and thoroughly than in many other advanced democracies. Secular values deeply influence the institutional culture of American public schooling (Maryl 2016).

The result is a tension between a procedural secularism, which guarantees that public institutions hold no religious preferences, and a programmatic secularism, which insists that the public sphere admit no religion (Williams 2015). This tension generates unease about addressing religion in school (Justice and Mcleod 2016; Hartmann et al. 2011: 330). School officials are trained to be zealous guardians of church-state boundaries and to embrace secular norms in public education. For example, Essex's (2016) widely-used *School law and the public schools: A practical guide for educational leaders* is very clear in this regard. It instructs principals and administrators that the law "compels public schools as state agencies to maintain a neutral position in their daily operations regarding religious matters" (17) and insists that they are legally obligated to refrain from endorsing religious symbols, devotions, and expressions at school and to avoid supporting students' religious activities (16–47).

Principals' commitment to programmatic secularism could lead to bias. A secular emphasis in educational training and professional norms could also motivate discrimination. Alternatively, discrimination could arise from a desire to avoid controversies and the anticipated costs of religious accommodation. In practice, whatever their private views, school administrators wish to avoid issues which create friction among students, between students and staff, with parents, and with the broader public (Bess and Goldman 2001). We posit that the mechanism producing bias is a moral judgment that religious families are prone to make illegitimate requests for special treatment. The result would be discrimination against parents who are affiliated with a religious group, especially if the parents explicitly inquire about the compatibility of the school with their beliefs or request religious accommodation.

H1: Bias toward the outwardly religious hypothesis: *Parents who reveal a religious*

affiliation when requesting a meeting with public school principals will experience more discrimination from principals than parents who are silent about their beliefs, regardless of the intensity with which the parents religious beliefs are communicated. Moreover, the strength of the discrimination against parents revealing a religious affiliation will increase if they inquire about the compatibility of the school with their beliefs or request religious accommodation.

B. Judeo-Christian nationalism and religious bias

Religion continues to play a central role in establishing the boundaries of American national identity (Bonikowski and DiMaggio 2016; Hartmann et al. 2011). Christianity is especially important in subjectively defining "legitimate" membership in the American nation (Gerteis 2011, 2014). In a representative sample of Americans, 65% of respondents reported that Christianity was a "fairly" or "very important" criterion for being considered "truly American" and nearly half (48%) said it was "very important" (Bonikowski and DiMaggio 2016: 955).

Nevertheless, Christian identity has expanded since the mid twentieth century, with intellectuals and politicians drawing heavily on explicitly "Judeo-Christian" religious discourses to construct the moral boundaries of America (see e.g., Neuhaus 1986). Historians have shown how denominationalism and immigrant-driven diversity suggested to many public elites a new conception of the political role of religion. The struggle against Fascism gave this project a special urgency, leading to a vision of national identity which included Protestants, Catholics, and Jews—three groups purportedly united by Judeo-Christian values into a new kind of multicultural nation (Carenen 2012; Sarna 2004). Over time, national political discourse reflected this broadened notion of religious values just as denominational groupings replaced sectarianism and ethnicity in discussions of collective identity. Explicit antisemitism went from mainstream politics to the margins of American life (Hartmann et al. 2005; Herberg 1983).

Judeo-Christian political ideology was particularly useful in helping to integrate generations of white immigrants from Southern and Eastern Europe. Although compared with

the sectarianism of the past, the Judeo-Christian formulation was inclusive, it was inclusive only up to a point (see e.g., Douthat 2013). With the acceleration of immigration from non-European countries that began in the late 1960s, the American religious landscape became far more diverse. In practice, newcomer religions have quickly adapted to American denominationalism (Berger 2007; Hirschman 2004; Olsen 2007). Nevertheless, the question for some Americans has become *which* religious groups belong and *which* religious voices should be heard in public life. From the 1990s onward, conservatives prominently reasserted claims about America as a Judeo-Christian nation. Such claims were used to bolster arguments about which religious groups should be politically influential and which could be integrated into the national community (Hartmann, Zhang, and Wischadt 2005; Wilcox and Robinson 2011). Since the 1970s, conservative Protestantism thrived, at least in part, because its leaders portrayed it as the embattled defender of "true" American values and mobilized to influence politics and culture (Lindsay 2007; Smith 2000; Sutton 2014).

The limits of inclusion are apparent in the unease of many Americans toward members of unusual religious groups and religious newcomers (Bonikowski and DiMaggio 2016; Edgell, Gerteis, and Hartmann 2006). Anti-Muslim discourse resonates with many Americans, particularly the more than 30% who identify with conservative Protestantism (Pew 2015) and the quarter of Americans who can be classified as ardent nationalists (Bonikowski and DiMaggio 2016). In the context of a Judeo-Christian understanding of national identity and moral belonging, Muslim Americans pose a special problem, particularly in the wake of the 9/11 attacks and the "war on terrorism," the ongoing Arab-Israeli conflict, and the rise of the Islamic State, which have amplified the belief among a large minority of Americans that Islam is at war with the U.S. and its values (Froese and Mencken 2009).

In their analysis of American nationalism, Bonikowski and DiMaggio (2016) find that those Americans they classify as "ardent" nationalists are overwhelmingly white conservative Protestants who are prone to exclude religious minorities from those they consider truly American. In recent General Social Surveys, about 65 percent of conservative Protes-

tants agreed with the notion that immigrants cannot become fully American (Sherkat 2014: 155). In 2016, Donald Trump, who received about 80% of the white conservative Protestant vote in the presidential elections, made the depiction of Muslims as outsiders and enemies a prominent theme in his campaign (Braunstein 2017). We posit that the mechanism producing bias against Muslims is moral judgment. In the context of the politicization of Islam as a supposed threat to American society and values, principals might be reluctant to assist Muslim families in enrolling in their schools either because of their own moral bias or because of what they consider to be prevailing community standards.

H2: Bias toward Muslims hypothesis: *Parents who reveal an affiliation with Islam when requesting a meeting with public school principals will experience more discrimination from principals than parents who reveal an affiliation with Protestantism or Catholicism, regardless of the intensity with which the parents religious beliefs are communicated. Moreover, the strength of the discrimination against parents revealing an affiliation with Islam will increase if they inquire about the compatibility of the school with their beliefs or request religious accommodation.*

C. Civil religion and bias against non-believers

In American society, religion has long been an important source of conceptions about political community and social belonging. Opinion research has shown that, even as they have become more accepting of ethnic, racial, gender, and cultural diversity generally, Americans across the political and racial spectrum remain notably hostile in their attitudes toward non-believers in roles of political leadership or as appropriate marriage partners for their children (Cragun et al. 2012; Edgell et al. 2016; Edgell, Gerteis, and Hartmann 2006).

Popular hostility and suspicion toward atheists is rooted in the historical evolution of American political culture. Religion has been important for a restless American civil society in which faith-based groups provide social attachments and serve secular needs such as charitable assistance, opportunities for dating and marriage, daycare, connections to potential business partners and employers, and the integration of immigrants into community

life (Sherkat and Ellison 1999; Hirschman 2004; Olson 2007). In spite of the growing disaffiliation of Americans, congregations of all kinds remain the most common form of civic membership in America (Putnam and Campbell 2010). The view that religion is important for society has persisted even as the salience of denominational boundaries has declined (Hout and Fischer 2002; Sherkat 2014; Wuthnow 1988).

In large part because of religious organizations' practical social importance and role in fostering American democracy, many Americans have come to understand the U.S. as a religious country (Noll 2002). Scholars of civil religion argue that this is envisioned not in terms of an established church or favoritism toward a particular denomination, but rather as a consensus about the importance of religion for society (Bellah 1967; Gorski 2011). American civil religion endorses religious pluralism and a broadly spiritual notion of national community, resulting in "a public religious dimension [that] is expressed in a set of beliefs, symbols, and rituals" (Bellah 1967: 4). The legacy of American civil religion is evident in comparative perspective. The U.S. is noteworthy among advanced democracies for the effectiveness with which religious interest groups influence public policy and the very high share of the population (about half) which believes that religious leaders *should* influence public policy (Gryzmala-Busse 2015; Author 2008).

Historically, civil religion has bolstered a bi-partisan strategy to seek political consensus around a vision of a diverse society united against "godless" Communism (Sutton 2014). Many influential religious leaders endorsed this vision of civil religion, ranging from the liberal theologian Reinhold Niebuhr to the evangelical preacher Reverend Billy Graham. America became for many the "new Israel," with Americans being the "chosen people" elected by Providence to safeguard freedom. Although Ronald Reagan was one of the most persuasive prophets of this vision, repeatedly evoking images of America as a God-given "city on a hill," he was hardly alone in his willingness to draw upon religious symbols and discourse to foster national unity. Religion has been a remarkably persistent feature of American political speech across both major political parties (Coe and Domke 2006).

The legacy of civil religion affects Americans' attitudes. Religious diversity is tolerated

but the explicit rejection of religion is intolerable to a majority of Americans (Edgell et al. 2016; Edgell, Gerteis, and Hartmann 2006). Edgell, Gerteis, and Hartmann (2006: 230) report that "[a]theists are at the top of the list of groups that Americans find problematic in both public and private life" (see also Cragun et al. 2012). Survey research finds that anti-atheist bias in the United States is "persistent, durable, and anchored in moral concern" (Edgell et al. 2016: 629). Recent experimental research reveals that people intuitively judge atheists as immoral (Gervais 2014) and regard them as lacking pro-social values (Simpson and Rios 2017). Consequently, the category of atheist may function as a marker signaling that the bearer is a "moral outsider" in American society, strongly associated with immorality and contempt for common values (Edgell, Gerteis, and Hartmann 2006: 227).

Distrust toward non-believers extends to attitudes about schools. Historically, Americans counted on public schools to be bulwarks against atheism (Reese 2011: 42). Even today, more than one third of Americans in recent General Social Surveys say that atheist teachers should be fired (Sherkat 2014: 159). Many Americans appear to believe that, by openly rejecting religion, atheists are rejecting the normative foundations of community and the broader civic good. We posit that the mechanism producing bias against atheists is moral judgment. If parents identify themselves as atheists they may invite suspicion from school principals who fear that atheists and their children would be immoral, ideologically strident, and likely to opt out of civil rituals (such as the Pledge of Allegiance). Principals might be reluctant to assist atheist families in enrolling in their schools either because of their own moral bias or because of what they consider to be prevailing community standards.

H3: Bias toward atheists hypothesis: *Parents who reveal their atheist beliefs when requesting a meeting with public school principals will experience more discrimination from principals than parents who reveal an affiliation with any religion, regardless of the intensity with which the parents beliefs are communicated. Moreover, the strength of the discrimination against parents revealing atheist beliefs will increase if they inquire about*

the compatibility of the school with their atheist beliefs or request the accommodation of their atheist beliefs.

III. Research design and data

We use an audit experiment to investigate religious discrimination by PK-12 principals. Our experimental sample consists of regular, operational, non-charter public PK-12 schools in 33 U.S. states. We included all states for which we were able to acquire principals' email addresses either by contacting state Departments of Education or by downloading contact information from the websites of those institutions. Within these 33 states, we dropped all schools with missing principal contact information. We also excluded schools that could not be uniquely matched to NCES (National Center for Education Statistics) data and schools with missing covariate data in the NCES or American Community Survey (ACS). Based on state and NCES data, we dropped inactive, private, charter, non-traditional, adult, and virtual schools as well as schools serving restricted populations such as schools for the blind and deaf and schools located on military bases. We also excluded schools with less than 100 students, schools that are majority American-Indian, and schools that offer pre-Kindergarten or Kindergarten as the highest grade. If several schools shared a principal we only kept one of the schools, chosen randomly. If several schools were located at the same physical address we only kept one of the schools, also chosen randomly.

Based on these selection criteria our sample size equaled 47,550 schools. When we conducted our audit experiment some of our emails could not be delivered due to misspelled or outdated principal email addresses. We dropped all schools with bounced emails from the experimental sample.[4] Our final sample size equals 45,710 schools.[5]

We observe a number of covariates drawn from the NCES (2013), ACS (2012), and the Religious Congregations & Membership Study (RCMS) (2010). From the NCES, we observe the share of Asian, Hispanic, Black, and White students at the school-level. We also observe the share of students eligible for free or reduced price lunches, the share of male students, the school size, and the pupil/teacher ratio. From the ACS, we observe the median household income, the share of adults holding bachelor degrees, and the share of

residents with income below the poverty line at the county level. We also observe county-level Republican vote shares in the 2012 presidential elections. From the RCMS, also at the county level, we have the rates of Black Protestant, evangelical Protestant, mainline Protestant, Catholic, Muslim, and total adherents per 1,000 capita.[6]

The plots in Figures A1–A3 in the SI (supplementary information) compare our sample to the NCES population of 78,348 regular, non-charter public schools without missing NCES data in the 48 contiguous U.S. states. While our experimental sample is not truly a random sample from the NCES universe, it tracks the NCES population rather well in terms of observed covariates.

Our experiment contains four different treatments: *Parent's gender* (male/female), *child's gender* (male/female), *religious affiliation/non-affiliation* (no information given, Protestant, Catholic, Muslim, atheist), and the *intensity* with which religious/non-religious beliefs are communicated (low [identification], medium [compatibility inquiry], high [accommodation request]). We randomized parent's and child's gender to rule out the possibility that our causal inferences about religious affiliation/non-affiliation or intensity are driven by a particular gender or gender combination. We signal parent's and child's gender by using different names: Isaac and Rebekah Adam for the parents and Jonah and Sarah for the children. We chose these first and last names because they frequently appear in the Old Testament, an important religious text for both Christians and Muslims.[7] As each of these names is relatively common in the United States, atheists with these names are also not unusual. Given the large number of emails we had to send out we used eight different email accounts to contact principals; email account names also signal parents' gender by including either Isaac or Rebekah. We sent emails throughout a one-week period in April 2016; the order in which principals were contacted was randomized. The text of our emails is shown in Figure 1.

We include a Catholic treatment in our experiment to ensure that discrimination against Christians is not being driven by political hostility toward conservative Protestants. Catholicism is liturgically and theologically distinct from Protestantism and readily culturally

identifiable. As a religious group, contemporary Catholics are a good benchmark because they are ethnically diverse and close to the U.S. mean on many demographic characteristics including education, income, and political preferences. In terms of the American religious spectrum, Catholics, on average, identify as religious moderates (Sherkat 2014). They furnish a better reference category than "mainline" Protestants because liberal Protestants are, on average, similar to the unchurched in attitudes and values. Accordingly, a mainline Protestant identification as a religious signal would not be as resonant as Catholicism.

The literature on religious discrimination has often focused on the Jewish experience (Davidson and Pyle 2011). When designing our study, we decided to focus on Islam rather than Judaism as our non-Christian minority religion of interest.[8] We are not claiming that Jews do not experience religious discrimination. Rather, our reasoning was that survey evidence consistently shows that Americans are more favorable toward Jews than toward any other religious group, that discrimination is ebbing, antisemitism has declined, and stereotypes have faded (Rebhun 2016). Jews have become, by most educational, occupational, and political measures, a very successful mainstream religious group in the U.S. (Burstein 2007). Moreover, our design was predicated on separating religion from ethnicity as bases for bias. The signal communicated by Jewish identification can be ambiguous because Jews are often thought of as an ethnic group and Jewish ethnicity does not imply religiosity.

We signal religious identity and the intensity with which beliefs are held in the following way. The low intensity condition signals religious affiliation/non-affiliation only by including an email signature at the bottom of the email, in purple color. The email signature contains a modified version of a Richard Dawkins quote ("[. . .] teaches that life is precious and beautiful. We should live our lives to the fullest, to the end of our days."). This quote is sufficiently bland (and obscure) that it could be reasonably attributed to virtually any source. We substitute "Christianity," "Catholicism," "Islam," or "Atheism" into the quote, depending on the religious affiliation/non-affiliation treatment. We also change the purported author of the quote to Rev. Billy Graham, Pope Benedict, The Prophet Muhammad, or Richard Dawkins, again depending on the religious affiliation/non-affiliation treatment.

The medium intensity condition keeps the signature but adds the following sentence, which is designed to signal the desire for compatibility between the school and the beliefs of the family: "One of the reasons we would like to meet with you is that we are raising [Jonah/Sarah] to be a good [Christian/Catholic/Muslim/Atheist Humanist] and want to make sure that this would be possible at your school." The high intensity condition likewise keeps the signature but adds the following sentence, which is designed to signal a request for the accommodation of the family's religious beliefs: "One of the reasons we would like to meet with you is that we are raising [Jonah/Sarah] to be a good [Christian/Catholic/Muslim/Atheist Humanist] and want to protect [him/her] from anything that runs counter to our beliefs. We want to make sure that this would be possible at your school." The no information given condition only contains the gender treatments, but no signature and no additional text.[9]

Treatments were randomly assigned within blocks defined by state, shares of Asian, Hispanic, Black, and White students, the percentage of students eligible for free or reduced price lunches, median household income, the share of adults holding bachelor degrees, the share of residents with income below the poverty line, and Republican vote share in the 2012 presidential elections. Table A1 in the SI shows that our sample is well-balanced.[10] We also compute an omnibus randomization inference p-value that tests for joint balance across all 18 covariates. This p-value equals 0.90, confirming that the blocked randomization procedure was successful in balancing observables.

We sent a single email to each principal, with no follow-up in case of non-response. We then observed whether principals replied to our email within a 14 day window from the time the email was sent. Automatic replies such as out-of-office replies were discarded. As is common practice (e.g., Bertrand and Duflu 2016; Butler 2014), we use receipt of a non-automated reply email as our binary outcome variable.[11] We recognize of course that non-response results from many sources besides bias. For example, each principal's responsiveness is undoubtedly affected by factors such as his or her work load. That being said, we are interested in systematic differences across groups of principals exposed

to different emails. While we cannot interpret non-reply by any individual principal as a sign of discrimination, the presence of systematic differences in responsiveness between randomly assigned treatment groups *is* evidence of discrimination (Bertrand and Duflu 2016; Butler and Broockman 2011).

IV. Empirical results

Among our 45,710 subjects, 19,696 sent at least one non-automated reply email within 14 days, for a response rate of 43.1%. This response rate is in line with response rates from other internet audit experiments with elected and unelected public officials (Costa 2016).[12]

Table 1 shows results from a probit model. Because we are interested in the interaction between the religious affiliation/non-affiliation treatment and the intensity treatment, we include dummy variables representing all *combinations* of the religious affiliation/non-affiliation and intensity treatment levels in the model. The model also includes dummy variables for parent's and child's gender as well as fixed effects for the eight email accounts we used to send emails (coefficient estimates not shown). Robust standard errors are clustered at the school district level.

We use two plots to visualize the main empirical results of our experiment. Based on the probit estimates in Table 1, Figure 2 plots the estimated probability of receiving a reply email for male and female parents, male and female children, the baseline condition in which we do not provide any information about the religious affiliation/non-affiliation of the family, and the twelve possible religious affiliation/non-affiliation and intensity combinations.[13]

Figure 3 complements Figure 2 by displaying treatment effects (i.e., *differences* in probabilities), making it easier to judge both the statistical and substantive significance of the various treatment effect estimates.

While not of primary interest for this paper, we find that compared to female names, male names for both the parent and the child reduce the probability of receiving a reply email. For parent's gender, the estimated effect size is 1.6 percentage points; for child's gender, it equals 0.8 percentage points. These effect estimates are statistically significant

at the 0.05 and 0.10 levels, respectively.[14]

Turning to the religious affiliation/non-affiliation and intensity treatments and comparing to the baseline (no information) condition (while averaging over the gender factors), we find the following patterns. For the religious affiliation/non-affiliation treatment paired with the low intensity condition (identification through signature line), the effects of Protestant and Catholic affiliation are slightly positive but not statistically significant at the 0.10 level. An affiliation with Islam, on the other hand, even if signaled solely through the email signature (and not the text of the email itself), reduces the probability of reply by 4.6 percentage points, an effect that is highly statistically significant. The effect size for atheist email signatures is very similar, reducing the probability of reply by 4.7 percentage points compared to the baseline condition.

In the medium intensity condition, in which parents inquire about the compatibility of the school with their beliefs, the extent of discrimination increases for all four religious affiliations/non-affiliations. For Protestants, the estimated treatment effect is -5.4 percentage points; for Catholics, it equals -6.6 percentage points. Effect estimates are even larger in absolute value for Muslims (-8.7 percentage points) and atheists (-13.8 percentage points). All four effect estimates are highly statistically significant; all four are also more negative than the corresponding effect estimates in the low intensity condition ($\chi^2_4 = 129.97, p \approx 0$). Moreover, the effect for Muslims is significantly larger (in absolute value) than the effects for Protestants and Catholics ($\chi^2_2 = 6.80, p = 0.03$) and the effect for atheists is significantly larger than the effect for Muslims ($\chi^2_1 = 17.72, p \approx 0$).

Effect estimates for the high intensity condition are very similar to estimates for the medium intensity condition. We cannot reject the null hypothesis that effects in the medium and high intensity conditions are the same ($\chi^2_4 = 0.95, p = 0.92$).

Our results are both substantively and statistically identical if we additionally control for block fixed effects (Table A2 in the SI). Using linear probability models instead of probit also does not affect our results (Table A3 in the SI). The same is true when we control for the covariates listed in Table A1 to address possible covariate imbalances (Table A3).

Finally, our findings are also completely unaffected by dropping Massachusetts from the sample (Table **??**).

Our results provide qualified support for all three hypotheses. Consistent with H3, discrimination against atheists is greater than discrimination against any religious group including Muslims, at least in the medium and high intensity conditions. When religious identity is merely signaled through email signatures, we still find sizable discrimination against atheists, but it is not significantly larger than discrimination against Muslims (it is significantly larger than discrimination against Protestants and Catholics). As predicted by H3, we also find that discrimination against atheists increases as we move from the low intensity condition to the medium intensity condition. Unexpectedly, moving from medium intensity to high intensity did not further increase the extent of discrimination against atheists (or any other group). It is possible that differences in language between the medium and high intensity conditions were not sufficiently large to induce additional discrimination.[15] Alternatively, it is conceivable that the language we used for the medium intensity condition already implied a possible request for religious accommodation, so that principals worried about the costs of accommodation refrained from responding to these emails at the same rate as in the high intensity condition.

In line with H2, we find that principals discriminate against Muslims even in the low intensity condition, where the only difference between Muslim parents' emails and Protestant and Catholic parents' emails is the email signature. Also consistent with H2, discrimination increases in the medium intensity condition.[16] In all three intensity conditions, Muslim parents are less likely to receive a reply than Protestant and Catholic parents.

Finally, our results are also partially consistent with H1. While we fail to observe any discrimination against Protestant or Catholic parents in the low intensity condition, we do find that principals are significantly less likely to reply if parents inquire about the compatibility of the school with their beliefs or request accommodation of the family's religious beliefs.

A. Limitations

Audit studies and randomized experiments more generally are designed to identify and estimate the average causal effects of specific treatments—in our study, signals of religious affiliation/non-affiliation and intensity of belief embedded in parents' emails. Although we have posited plausible mechanisms that could explain the biases we find, our experiment cannot conclusively demonstrate that it is indeed these mechanisms which produce the biases we document. Nor does our audit study speak directly to the many other situations in which principals (or other public officials) might engage in religious discrimination. Our study shares these limitations with other audit studies and randomized experiments more generally (e.g., Bullock, Green, and Ha 2010). Despite these limitations, our paper makes an important contribution to the literature by documenting for the first time that significant religious discrimination takes place among PK-12 principals in American public schools.

V. Conclusion and Implications

In 2014, American public school enrollment crossed the "majority-minority" threshold—non-white children had become the majority of children enrolled in public schools. This transition has been accelerated by immigration. By 2050, the NCES expects that the share of school-age children with one or more immigrant parents will reach one third of all students (Maxwell 2014). These demographic trends will lead to greater religious diversity in American schools because of the greater share of nonbelievers among younger cohorts and because about one third of immigrants are affiliated with religious traditions outside the Judeo-Christian constellation.

In a rapidly changing society, PK-12 principals occupy a particularly challenging role as mediators between teachers and parents and between schools and the public. They are expected to ensure competence, access, and equity in public education. Their task is complicated by changes that unsettle the established moral consensus and strain the capacity of public schools to ensure equity and fair treatment. Routine bureaucratic behavior in public education must simultaneously respect the separation of church and state and safeguard

the protection of individuals' religious liberties. Discrimination on the basis of religious affiliation/non-affiliation could damage educational performance and attainment by undercutting school attachment and academic engagement (Johnson, Crosnoe, and Elder 2001) and by posing a substantial barrier to parental involvement in schooling (Turney and Kao 2009).

Because education is a primary factor in occupational and income mobility as well as formal socialization, the potential consequences of discrimination in this domain are far-reaching. As leading experts in education law observe, the U.S. is "a country that has developed an extraordinary reliance on public schools as a mechanism for social and economic justice and improvement" (Alexander and Alexander 2012: xxxvii). A prominent historian of education observes that "[i]n the early twenty-first century, America's schools remain central to most public debates over how to define and secure the good life for the nation's children" (Reese 2011: 8). Our evidence of discrimination by public school principals raises serious concerns about one of the fundamental institutions of modern democratic government and its capacity for integration and equity.

We expected bias against those outwardly identifying as religious because of the secular professional norms of public school principals and the anticipated costs and conflicts associated with religious accommodation. We found evidence for such bias when Protestant and Catholic parents inquire about the compatibility of the school with their beliefs or request religious accommodation. We also expected that Judeo-Christian values in American culture would result in discrimination against Muslims. We found this to be the case, even when religious affiliation with Islam was signaled merely through email signatures. Finally, based on the well-documented distrust of non-believers and the role of American civil religion, we expected discrimination against atheists. Our findings show that atheists are highly discriminated against when they explicitly inquire about the compatibility of the school with their non-beliefs or request accommodation of their non-beliefs. Discrimination against atheists is of the same magnitude as discrimination against Muslims when they signal their non-beliefs solely through their email signatures. Our findings are thus

consistent with prior research demonstrating that atheists are among the most distrusted groups in American society (Cragun et al. 2012; Edgell et al. 2016).

Although previous research has shown a prevalence of hostile and exclusive attitudes toward religious outsiders and demonstrated religious biases in hiring, to the best of our knowledge our paper reports the first experimental research examining religious discrimination in the American public school system. Our findings strongly suggest that stories about discrimination in public schools on the basis of religious identification/non-identification are not merely anecdotal but reflective of patterns of discrimination worthy of much more detailed investigation. Now that we have established that such discrimination is common when parents interact with school principals in the process of school enrollment, future research should make it a priority to investigate *when* and *why* principals engage in such discrimination and whether other public officials, both elected and unelected, also engage in religious discrimination.

Notes

[1]By religious discrimination we refer to differential treatment based on religious affiliation or non-affiliation.

[2]The Constitution of the United States prohibits the state from discriminating against individuals or groups based on their religious identification. Title IV and Title IX of the Civil Rights Act of 1964 prohibit discrimination against students based on their faith. Case law clarifies that students can religiously identify at school and can take part in religious activities of their own devising. State constitutions also generally either make public education a fundamental right or contain protection clauses or their equivalents that prohibit religious discrimination (Alexander and Alexander 2012: 46).

[3]In line with the literature (e.g., Butler 2014), we use the terms bias and discrimination interchangeably.

[4]This is unproblematic because invalid or outdated email addresses are orthogonal to treatment assignment by virtue of randomization.

[5]The following 33 states make up the experimental sample (with number of schools in parentheses): AL (851), AR (761), CA (5892), CO (832), DE (135), FL (1902), GA (1636), IA (864), ID (389), IL (2519), IN (1486), LA (821), MA (1314), MI (1979), MN (813), MO (1412), MS (702), NC (2027), ND (161), NE (571), NH (326), NJ (1759), NM (412), NY (2904), OH (1974), RI (189), SC (930), TN (1067), TX (4723), VA (1427), VT (179), WA (1336), WI (1417). One complication arose during our experiment. In Massachusetts, our emails coincided with a malware attack targeting public schools. At least one principal thus forwarded our email to Massachusetts state police, which contacted all Massachusetts public school principals warning them that our emails were probably spam. We chose to

keep Massachusetts in our sample, but the results are entirely unchanged if we drop all Massachusetts schools from the sample. We have included this set of results in the SI (supplementary information).

[6]Using Amelia II (Honaker, King, and Blackwell 2012), missing data in RCMS variables have been multiply imputed using the NCES and ACS variables listed above, the outcome variable, and an additional set of 18 ACS variables plausibly prognostic of religious adherence or missingness. All standard errors and statistical tests have been adjusted to account for multiple imputation as described in Little and Rubin 2002 and Schafer 1997.

[7]Of course, Muslims might be more likely to have Arabic versions of these names; "Jonah", for example, might be rendered as "Yunus." However, if we had used different names for different religious affiliations we would have conflated signals of religious affiliation with signals of race/ethnicity, a problem which has plagued research on discrimination against Muslims (Zainiddinov 2016).

[8]We designed our study well in advance of the 2016 presidential election and failed to foresee the reappearance of anti-Semitic tropes during the election campaign and, after the election, among the so-called alt-right. In hindsight, the inclusion of Jewish families would have been valuable.

[9]We used "atheist humanist" as opposed to merely "atheist" in our emails since atheism as such does not have any ethical content. It would have sounded odd if parents had announced their intention to raise children to be "good atheists." Moreover, we used "Christian" as opposed to "Protestant" in our emails since American Protestants typically refer to themselves as "Christian" and not as "Protestant." In order to ensure that respondents would recognize the Protestant treatment in the signature line we attributed it to Billy Graham, one of the most famous Protestant clergymen of the late 20th Century.

[10]We approximated exact randomization-based p-values using 10,000 randomly chosen blocked treatment assignments. The test statistic is the maximum KolmogorovSmirnov statistic across all two-way comparisons of treatment groups. The p-value is the fraction of test statistics at least as large as the test statistic in our sample.

[11]We randomly sampled and read 500 reply emails. In almost all of them, principals either asked for times that we would be able to meet or proposed times for a meeting. In a few emails, principals asked us to provide additional information such as our moving date or our child's grade level. In 8 emails, principals informed us that their schools did not offer school tours at the moment, typically because state testing was currently taking place. 17 reply emails were from former principals who suggested contacting the current principal, almost always either providing contact information for or cc'ing the current principal. After discarding automatic replies, we thus feel confident in treating the receipt of a reply email as indication of a principal's willingness to meet with us. We originally had planned to use supervised text analysis tools to analyze and code the content and tone of reply emails, but the reply emails proved to be too uniform in content and too succinct to make such an endeavor worthwhile. A typical reply was something like "Sure! When can you come in?" The major variation in responses is thus between getting a reply email and not getting a reply email.

[12]A small number of principals sent several emails, typically to update times they had mentioned in a previous email during which they would be available for a meeting.

[13]Probability estimates are simulated using the observed values approach throughout (Hanmer and Kalkan 2013). When displaying estimated probabilities for a given factor or factor combination we average over the remaining factors. For example, the probabilities of reply for male and female parents average over the child's gender, affiliation/non-affiliation, and intensity factors. The probabilities of reply for the twelve affiliation/non-affiliation and

intensity combinations average over the parent's and child's gender factors.

[14]There is no evidence of an interaction between parent's gender and child's gender: $\chi^2_1 = 0.19$, p-value $= 0.67$. There is also no evidence of interactions between the gender treatments and the religious affiliation/non-affiliation or intensity treatments: $\chi^2_{24} = 26.73$, p-value $= 0.32$. All Wald tests are two-sided.

[15]When designing the experiment we considered even more strident language for the high intensity condition but decided against it in order to safeguard the realism of our emails and the internal and external validity of our experiment.

[16]As with atheists, we see no further increase in discrimination in the high intensity condition.

VI. References

Adida, Claire L., David D. Laitin, and Marie-Anne Valfort. 2010. "Identifying barriers to Muslim integration in France." *Proceedings of the National Academy of Sciences* 107/52: 22384–22390.

Alexander, Kern and M. David Alexander. 2012. *American public school law.* 8th edition. Belmont, CA: Wadsworth.

Ammerman, Nancy T. 2013. "Spiritual but not religious: Beyond binary choices in the study of religion." *Journal for the Scientific Study of Religion* 52 (2): 258–278.

Baker, Joseph O. and Buster G. Smith. 2015. *American secularism: Cultural contours of nonreligious belief systems.* New York: New York University Press.

Banks, James A. and Cherry A. McGee Banks. 2010. *Multicultural education: Issues and perspectives.* New York: Wiley.

Bellah, Robert N. 1967. "Civil religion in America." *Daedalus* 96 (1): 1–21.

Berger, Peter L. 1967. *The sacred canopy.* New York: Doubleday.

Berger, Peter L. 2007. "Pluralism, protestantization, and the voluntary principle." In Thomas Banchoff (ed.). *Democracy and the new religious pluralism,* pp. 19–30. New York: Oxford University Press.

Berry, William D., Jacqueline H.R. DeMeritt, and Justin Esarey. 2015. "Bias and overconfidence in parametric models of interactive processes." *American Journal of Political Science* 60 (2): 521–539.

Bertrand, Marianne and Esther Duflo. 2016. "Field experiments on discrimination." NBER working paper.

Bess, James L. and Paul Goldman. 2001. "Leadership and ambiguity in universities and K-12 schools and the limits of contemporary leadership theory." *Leadership Quarterly* 12: 419–450.

Blommaert, Lieselotte, Marcel Coenders, and Frank van Tubergen. 2014. "Discrimination of Arabic-named applicants in the Netherlands: An internet-based field experiment examining different phases in online recruitment procedures." *Social Forces* 92 (3): 957–982.

Bonikowski, Bart and Paul DiMaggio. 2016. "Varieties of American popular nationalism." *American Sociological Review* 81 (5): 949–980.

Braunstein, Ruth. 2017. "Muslims as outsiders, enemies, and others: The 2016 Presidential election and the politics of religious exclusion." *American Journal of Cultural Sociology* 5 (3): 355–372.

Bruce, Steve. 2001. *God is dead: Secularization in the West.* Oxford: Blackwell's Press.

Burstein, Paul. 2007. "Jewish educational and economic success in the United States: A search for explanations." *Sociological Perspectives* 50 (2): 209–228.

Bullock, John G., Donald P. Green, and Shang E. Ha. 2010. "Yes, but what's the mechanism? (Don't expect an easy answer)." *Journal of Personality and Social Psychology* 98 (4): 550–558.

Butler, Daniel M. and David E. Broockman. 2011. "Do politicians racially discriminate against constituents? A field experiment on state legislators." *American Journal of Political Science* 55 (3): 463–477.

Butler, Daniel M. 2014. *Representing the advantaged: How politicians reinforce inequality.* New York: Cambridge University Press.

Carenen, Caitlyn. 2012. *The fervent embrace: Liberal protestants, evangelicals and Israel.* New York: NYU Press.

Chipman, Hugh A., Edward I. George, and Robert E. McCulloch. 2010. "BART: Bayesian additive regression trees." *Annals of Applied Statistics* 4 (1): 266–298.

Coe, Kevin and David Domke. 2006. "Petitioners or prophets? Presidential discourse, God, and the ascendancy of religious conservatives." *Journal of Communication* 56 (2): 309–330.

Costa, Mia. 2016. "How responsive are political elites? A meta-analysis of experiments on public officials." Working paper.

Cragun, Ryan T., Barry Kosmin, Ariela Keysar, Joseph H. Hammer, and Michael Nielsen. 2012. "On the receiving end: Discrimination toward the nonreligious in the United States." *Journal of Contemporary Religion* 27 (1): 105–127.

Davidson, James D. and Ralph E. Pyle. 2011. *Ranking faiths: Religious stratification in America.* Lanham, MA: Rowan and Littlefield.

Douthat, Ross. 2013. *Bad religion: How we became a nation of heretics.* New York: Free Press.

Drydakis, Nick. 2010. "Religious affiliation and employment bias in the labor market." *Journal for the Scientific Study of Religion* 49 (3): 477–493.

Edgell, Penny, Joseph Gerteis, and Douglas Hartmann. 2006. "Atheists as 'other': Moral boundaries and cultural membership in American society." *American Sociological Review* 71 (2): 211–234.

Edgell, Penny, Douglas Hartmann, Evan Stewart, and Joseph Gerteis. 2016. "Atheists and other cultural outsiders: Moral boundaries and the non-religious in the United States." *Social Forces* 95 (2): 607–638.

Einstein, Katherine Levine and David M. Glick. 2017. "Does race affect access to government services? An experiment exploring street-level bureaucrats and access to public housing." *American Journal of Political Science* 61 (1): 100–116.

Essex, Nathan L. 2016. *School law and the public schools: A practical guide for educational leaders.* Sixth edition. Pearson.

Finke, Roger and Rodney Stark. 1992. *The churching of America, 1776–1990: Winners and losers in our religious economy.* New Brunswick, NJ: Rutgers University Press.

Froese, Paul and F. Carson Mencken. 2009. "A U.S. Holy War? The effects of religion on Iraq War policy attitudes." *Social Science Quarterly* 90 (1): 103–116.

Gaddis, S. Michael. 2015. "Discrimination in the credential society: An audit study of race and college selectivity in the labor market." *Social Forces* 93 (4): 1451–1479.

Gerteis, Joseph. 2011. "Civil religion and the politics of belonging." In Julian Go (ed.). *Rethinking Obama.* Emerald Group Publishing Limited, pp. 215–223.

Gerteis, Joseph. 2014. "The social functions of religion in American political culture." In Douglass Hartmann and Christopher Uggen (eds.). *The social side of politics.* New York: Norton.

Gervais, Will. 2014. "Everything is permitted? People intuitively judge immorality as representative of atheists." *PLOS One* 9 (4) e92302: 1–9.

Gorski, Philip S. 2011. "Barack Obama and civil religion." In Julian Go (ed.). *Rethinking Obama.* Emerald Group Publishing Limited, pp. 179–214.

Grzymala-Busse, Anna. 2015. *Nations under God: How churches use moral authority to influence policy.* Princeton: Princeton University Press.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2017. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." Working paper.

Hammer, Joseph H., Ryan T. Cragun, Karen Hwang, and Jesse M. Smith. 2012. "Forms, frequency, and correlates of perceived anti-atheist discrimination." *Secularism and Nonreligion* 1 (1): 43–58.

Hanmer, Michael J. and Kerem Ozan Kalkan. 2013. "Behind the curve: Clarifying the best approach to calculating predicted probabilities and marginal effects from limited dependent variable models." *American Journal of Political Science* 57 (1): 263–277.

Hartmann, Douglas, Xuefeng Zhang, and William Wischstadt. 2005. "One (multicultural) nation under God? Uses and meanings of the term 'Judeo-Christian' in the American media." *Journal of Media and Religion* 4 (4): 207–234.

Hartmann, Douglas, Daniel Winchester, Penny Edgell, and Joseph Gerteis. 2011. "How Americans understand racial and religious differences: A test of parallel items from a national survey." *Sociological Quarterly* 52 (3): 323–345.

Herberg, Will. 1983. *Protestant-Catholic-Jew: An essay in American religious sociology.* Third ed. Chicago: University of Chicago Press.

Hill, Jennifer L. 2011. "Bayesian nonparametric modeling for causal inference." *Journal of Computational and Graphical Statistics* 20 (1): 217–240.

Hirschman, Charles. 2004. "The role of religion in the origins and adaptation of immigrant groups in the United States." *International Migration Review* 38 (3): 1206–1233.

Hout, Michael and Claude S. Fischer. 2002. "Why more Americans have no religious preference: Politics and generations." *American Sociological Review* 67 (2): 165–190.

Hout, Michael and Claude S. Fischer. 2014. "Explaining why more Americans have no religious preference: Political backlash and generational succession, 1987–2012." *Sociological Science* 1: 423–447.

Johnson, Monica Kirkpatrick, Robert Crosnoe and Glen H. Elder, Jr. 2001. "Students' attachment and academic engagement: The role of race and ethnicity." *Sociology of Education* 74 (4): 318–340.

Justice, Benjamin and Colin Mcleod. 2016. *Have a little faith: Religion, democracy, and the American public school.* Chicago: University of Chicago Press.

Kao, Grace and Jennifer S. Thompson. 2003. "Racial and ethnic stratification in educational achievement and attainment." *Annual Review of Sociology* 29: 417–42.

Lindsay, D. Michael. 2007. *Faith in the halls of power: How Evangelicals joined the American elite.* New York: Oxford University Press.

Lippy, Charles (ed). 2006. *Faith in America: Changes, challenges, new directions.* Westport, CT: Praeger.

Manning, Christel. 2015. *Losing our religion: How unaffiliated parents are raising their children.* New York: New York University Press.

Marler, Penny Long and Hadaway C. Kirk. 2002. "'Being religious' or 'being spiritual' in America: A zero-sum proposition?" *Journal for the Scientific Study of Religion* 41 (2): 289–300.

Maryl, Damon. 2016. *Secular conversions: Political institutions and religious education in the United States and Australia, 1800–2000.* New York: Cambridge University Press.

Maxwell, Lesli A. 2014. "U.S. schools become 'majority-minority'." *Education Week* 10/20/14.

National Center for Education Statistics (NCES). 2013. *Common core of data [data file and codebook].* Retrieved from https://nces.ed.gov/ccd/pubschuniv.asp.

Neuhaus, Richard John. 1986. *The naked public sphere: Religion and democracy in America.* Grand Rapids: Eerdmans.

Neumark, David. 2012. "Detecting discrimination in audit and correspondence studies." *Journal of Human Resources* 47 (4): 1128–1157.

Noll, Mark A. 2002. *America's God: From Jonathan Edwards to Abraham Lincoln.* New York: Oxford University Press.

Norris, Pippa and Ronald Inglehart. 2004. *Sacred and secular: Religion and politics worldwide.* Cambridge: Cambridge University Press.

Olson, Daniel V. A. 2007. "American religion: Will it heal or divide the nation?" *Contemporary Sociology* 36 (1): 8–11.

Pager, Devah and Hana Shepherd. 2008. "The sociology of discrimination: Racial discrimination in employment, housing, credit and consumer markets." *Annual Review of Sociology* 34: 181–209.

Pedulla, David S. 2016. "Penalized or protected? Gender and the consequences of nonstandard and mismatched employment histories." *American Sociological Review* 81 (2): 262–89.

Peek, Lori. 2010. *Behind the backlash: Muslim Americans after 9/11.* Philadelphia: Temple University Press.

Pew. 2015. *America's changing religious landscape.* Washington, D.C.: Pew Research Center.

Putnam, Robert D. and David E. Campbell. 2010. *American grace: How religion unites and divides us.* NY: Simon and Schuster.

Read, Jen'nan Ghazal. 2008. "Muslims in America." *Contexts.* Fall: 39–43.

Rebhun, Uzi. 2016. *Jews and the American religious landscape.* New York: Columbia University Press.

Reese, William J. 2011. *America's public schools: From the common school to 'No Child Left Behind'.* Baltimore: Johns Hopkins University Press.

Roscigno, Vincent J. 1998. "Race and the reproduction of educational disadvantage." *Social Forces* 76 (3): 1033–61.

Roscigno, Vincent J. 2000. "Family/school inequality and African-American/Hispanic achievement." *Social Problems* 47 (2): 266–90.

Sarna, Joseph. 2004. *American Judaism, a history.* New Haven, CT: Yale University Press.

Sherkat, Darren E. and Christopher G. Ellison. 1999. "Recent developments and current controversies in the sociology of religion." *Annual Review of Sociology* 25: 363–394.

Sherkat, Darren E. 2014. *Changing faith: The dynamics and consequences of Americans' shifting religious identities.* New York: New York University Press.

Simpson, Ain and Kimberly Rios. 2017. "The moral contents of anti-atheist prejudice (and why atheists should care about it)." *European Journal of Social Psychology* 47 (4): 501–8.

Smith, Christian. 2000. *Christian America? What Evangelicals really want.* Berkeley, CA: University of California Press.

Steele, Claude M. and Joshua Aronson. 2005. "Stereotypes and fragility of academic competence, motivation, and self-concept." In Andrew J. Elliott and Carol S. Dweck (eds.). *Handbook of competence and motivation.* New York: Guildford Press, pp. 436–56.

Sutton, Matthew Avery. 2014. *American apocalypse: A history of modern Evangelicalism.* Cambridge, MA: Harvard University Press.

Turney, Kristin and Grace Kao. 2009. "Barriers to school involvement: Are immigrant parents disadvantaged?" *Journal of Educational Research* 102 (4): 257–71.

Voas, David and Mark Chavez. 2016. "Is the United States a counterexample to the secularization thesis?" *American Journal of Sociology* 121 (5): 1517–56.

Wallace, Michael, Bradley R. E. Wright, and Allen Hyde. 2014. "Religious affiliation and hiring discrimination in the American South: A field experiment." *Social Currents* 1/12: 189–207.

Warikoo, Natasha, Stacey Sinclair, Jessica Fel, and Drew Jacoby-Senghor. 2016. "Examining racial bias in education: A new approach." *Educational Researcher* 45 (9): 508–514.

White, Ariel R., Noah L. Nathan, and Julie K. Faller. 2015. "What do I need to vote? Bureaucratic discretion and discrimination by local election officials." *American Political Science Review* 109 (1): 129–142.

Wilcox, Clyde and Caren Robinson. 2011. *Onward Christian soldiers? The religious right in American politics.* Boulder, CO: Westview Press.

Williams, Rowan. 2015. *Faith in the Public Sphere.* London: Bloomsbury.

Wright, Bradley R. E., Michael Wallace, John Bailey, and Allen Hyde. 2013. "Religious affiliation and hiring discrimination in New England." *Research on Social Stratification and Mobility* 34 (1): 111–126.

Wuthnow, Robert. 1988. *The restructuring of American religion: Society and faith since World War II.* Princeton: Princeton University Press.

Zainiddinov, Hakim. 2016. "Racial and ethnic differences in perceptions of discrimination among American Muslims." *Ethnic and Racial Studies* 39 (15): 2701–2721.

VII. Figures

Figure 1: Email to principals

Subject: School visit?

Dear principal,

Hello. My family and I will be moving into the area sometime this summer. Right now, we are deciding where exactly to move and are looking at schools for our [son/daughter], [Jonah/Sarah]. Before we pick a place to live, we would like to meet with you or a member of your staff and chat a bit about your school. Would that be possible?

[A] [One of the reasons we would like to meet with you is that we are raising [Jonah/Sarah] to be a good [Christian/Catholic/Muslim/Atheist Humanist] and want to make sure that this would be possible at your school.]

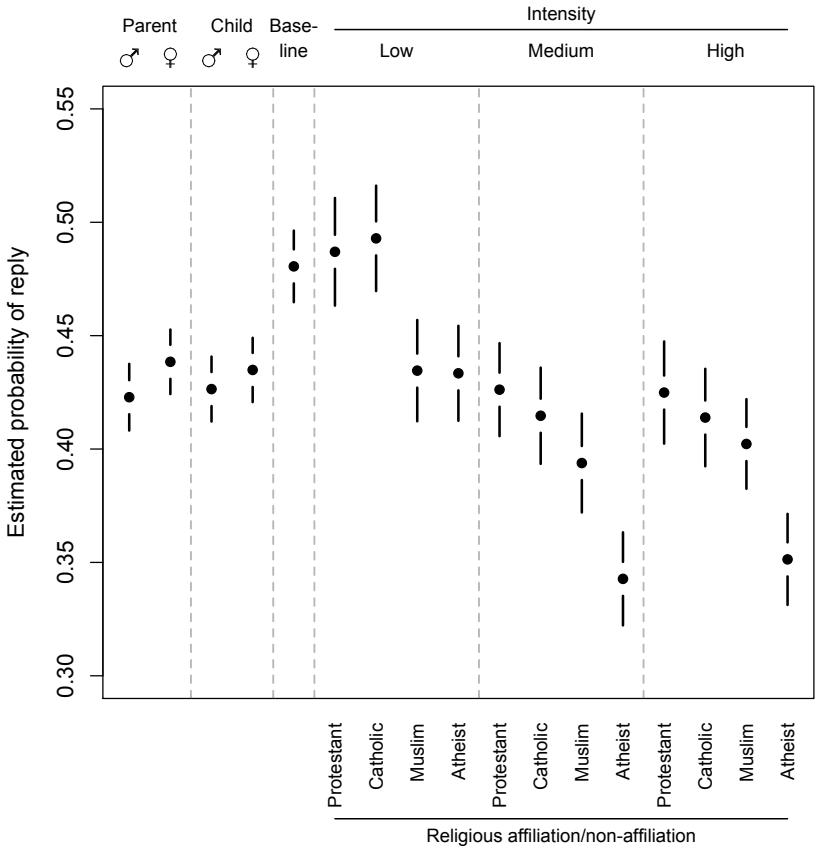
[B] [One of the reasons we would like to meet with you is that we are raising [Jonah/Sarah] to be a good [Christian/Catholic/Muslim/ Atheist Humanist] and want to protect [him/her] from anything that runs counter to our beliefs. We want to make sure that this would be possible at your school.]

Sincerely,
[Isaac Adam/Rebekah Adam]

[C] [Catholicism/Christianity/Islam/Atheism teaches that life is precious and beautiful. We should live our lives to the fullest, to the end of our days. - Pope Benedict/Rev. Billy Graham/The Prophet Muhammad/ Richard Dawkins]

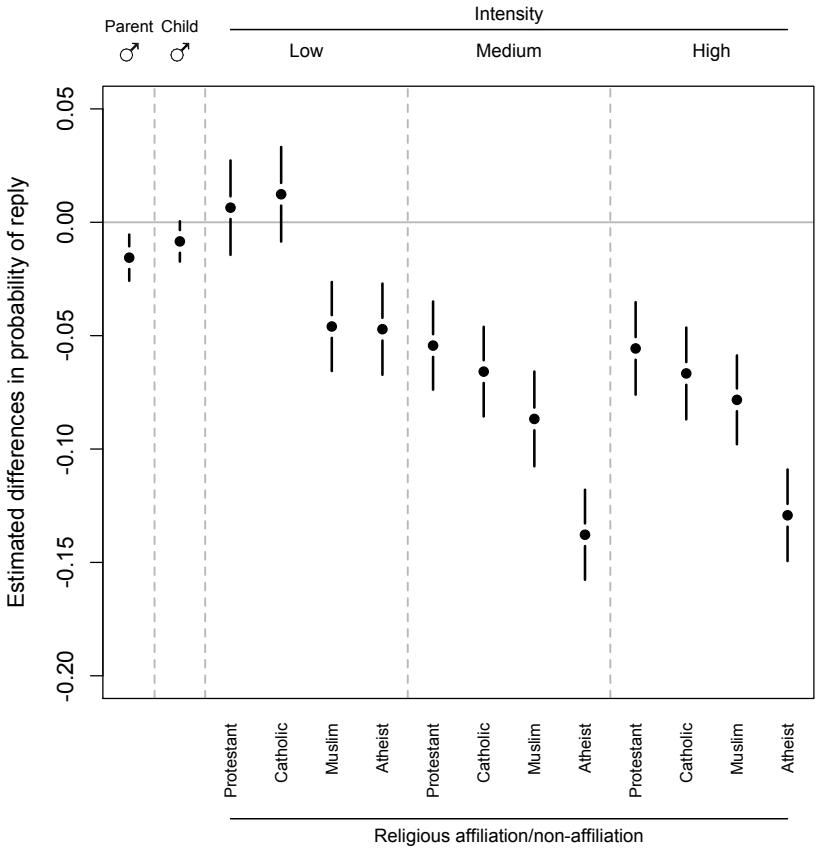
Note: Emails revealing no information about the parent's religious affiliation/non-affiliation exclude text blocks A, B and C. Among emails that do reveal religious affiliation/non-affiliation, low intensity requests include C (but not A or B), medium intensity requests include A and C (but not B), and high intensity requests include B and C (but not A).

Figure 2: Estimated probabilities of reply based on model in Table 1



Note: The plot shows estimated probabilities of receiving a reply and 95% confidence intervals based on the probit model in Table 1. Robust standard errors are clustered at the school district level.

Figure 3: Estimated treatment effects based on model in Table 1



Note: The plot shows estimated differences in probabilities of receiving a reply (i.e., treatment effects) and 95% confidence intervals based on the probit model in Table 1. Robust standard errors are clustered at the school district level.

VIII. Tables

Table 1: Parameter estimates (probit)

	est.	se
intercept	0.131^{***}	(0.037)
Male parent	-0.040^{***}	(0.014)
Male child	-0.022^{*}	(0.012)
Protestant, low intensity	0.018	(0.026)
Protestant, medium intensity	-0.137^{***}	(0.026)
Protestant, high intensity	-0.141^{***}	(0.026)
Catholic, low intensity	0.032	(0.026)
Catholic, medium intensity	-0.165^{***}	(0.025)
Catholic, high intensity	-0.169^{***}	(0.027)
Muslim, low intensity	-0.116^{***}	(0.026)
Muslim, medium intensity	-0.220^{***}	(0.027)
Muslim, high intensity	-0.200^{***}	(0.026)
Atheist, low intensity	-0.118^{***}	(0.026)
Atheist, medium intensity	-0.357^{***}	(0.027)
Atheist, high intensity	-0.334^{***}	(0.027)

Note: $N = 45{,}710$. The table shows estimates and robust standard errors clustered at the school district level from a probit model. The model contains email account fixed effects (not shown). Omitted categories are female parent, female child, and no information given.
* denotes statistical significance at 0.10 level. ** denotes statistical significance at 0.05 level. *** denotes statistical significance at 0.01 level.

IX. Supplementary information, to be published online

A. Treatment effect heterogeneity

We might expect the effects of treatment to vary by the social context in which principals are embedded. Examining treatment effect heterogeneity is complicated by the fact that we have a number of different treatments and many covariates. We simplify our analysis by collapsing the medium and high intensity levels, which, as shown in Figure 3 in the paper, lead to very similar effect estimates. We do the same with Protestant and Catholic religious affiliation, again for the same reason. Moreover, parametric models such as probit are poorly suited for modeling a large number of interactions of unknown functional form (Berry, DeMeritt, and Esaray 2015; Hainmueller, Mummolo, and Xu 2017). We therefore use Bayesian Additive Regression Trees (BART), one of the best off-the-shelf statistical learning estimators, to nonparametrically model treatment effect heterogeneity (Chipman, George, and McCulloch 2010; Hill 2011). We investigate how the treatment effects vary as a function of school-level characteristics (the share of Asian, Hispanic, Black, and White students as well as school type (primary, middle, high)) and county-level characteristics (median household income, the share of adults holding bachelor degrees, the share of residents with income below the poverty line, Republican vote shares in the 2012 presidential elections, and religious adherence rates from the RCMS). We also include a dummy variable for the South. Finally, since it is possible that principals in schools/communities with greater diversity discriminate less, we also compute Herfindahl indices for racial and religious diversity and investigate whether treatment effects vary systematically with these indices.

Figures A4 and A5 show treatment effect estimates from a BART fit with 95% credible intervals. We find very little treatment effect heterogeneity. This is perhaps not too surprising given that covariates are measured at the school- or county-level and do not directly measure attitudes toward minority religions, non-believers, or the separation of church and state. (Fitting a linear probability model regressing the experimental outcome on all co-

variates produces an adjusted R^2 of merely 0.029, indicating that covariates not only fail to capture treatment effect heterogeneity but also variation in principals' responsiveness more generally.) We thus conclude that we observe substantively large levels of discrimination on average and that discrimination does not appear to be unique to any single social context. The available covariates do not allow us to more finely distinguish between types of schools or communities with low and high levels of discrimination.

B. Generalizing impact estimates to NCES universe

Finally, we formally generalize our results to the NCES population of 78,348 regular, non-charter public schools in the 48 contiguous U.S. states without missing data. Following Author (2016), we generalize effect estimates by reweighting our experimental sample so that it matches the NCES target population in terms of covariate means for the covariates listed in Table A1. Weights are generated using maximum entropy weighting (Hainmueller 2012), which guarantees that reweighted sample covariate means equal the covariate means in the NCES population. Estimates from Weighted Least Squares regressions are reported in Table A4. Reweighting the sample so that it matches the NCES population has only a small effect on our treatment effect estimates, in line with the earlier findings that our sample and the NCES population match closely (Figures A1–A3) and that treatment effect heterogeneity is limited (Figures A4 and A5). Based on these results, it seems plausible that the patterns of discrimination we observe in our experimental sample are not unique to the schools in our experiment. It is likely that similar discrimination occurs across public schools throughout the U.S.

Table A1: Balance

covariate	p-value
% Asian students	0.97
% Hispanic students	0.28
% Black students	0.57
% White students	0.53
% Male students	0.48
School size	0.35
Pupil/teacher ratio	0.44
% Free or reduced price lunch students	0.57
% GOP (2012 Presidential election)	0.39
% Median household income	0.59
% Bachelor degree	0.40
% Below poverty line	0.75
% Total adherents	0.99
% Black protestant adherents	0.74
% Evangelical protestant adherents	0.52
% Mainline protestant adherents	0.15
% Catholic adherents	0.50
% Muslim adherents	0.56

Note: The table shows exact p-values from univariate randomization inference tests of the null hypothesis that balance is as good as one would expect under block random assignment. Exact p-values are approximated using 10,000 randomly chosen blocked treatment assignments. The test statistic is the maximum KolmogorovSmirnov statistic across all two-way comparisons of treatment groups. The p-value is the fraction of test statistics at least as large as the test statistic in our sample. For the religious adherence covariates, the balance tests use the average of five multiply imputed datasets.

Table A2: Parameter estimates (probit) controlling for blocks

	est.	se
intercept	−0.411	(0.253)
Male parent	−0.045***	(0.014)
Male child	−0.024**	(0.012)
Protestant, low intensity	0.017	(0.028)
Protestant, medium intensity	−0.147***	(0.027)
Protestant, high intensity	−0.156***	(0.027)
Catholic, low intensity	0.032	(0.028)
Catholic, medium intensity	−0.180***	(0.027)
Catholic, high intensity	−0.183***	(0.028)
Muslim, low intensity	−0.124***	(0.027)
Muslim, medium intensity	−0.241***	(0.029)
Muslim, high intensity	−0.217***	(0.028)
Atheist, low intensity	−0.127***	(0.027)
Atheist, medium intensity	−0.384***	(0.028)
Atheist, high intensity	−0.362***	(0.028)

Note: $N = 45{,}710$. The table shows estimates and robust standard errors clustered at the school district level from a probit model. The model contains email account fixed effects as well as block fixed effects (not shown). Omitted categories are female parent, female child, and no information given.
* denotes statistical significance at 0.10 level. ** denotes statistical significance at 0.05 level. *** denotes statistical significance at 0.01 level.

Table A3: Parameter estimates (OLS)

model	(1)		(2)		(3)	
	est.	se	est.	se	est.	se
intercept	0.551***	(0.014)	0.353***	(0.077)	0.670***	(0.157)
Male parent	−0.015***	(0.005)	−0.016***	(0.005)	−0.016***	(0.005)
Male child	−0.009*	(0.004)	−0.009**	(0.004)	−0.009**	(0.004)
Protestant, low intensity	0.007	(0.011)	0.006	(0.010)	0.006	(0.010)
Protestant, medium intensity	−0.054***	(0.010)	−0.053***	(0.010)	−0.053***	(0.010)
Protestant, high intensity	−0.056***	(0.010)	−0.056***	(0.010)	−0.056***	(0.010)
Catholic, low intensity	0.013	(0.010)	0.012	(0.010)	0.012	(0.010)
Catholic, medium intensity	−0.065***	(0.010)	−0.066***	(0.010)	−0.066***	(0.010)
Catholic, high intensity	−0.067***	(0.010)	−0.066***	(0.010)	−0.066***	(0.010)
Muslim, low intensity	−0.046***	(0.010)	−0.045***	(0.010)	−0.046***	(0.010)
Muslim, medium intensity	−0.086***	(0.011)	−0.087***	(0.010)	−0.086***	(0.010)
Muslim, high intensity	−0.079***	(0.010)	−0.079***	(0.010)	−0.079***	(0.010)
Atheist, low intensity	−0.047***	(0.010)	−0.046***	(0.010)	−0.045***	(0.010)
Atheist, medium intensity	−0.138***	(0.010)	−0.136***	(0.010)	−0.135***	(0.010)
Atheist, high intensity	−0.129***	(0.010)	−0.129***	(0.010)	−0.129***	(0.010)
Email account fixed effects	yes		yes		yes	
Block fixed effects	no		yes		yes	
Covariates	no		no		yes	

Note: $N = 45{,}710$. The table shows estimates and robust standard errors clustered at the school district level from three linear probability models. All models contain email account fixed effects (coefficients not shown), model (2) additionally contains block fixed effects (coefficients not shown), and model (3) additionally contains the covariates listed in Table A1 (coefficients not shown). Omitted categories are female parent, female child, and no information given.

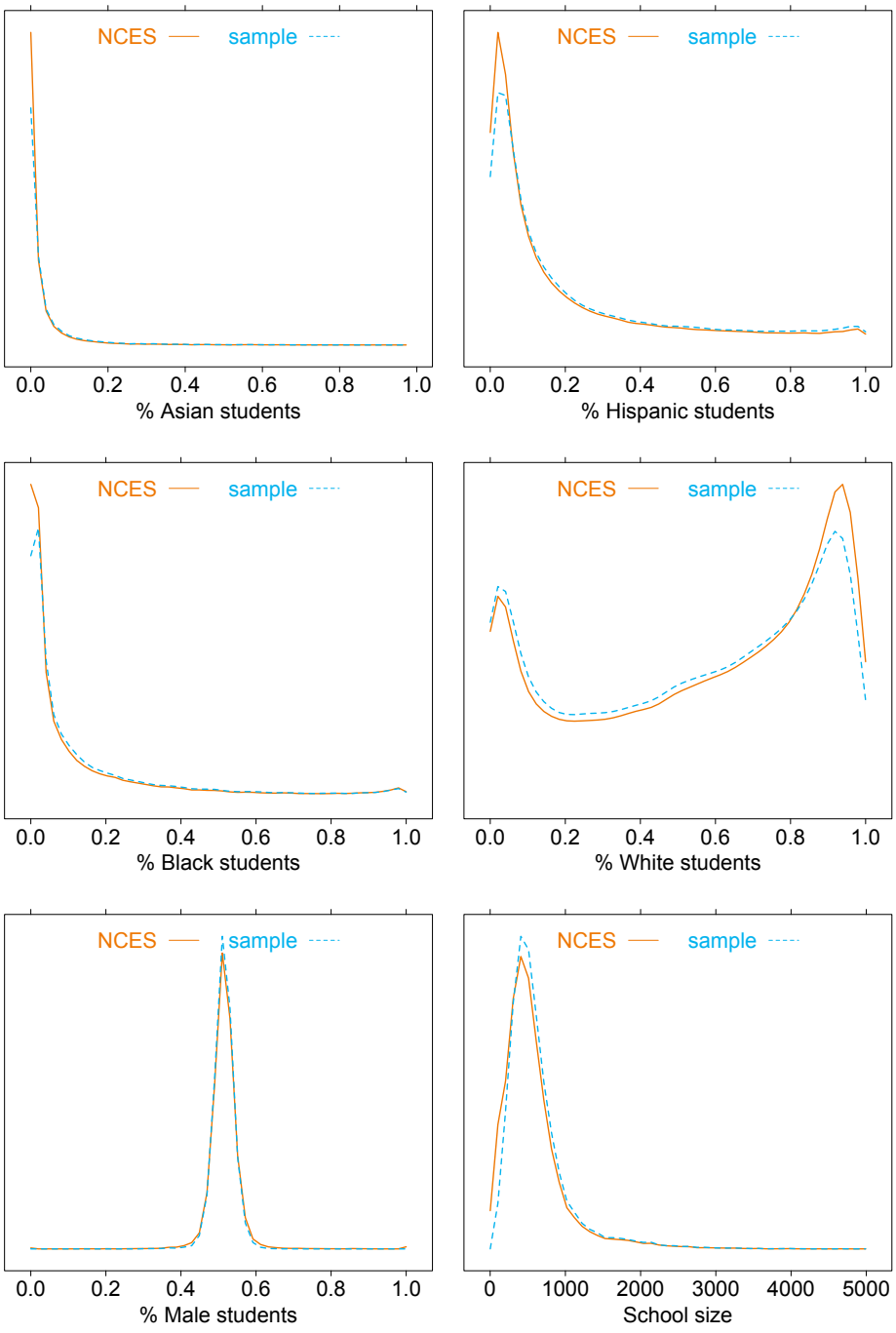
* denotes statistical significance at 0.10 level. ** denotes statistical significance at 0.05 level. *** denotes statistical significance at 0.01 level.

Table A4: Parameter estimates from population-weighted sample (WLS)

	est.	se
intercept	0.554^{***}	(0.017)
Male parent	-0.012^{**}	(0.006)
Male child	-0.009^{*}	(0.005)
Protestant, low fervor	0.014	(0.011)
Protestant, medium fervor	-0.045^{***}	(0.011)
Protestant, high fervor	-0.056^{***}	(0.011)
Catholic, low fervor	0.018	(0.011)
Catholic, medium fervor	-0.063^{***}	(0.011)
Catholic, high fervor	-0.066^{***}	(0.011)
Muslim, low fervor	-0.038^{***}	(0.011)
Muslim, medium fervor	-0.086^{***}	(0.011)
Muslim, high fervor	-0.081^{***}	(0.011)
Atheist, low fervor	-0.046^{***}	(0.011)
Atheist, medium fervor	-0.139^{***}	(0.011)
Atheist, high fervor	-0.124^{***}	(0.012)

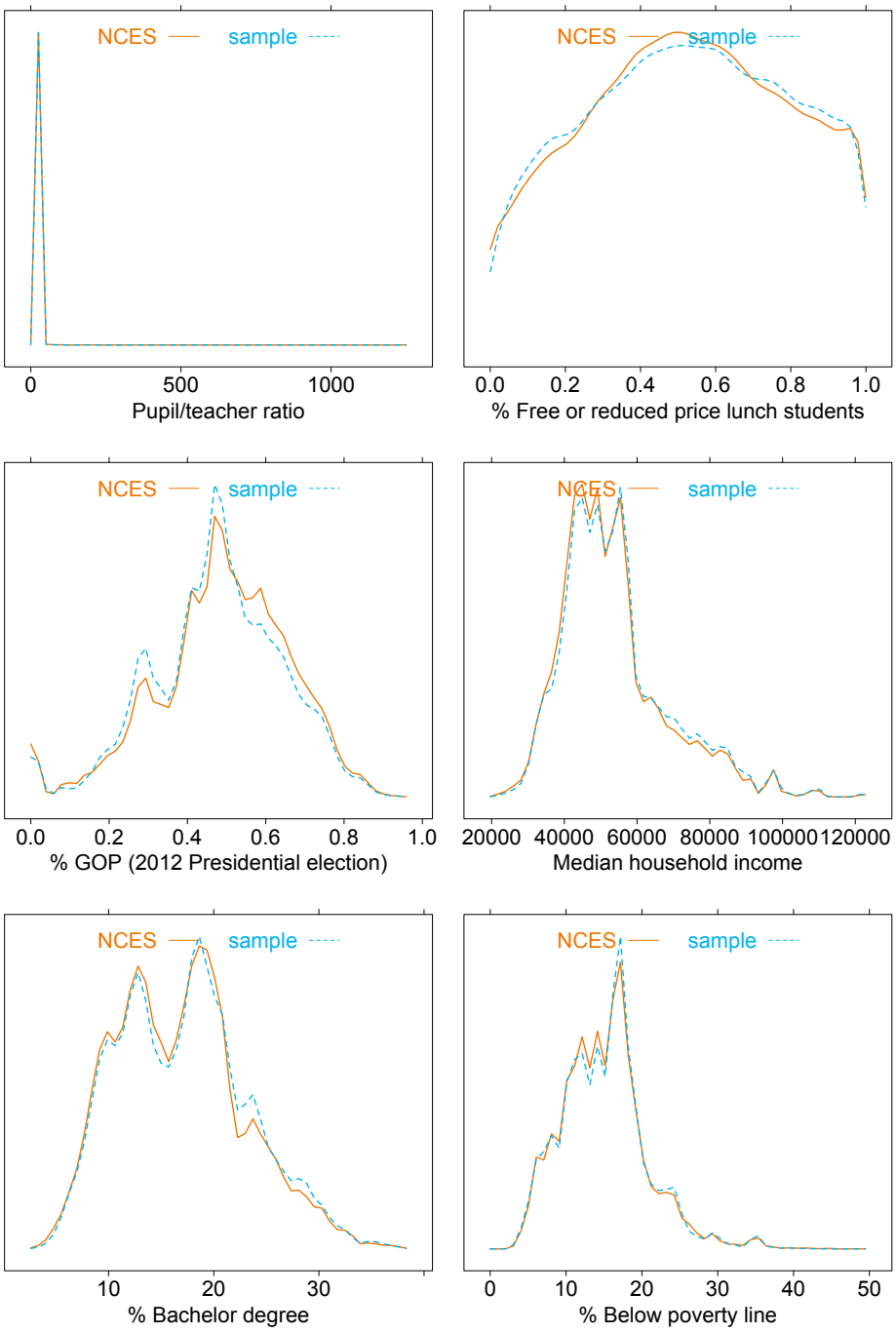
Note: $N = 45{,}710$. The table shows estimates and robust standard errors clustered at the school district level from a linear probability model weighted with maximum entropy weights (see text). The model also contains email account fixed effects (coefficient estimates not shown). Omitted categories are female parent, female child, and no information given.
* denotes statistical significance at 0.10 level. ** denotes statistical significance at 0.05 level. *** denotes statistical significance at 0.01 level.

Figure A1: Comparison between sample and NCES population I



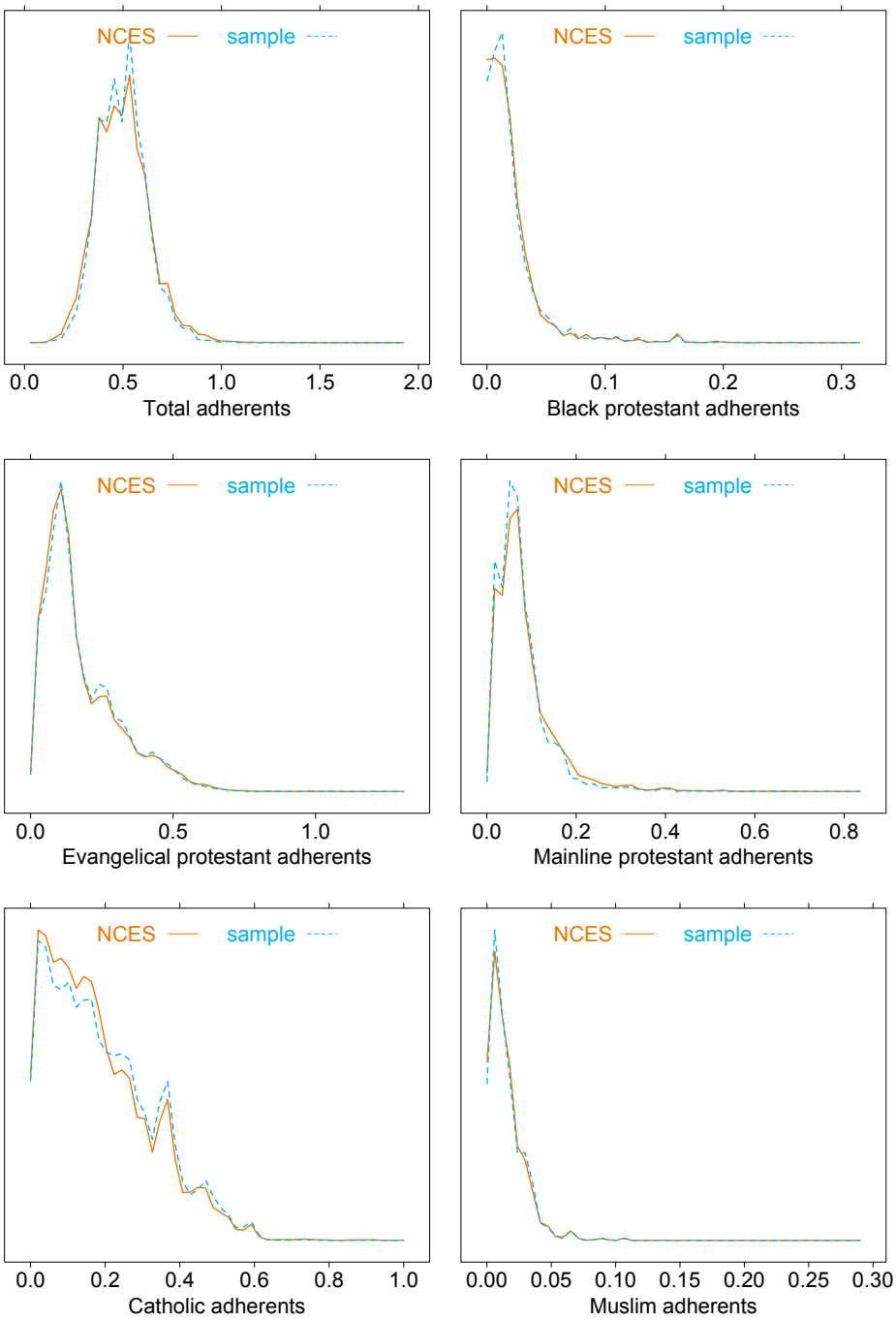
Note: The kernel density plots compare covariate distributions in the experimental sample with the NCES population of 78,348 regular, non-charter public schools without missing data in the 48 contiguous U.S. states.

Figure A2: Comparison between sample and NCES population II



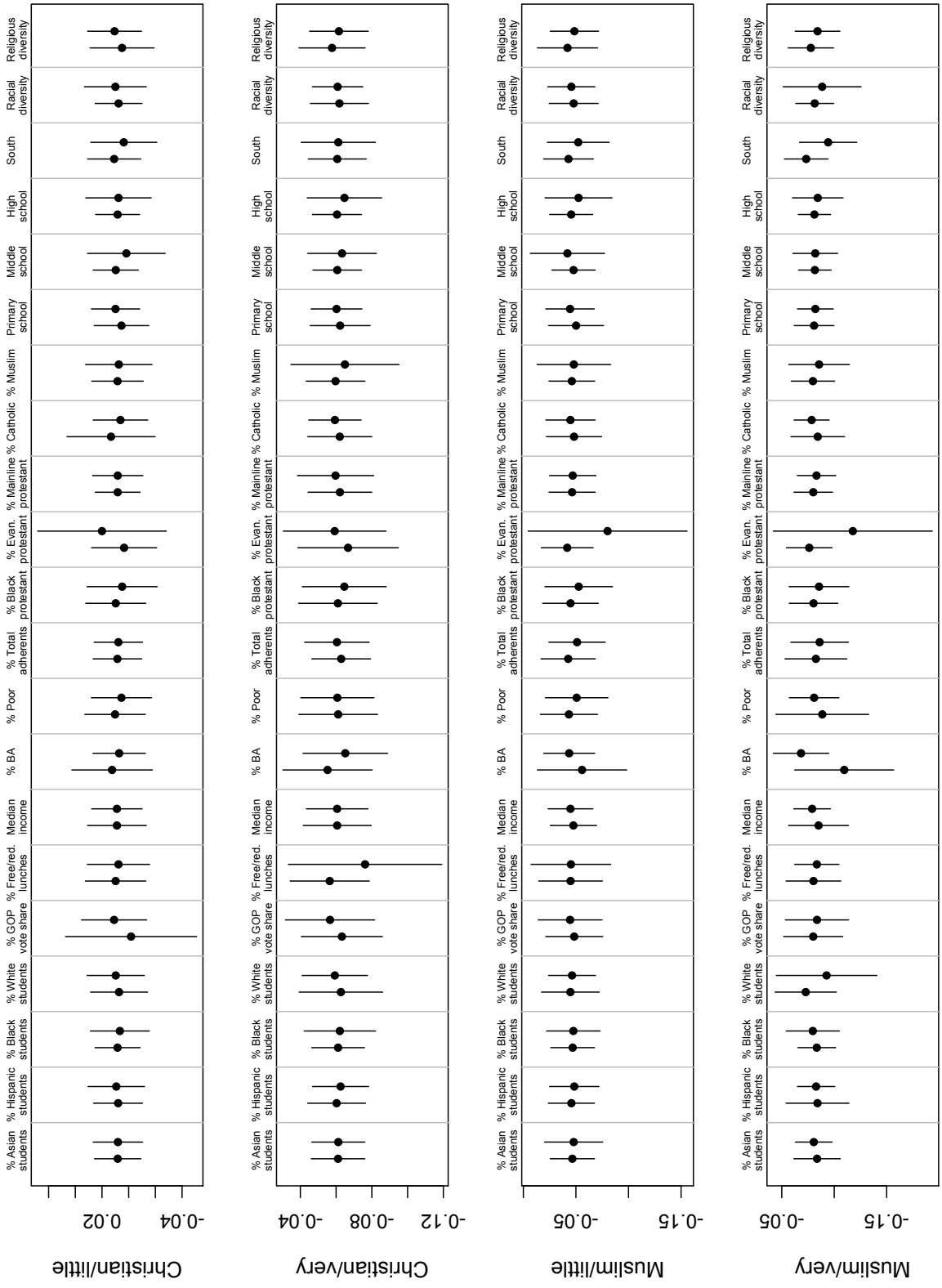
Note: The kernel density plots compare covariate distributions in the experimental sample with the NCES population of 78,348 regular, non-charter public schools without missing data in the 48 contiguous U.S. states.

Figure A3: Comparison between sample and NCES population III



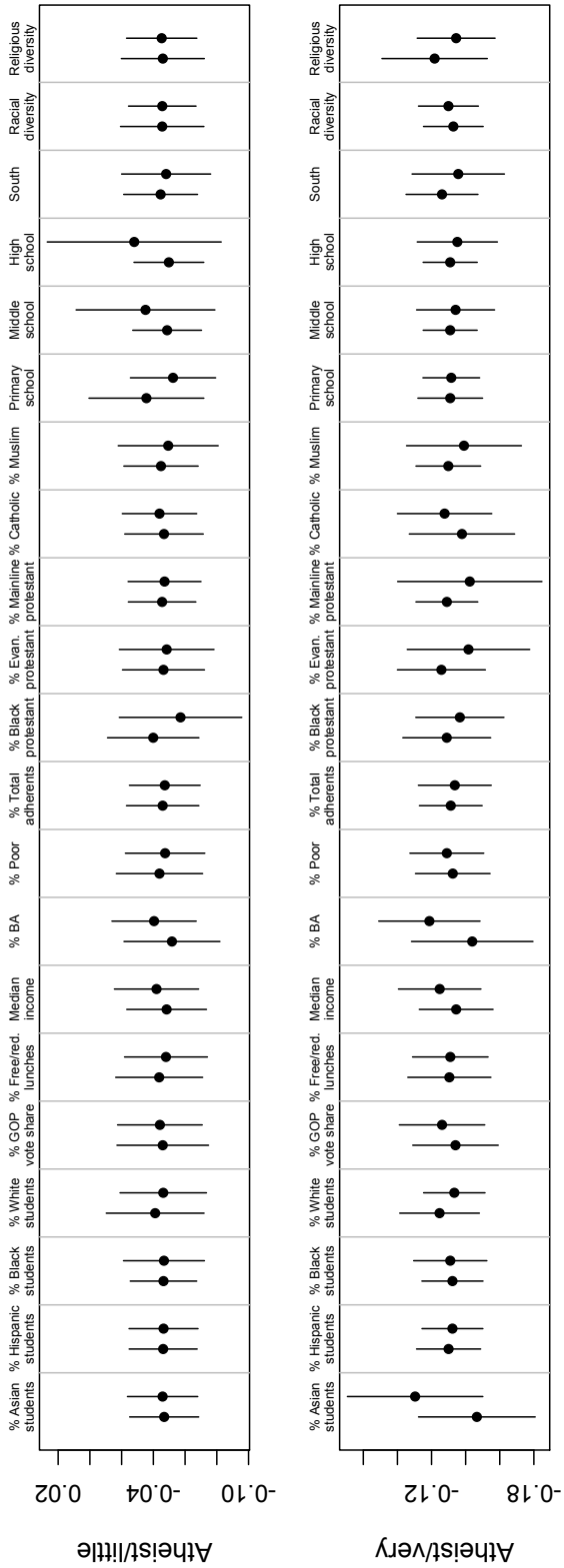
Note: The kernel density plots compare covariate distributions in the experimental sample with the NCES population of 78,348 regular, non-charter public schools without missing data in the 48 contiguous U.S. states.

Figure A4: Treatment effect heterogeneity I



Note: The plots show treatment effect heterogeneity as a function of covariates from a Bayesian Additive Regression Tree (BART) fit, contrasting treatment effect estimates for when a given covariate is set to the 0.05 sample quantile and for when it is set to the 0.95 sample quantile.

Figure A5: Treatment effect heterogeneity II



Note: The plots show treatment effect heterogeneity as a function of covariates from a Bayesian Additive Regression Tree (BART) fit, contrasting treatment effect estimates for when a given covariate is set to the 0.05 sample quantile and for when it is set to the 0.95 sample quantile.

Berry, William D., Jacqueline H.R. DeMeritt, and Justin Esarey. 2015. "Bias and overconfidence in parametric models of interactive processes." *American Journal of Political Science* 60 (2): 521–539.

Chipman, Hugh A., Edward I. George, and Robert E. McCulloch. 2010. "BART: Bayesian additive regression trees." *Annals of Applied Statistics* 4 (1): 266–298.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2017. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." Working paper.

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis* 20 (1): 25–46.

Hill, Jennifer L. 2011. "Bayesian nonparametric modeling for causal inference." *Journal of Computational and Graphical Statistics* 20 (1): 217–240.