

NBER WORKING PAPER SERIES

THE SURPRISINGLY SMALL EFFECTS OF RELIGION-BASED DISCRIMINATION
IN EDUCATION

Victor Lavy
Edith Sand
Moses Shayo

Working Paper 24922
<http://www.nber.org/papers/w24922>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue
Cambridge, MA 02138
August 2018, Revised November 2020

A previous version of this paper circulated as “Charity Begins at Home (and at School): Religion-Based Discrimination in Education”. We thank Josh Angrist, James Fenske, Jonathan Guryan, Imran Rasul, and participants at seminars at Ben Gurion University, IFS London, Bank of Israel, IDC Herzliya, Northwestern University, University of Bonn, University of Zurich, University of Warwick, and CAGE Venice Applied Micro Economics Conference for useful comments and suggestions. We also thank the Israel’s Ministry of Education and Dr. Haim Gat and Eliad Trefler for allowing restricted access to schooling data in the Ministry online protected research lab. Evgeni Rachkovski provided excellent research assistance. Lavy acknowledges financial support from the European Research Council through ERC Advanced Grant 323439. Shayo thanks the I-Core Program at the Israel Science Foundation (grant no. 1821/12). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Victor Lavy, Edith Sand, and Moses Shayo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Surprisingly Small Effects of Religion-Based Discrimination in Education
Victor Lavy, Edith Sand, and Moses Shayo
NBER Working Paper No. 24922
August 2018, Revised November 2020
JEL No. J24,J48,J71

ABSTRACT

Religions often preach preferential treatment of fellow believers, but the magnitude and economic implications of religion-based discrimination are largely unknown, partly because religiosity is often confounded with ethnicity. We analyze grading decisions in national matriculation exams in Israel, exploiting unique features that reveal student religiosity to the graders, and grader religiosity to the researcher. We find evidence of religiosity-based ingroup bias. Substantively, however, the effects of this bias are small. One reason is that religious bias is entirely driven by men. Furthermore, patterns of bunching in the grade distribution suggest the primary source of bias is the religious (rather than secular) men – a small fraction of the grader population. A second potential reason is that many graders live in integrated communities. Indeed, we find that living and working in close proximity to people with different levels of religiosity appears to attenuate religion-based discrimination.

Victor Lavy
Department of Economics
University of Warwick
Coventry, CV4 7AL
United Kingdom
and Hebrew University of Jerusalem
and also NBER
v.lavy@warwick.ac.uk

Moses Shayo
Department of Economics
The Hebrew University of Jerusalem
Jerusalem 91905
Israel
mshayo@huji.ac.il

Edith Sand
Bank of Israel
P.O. Box 780
91007, Jerusalem
Israel
edith.sand@boi.org.il

A data appendix is available at <http://www.nber.org/data-appendix/w24922>

1. Introduction

Economists of religion have long been interested in the causes and effects of religiosity and secularization (Barro and McCleary 2003; Gruber and Hungerman 2008; Iyer 2016; Becker, Rubin and Woessmann 2020). But what is the impact of the emerging rift between secular and religious groups in modern economies? Many religions openly preach preferential treatment of fellow believers, and sometimes harsh treatment of non-believers. Secular people might similarly show bias against religious individuals. However, since one's level of religiosity is rarely observed in administrative data, and since groups that differ in their religiosity often vary in other dimensions (including race or ethnicity), religion-based discrimination is hard to isolate and quantify. This paper makes a first attempt.

Israel's high school matriculation system offers a unique opportunity to study religion-based discrimination. It is a centralized country-wide scheme of exams that affect both a student's prospects for continuing to higher education as well as their field of study. The exact same exams are taken by both religious and secular students.¹ Each exam booklet is randomly assigned to two independent graders (or "examiners"). Grading decisions are made under anonymous conditions, reducing the possibility of social pressure or reciprocity effects. Nonetheless, as explained in the next section, certain features of this setting allow the grader to infer student religiosity. Identifying the religiosity of both students and examiners is feasible in our setting because the Israeli public school system is divided into religious and secular schools. We are thus able to infer student religiosity from the schools they go to, and examiner religiosity from the schools they send their children to. Finally, we have detailed data on the grades given to each exam booklet, where the grades range from 0 to 100. Observing the entire distribution of grades allows us to exploit bunching at certain points in the distribution to go beyond measuring relative ingroup bias, and identify the source of discrimination

We begin with a difference-in-differences model, exploiting the random assignment of exam booklets and allowing for systematic differences across levels of religiosity both in student ability and in examiner standards. Intuitively, we compare the mean difference in grades given

¹ Ultra-orthodox students have a separate system and are not included in our data.

to religious versus secular students by religious and secular examiners, controlling for student and questionnaire fixed effects (“questionnaire” refers to subject by level of proficiency, e.g. “math at level 4”). Using data from over 3.5 million grades given in 112 questionnaires in the years 2010–2015, we find evidence of a statistically significant but substantively small tendency toward religion-based ingroup bias. An exam grade is on average about 0.01 standard deviations higher when assigned to an examiner of the same (rather than different) level of religiosity as the student.

One clue into why the overall effect is so small comes from the fact that 83% of the examiners in our data are female. Research across cultures, time and samples, has demonstrated that, on average, men display more self-reported xenophobic and ethnocentric attitudes than do women. This has also been shown in lab experiments studying discrimination against outgroups and cooperation with the ingroup (see McDonald, Navarrete and Van Vugt 2012 for a review). Consistent with this body of work, we find that the bias in grading is driven almost entirely by male examiners. Female examiners exhibit little if any religion-based discrimination.

To further zoom in on the source of the bias, we examine the grade distributions. While the difference-in-differences analysis can detect ingroup bias, it does not identify the source of this discriminating behavior. However, using the existence of bunching of test scores at important test score thresholds (the failing mark and the perfect 100 grade), we can test separately for religious and secular examiners, if the likelihood of just crossing the threshold is higher when the student is religious rather than secular. The results suggest that the main source of the discriminating behavior is the (male) religious examiners. Male religious examiners account for only 5% of the examiners in our sample.

Another possible reason for the small overall bias is that, for the most part, secular and (non ultra-orthodox) religious Jews live and work in the same places. In line with inter-group contact theory, we find that religiosity-based discrimination might indeed be affected by exposure to people from other groups: in our case, people with a different level of religiosity. We examine several measures of exposure both at the community level (the neighborhoods where the examiners live) and at the workplace level (the schools where they teach). We also find that

religion-based discrimination is much more pronounced among the 3% of examiners who teach in segregated religious communities.

The paper relates most directly to the literature on the economics of religion which has studied the effects of religiosity and secularization at both the national and individual levels (Iyer 2016 provides a recent review). At the individual level, the literature has focused on such outcomes as income, education, and health-related behavior (Gruber and Hungerman 2008; Bryan, Choi, and Karlan 2018). Our analysis provides an important complement: while religiousness may have positive (or negative) effects relative to secularism, the cleavage itself might also have important implications as it can generate prejudice and discrimination, leading to bad allocations.

Another related literature studies racial and gender discrimination in settings such as the labor market and law enforcement (see Charles and Guryan 2013 and Bertrand and Duflo 2017 for reviews). We contribute to this literature in three important ways. First, we study discrimination along a very salient but little-studied dimension, namely, religiosity. Since we focus on a population with a similar ethnic and cultural background, we are able to isolate discrimination which is based on individuals' religious beliefs. Second, we study discrimination in the school system, which can have long-term implications for professional development and lifetime earnings.² Third, we provide evidence on inter-group contact theory, which has received increasing attention from economists in recent years (see Bertrand and Duflo 2017).

Finally, there seem to be both theoretical and empirical reasons to think that men are more prone than women to discriminate between ingroup and outgroup (see for example Balliet, Wu and De Dreu 2014). The argument is that male humans have evolved a specialized psychology that strengthens inter-group discrimination (e.g. Sidanius et al. 2000; Van Vugt, Cremer and Janssen 2007; Navarrete et al. 2010). However, most of the evidence in this literature is based on

² The literature on the economics of education often uses teachers' grading biases as a measure of discrimination. For example, Lavy (2008), Björn, Höglin, and Johannesson (2011), Burgess and Greaves (2013), Diamond and Persson (2016), , Lavy and Sand (2015), and Terrier (2016) use the systematic difference between non-blind and blind assessment across groups as a measure of such discrimination.

lab experiments and survey data. Our analysis suggests that this stronger male tendency for ingroup bias extends to professional high-stakes decisions.

2. Institutional Background

We focus in this study on schools in the academic track where the language of instruction is Hebrew. The vast majority of students in these schools are Jewish. These public schools can belong to two distinct sectors, according to level of religiosity. “State schools” are secular and serve the secular Jewish population. “State-religious schools” serve mainly the religious Jewish population, observe religious practices, emphasize religious teachings and in some subjects follow a different curriculum. Students receive a matriculation certificate (a prerequisite for university admission) after passing a series of state exams administered mostly at end of twelfth grade. Some subjects are mandatory and at least one elective is required at an advanced level. Religious and secular schools share over half the matriculation test questionnaires.

The final matriculation score in a given questionnaire is the mean of the test scores in a school-level (“internal”) exam and a national (“external”) exam. The latter is graded independently by two examiners, randomly assigned by a computer algorithm. Furthermore, the computerized process sends all exam booklets that were distributed in a specific classroom to the same two examiners. The final external score is the average of these two examiners’ evaluations.

The external exam booklets do not reveal a student’s identity to the grader: they only include the student’s ID number and school code. Nonetheless, the examiner can, in principle, look up the school code and find out whether the school is religious or not. More importantly, even if the examiner does not look up the school code, most religious Jews write a special inscription—*BS”D*—at the top of every written document. Thus, the level of religiosity of the student is effectively revealed to the examiners.³

³ The inscription *BS”D* is an acronym for *Besiyata DiShmaya*, an Aramaic phrase meaning “with the help of heaven.” Religious Jews write this inscription (or a variation thereof) at the top of the first page of every written document. See the online Appendix for more detailed discussion and examples of notebooks with the *BS”D* notation.

3. Data and Descriptive Statistics

Our data include all matriculation questionnaires taken in the summer session in schoolyears 2010–2015 by Jewish students in the state education system (both religious and secular). We start with the matriculation test scores database. Each matriculation test score record contains student, school, and class identifiers, as well as the grade, questionnaire number, number of credit units, scores given by the first and second examiners, and the school-level (“internal”) score. Importantly, we have identifiers for both examiners. We merge the matriculation exam record with the student database of the same year to obtain student characteristics (grades, class and school assignment and school zip code, gender, ethnicity, number of siblings, and parents’ education). Merging the data with the school file allows us to determine student religiosity according to their schools’ religious orientation.

All examiners are teachers. We can thus obtain information on examiners from teachers’ files for the years 2000–2015. This includes main field of instruction, main school assignment, gender, number of children, age, education and ethnicity. Examiners are defined as religious if at least one of their children attended a religious school. To determine examiner religiosity, we construct a new database that contains each parent who had a child enrolled in high school during 1998–2016. Using this database, we can determine the level of religiosity of about 85% of the examiners in our sample.⁴ We also develop several measures of examiners’ exposure to different environments, which we describe below when discussing the contact hypothesis. The online Appendix provides a detailed description of the construction of the database.

Tables A3 and A4 in the appendix present descriptive statistics at the student level and test level respectively. The total number of students who took at least one summer exam in Hebrew during 2010–2015 is 423,002 students. One-quarter of these students came from religious schools. The proportion of girls and the number of siblings are both higher among religious students (the proportion of girls is 62% versus 51% and the average number of siblings is 2.25 versus 0.9). Other characteristics are similar across sectors. With respect to test scores, secular and religious students

⁴ A series of balancing tests finds that students assigned to examiners with missing religiosity do not differ significantly from other students. See Appendix Table A2.

have similar external test scores on average (70.5 versus 70), as well as similar probabilities of passing the exam.

Our sample includes 2508 examiners, 83% of whom are female and one-third are religious. Of the religious examiners, one third are Ultra-Orthodox and about 13% teach at schools located in segregated religious areas (religious settlements). Overall, secular and religious examiners have similar observed characteristics. See Appendix Tables A5. Appendix Tables A6 and A7 provide additional descriptive statistics on examinations. The dataset includes around 2 million exam booklets, from one thousand schools. The number of exam booklets graded by an examiner is on average 1650. The mean number of booklets per student is 4.88 (std.=2.77) and the total number of questionnaires is 112.

4. Identification and Estimation

To identify religion-based ingroup bias, we rely on the random assignment of students' exam booklets to examiners within a given questionnaire. To evaluate this identifying assumption, we test whether booklets assigned to religious examiners were systematically different from booklets assigned to secular examiners within a given questionnaire-year, in terms of a host of student characteristics. Appendix Table A8 presents the results for all examiners, and separately for male and female examiners. Except for one case, none of the estimated differences are significantly different from zero. These balancing tests confirm that the computer algorithm that assigns exam booklets is indeed random with respect to examiner religiosity.

Consider the following benchmark difference-in-differences specification:

$$(1) \quad y_{bijqt} = \alpha_0 + \alpha_1 ReligStudent_i + \alpha_2 ReligExaminer_j \\ + \alpha_3 ReligStudent * ReligExaminer_{ij} + \beta_i + \gamma_q + \delta_t + \varepsilon_{bijqt}$$

where y_{bijqt} is the outcome (e.g., test score) of exam booklet b , written by student i , assigned to examiner j , in questionnaire q , in year t . $ReligStudent_i$ and $ReligExaminer_j$ are indicator variables for religious student and religious examiner. The baseline specification includes

questionnaire (γ_q) and year (δ_t) fixed effects. We further include student fixed effects (β_i). ε_{bijqt} is an error term clustered within examiner (the level of treatment).

Equation (1) allows for two possible differences across religiosity groups that do not necessarily indicate religious bias. First, it is possible that exams written by religious students have different unobserved characteristics (including, but not limited to, different quality) than those written by secular students. Thus, α_1 may be nonzero even in the absence of religious bias. Second, it is possible that religious and secular examiners have different grading standards (e.g., religious examiners may be more lenient). In other words, α_2 may be nonzero even in the absence of religious bias. Examiner religious bias is captured by α_3 . This coefficient reflects a difference-in-differences: by how much religious examiners are more generous than secular examiners when grading an exam written by a religious student rather than a secular one.

5. Results

Table 1 shows baseline results. The unit of observation is an exam booklet graded by a particular examiner. The dependent variable is the (normalized) score. The number of observations is twice the number of exam booklets, since each booklet is graded by two different examiners.

Before estimating equation (1), columns 1 and 2 estimate separately for religious and secular examiners, the difference in grades given to religious versus secular students, controlling for questionnaire and year fixed effects. Both religious and secular examiners give lower grades to religious students, but the difference is larger among secular examiners. Column 3 estimates equation (1) but for comparability only includes questionnaire and year fixed effects. Religious students' test scores are lower by 0.05 of a SD and religious examiners are marginally more generous. The ingroup bias estimate is reported in the third row and equals 0.011, which is the difference between the two religious student indicators' estimates in the first two columns. Thus, test scores are on average higher by 1% of a SD when the exam booklet is assigned to an examiner with similar religiosity to the student. This effect is quite small and statistically insignificant.

In column 4 we add student fixed effects (hence the religious student indicator drops out). The estimated ingroup bias remains unchanged though it is now statistically significant at $p=0.077$.⁵ Column 5 further includes booklet fixed effects. The estimated ingroup bias is slightly smaller and much more precise. This last specification captures within-booklet differences in test scores given by examiners of a different religious orientation than both types of students. However, once we stratify the sample to different subgroups, we will not be able to perform this more demanding estimation strategy.

In Appendix Table A10 we report results of a similar exercise when restricting attention only to examiners who send *all* their children to one type of school (religious vs secular). This provides a sharper contrast between secular and religious examiners though based on a smaller sample. As one might expect, the results indicate larger ingroup bias for both male and female examiners compared to Table 1. However, the estimated overall ingroup bias in our preferred specification in column 4 remains small at 0.015 (SE=0.007).

Why is the effect so small? The psychology literature suggests that group biases are less common among females. As the last two columns of Table 1 show, the estimated ingroup bias of male examiners is 0.030 (se=0.015), three times larger than the average effect shown in column 3. Female ingroup bias is much smaller and not significantly different from zero. This is consistent with patterns seen in lab experiments studying ingroup favoritism (Balliet et al. 2014), but emerges here on a much larger scale and concerning decisions that have important lifetime implications.

One concern when interpreting α_3 is that it might capture differential treatment by religious and secular examiners of some other student characteristic, rather than their religiosity. In Table A11 we augment equation (1) with interaction terms of religious examiner and various student characteristics (gender, mother's and father's education, number of siblings, parental country of birth). The ingroup bias estimate is stable and virtually unaffected by the inclusion of

⁵ Appendix Table A9 presents estimations of ingroup biases based on raw test scores, instead of standardized scores. The magnitude and significance of the estimated ingroup bias align with the results in Tables 1.

the additional interaction terms. Furthermore, six of the eight additional interaction terms are not statistically significant. Thus, our estimated ingroup bias does not appear to capture omitted interaction bias due to examiners favoring other student characteristics.

Another concern is that religious and secular examiners may grade a given booklet differently because they differentially like a particular feature in it, for example, the student's way of reasoning. In other words, we might be capturing a coincidence of taste or style between the student and the examiner and not religion-based discrimination by examiners. Appendix Table A12 splits the sample to STEM and non-STEM subjects. The latter include social studies and humanities, where the examiner might be more prone to bias grades because of writing style or expressed views. In STEM subjects the correct answer tends to be more definitive. However, the estimates of ingroup bias are very similar: 0.012 in STEM and 0.010 in non-STEM subjects.

5.1 Final Matriculation Outcomes and a Placebo Exercise

Table 2 (column 1) estimates the impact of ingroup bias on the *external score* (the average score of the two examiners). The treatment is the proportion of religious examiners for each exam booklet (zero, 0.5, or 1) times the indicator of religious student. The estimated ingroup bias is 0.02 standard deviations (statistically significant). When the treatment indicator is equal to 0.5 (one of the two examiners is of the same religiosity as the student), the ingroup bias effect is equal to 0.01, the respective estimate that we report in Table 1.

Recall that the *internal score* is provided by student's own teacher. Since it is filed prior to the external exam, it serves as a useful placebo outcome. As seen in column 2, the treatment effect on the internal grade is an order of magnitude smaller and not significantly different from zero. This supports the identification strategy in Table 1.

Column 3 reports the impact of ingroup bias on the final grade, which is an average of the external and internal scores. This estimate is 0.010 (se=0.004), close to the average of the estimates reported in columns 1 and 2.

Columns 4–6 in Table 2 examine the likelihood of passing the exam (the mean probability of passing in the sample is 89%). Overall, the probability of passing a matriculation exam increases by 0.005, or half a percentage point, when both examiners share the student’s religiosity. Columns 5-6 break this effect by student background. Not surprisingly, students from high-education families are unaffected, since they have a much lower likelihood of being at the margin of failing or passing a matriculation exam. Among students from low-education families, however, the point estimate is 0.009 (relative to a mean probability of 0.83 of passing an exam in this group).

5.2 Sources of Bias: Evidence from Test Score Bunching

Figure 1 shows the distribution of test scores by examiner and student religiosity. In all distributions, we observe substantially larger mass at two points in the distribution: at 55, the passing score in a matriculation exam, and at 100, the highest score possible in these exams. This bunching can be viewed as evidence that examiners systematically adjust grades to be just enough to pass the exam or, for the best students, to get a perfect score. In this section we exploit these adjustments to explore the sources religion-based discrimination.

As in our baseline regressions, we continue to allow religious examiners to systematically display more (or less) of this bunching behavior. We also allow religious students to systematically receive more (or less) of these upward adjustments.⁶ In the regressions reported in columns 1–3 of Table 1, we cannot infer that religious students receive lower grades unjustifiably, as they may be systematically weaker. However, being more (or less) likely to receive an upward adjustment might indicate general discrimination against one group, beyond any preference for one’s own group.

⁶ This may be due to a general bias for or against one of the groups, but in the case of the bunching at 100, it might in principle also be due to one group having a higher proportion of students who write outstanding exams that get censored at 100. However, as we will see below, religious students have the same likelihood as secular students to score 100 rather than any score in the range 90-99.

Table 3 focuses on the passing grade threshold, noting that examiners may push up a grade within a close range of the passing grade and not necessarily from 54 to 55. Panel A estimates a variant of equation (1) where the dependent variable is the probability of passing the exam (a grade of at least 55). We estimate these regressions using four different subsamples according to test scores, beginning with all exam booklets with test scores between 50 and 60 and then looking at narrower intervals: [54,60], [54,57], and [54,56]. Each column is a separate regression that includes questionnaire fixed effects.

Notice first that we find little consistent evidence of general discrimination in favor (or against) religious students (first row). The ingroup bias estimates are consistently positive but are largely restricted to male examiners. Ingroup bias among male examiners seems particularly large when focusing on the two ranges closest to the passing threshold: the likelihood of “bumping” a student from one’s religious group from 54 to 55/56 or from 54 to 55-57 is 4.3 and 3.2 percentage points higher ($p=0.101$ and $p=0.077$), respectively. This effect is sizeable and is equivalent to about 5–6% of the mean passing rate in the whole sample. By contrast, the estimated ingroup bias of female examiners in these two ranges is zero.

Panel A of Table 4 repeats this exercise at the margin of scoring 100, restricting the sample to test scores within the following ranges: [90,100], [95,100], [98,100], and [99,100]. Importantly, there is no evidence that religious students are overall more likely to receive a grade of 100 rather than any grade in the 90-100 range (first row of first column). Furthermore, we again observe sharp differences in ingroup bias between male and female examiners. When looking only at exams graded by men, the likelihood of getting 100 versus 99 is almost 11 percentage points higher when the exam is assigned to an examiner of the same religiosity as the student. Strikingly, ingroup bias estimates among the female examiners in all four ranges are zero.

Before continuing, it is important to note that the overall ingroup bias we documented in the previous sections is not limited to these ranges. The ingroup bias estimate (in the preferred specification in column 4 of Table 1) remains 0.10 ($se=0.06$, $p=0.096$), even when we remove from the sample test scores in the ranges 55–60 and 95–100.

The difference-in-differences estimate we have been studying so far is a *relative* measure of ingroup bias. We cannot tell whether the source of discriminating behavior is the secular or religious examiners. The difficulty is due to the lack of an objective test score for each exam. It may be the case that secular students perform better on exams and hence the extent to which secular examiners give them higher grades is not an indication of a bias. In this example the bias is entirely due to religious examiners. But, of course, the reverse is also possible and the bias might be entirely due to the secular examiners. This limitation is common in studies that attempt to identify ingroup bias in naturally occurring (non-experimental). For example, Shayo and Zussman (2011) detect ingroup bias among Arab and Jewish judges in Israel, but absent an objective measure of the “correct” outcome, cannot definitively determine whether the bias is driven by Jewish or Arab judges (or both). Similarly, Anwar Bayer and Hjalmarsson (2012) find that in Florida, the presence of a member of one’s race in the jury pool entails a better outcome for the defendant, but again cannot pin down the source of the bias. Here, we propose a simple way to help address this limitation.

Our approach is based on test score bunching. We examine whether the likelihood of increasing test scores above the failing grade or to the 100 score is higher among, say, religious examiners when they grade exam booklets of religious students versus secular students. Note that while secular and religious students may well write different quality exams on average, it is less likely that they systematically vary in the likelihood of writing an exam worth 99 versus 100 (or 54 versus 55). This allows us to test for discrimination separately for secular and religious examiners in these ranges.

Panels B-C of Table 3 focus on the probability of passing the exam. The dependent variable is an indicator for scoring 55 or higher and the main explanatory variable is a dummy for religious student. Consider first the male examiners. Among secular examiners (panel B), the coefficient on religious student is negative in all four columns, consistent with discrimination against religious students. However, all the estimates are imprecisely measured and, for the most part, are not statistically different from zero. At the same time, the estimated coefficients for *religious* examiners (panel C) are all positive, implying a pro-religious student bias, but again only one of

the estimates is statistically different from zero (0.024, $se=0.015$, $p=0.1$). Note that the difference between the estimated pro-religious bias of the religious and secular examiners gives us the ingroup bias reported in panel A.

The patterns of bias towards the best students are sharper (Table 4, panels B-C). For male religious examiners the estimates in columns 5-8 are positive, high, and significant. For male secular examiners, they are negative as expected, but much smaller, and mostly insignificant. The bias toward religious students among male religious examiners is especially large in the 99–100 range. The probability of a score of 100 is higher by almost 10 percentage points when it is a religious student. The respective bias of the secular males is much lower at 0.017 ($se=0.021$). Clearly, the religious examiners drive most of the ingroup bias at this bunching of test scores.

Turning to female examiners, columns 9–12 of both Table 3 and 4 show little evidence of bias among either the secular or the religious examiners. In other words, the lack of overall bias among women in Tables 1 is unlikely to be masking differences between religious and secular women (e.g., due to ingroup bias in one group and out-group bias in the other). This lends support to the argument that gender differences in intergroup relations is not limited to a particular culture.

An intriguing question about the nature of the discrimination of male religious examiners is whether they increase the grades of students from their own group (“ingroup love”) or whether they lower the grades of students from the other group (“out-group hate”). The surplus mass at test scores 55 and 100 and the “hole” in the test score distribution at 54 and 99 suggest that male religious examiners inflate test scores of religious students and do not lower test scores of secular students. This is consistent with Feld, Salamanca, and Hamermesh (2016) who find that examiners’ favoritism toward their own group, rather than discrimination against the outgroup, explains relative ingroup bias by nationality and by gender.

In Appendix Tables A13 and A14 we report results on the variation in bias by examiner characteristics. Most interesting are the results on differences across religious orientation *within* the religious group. The results in Appendix Table A14 (column 1) indicate that religious bias of Ultra-Orthodox examiners is small and not significantly different from zero. This result is

consistent with the often-expressed opinion that Ultra-Orthodox Jews do not view the Religious-National Jews (who attend the state schools) as “truly” religious. Hence, they do not treat them as part of their ingroup.

5.3 Does Ingroup Bias Decline with Exposure to the Out-Group?

In this section, we examine whether religion-based discrimination declines with examiners’ exposure to people of different levels of religiosity at home (the neighborhood where they live) and at work (the school where they teach). Since many secular and (non-ultra-orthodox) religious Jews live and work in close proximity to each other, this might help further explain the surprisingly small effects of religious ingroup bias documented above.⁷ It is also an interesting question in its own right. A long line of literature dating back to the 1940s suggests that intergroup contact can reduce intergroup prejudice (Pettigrew and Tropp 2006). However, Paluck et al. (2019) argue that this literature tends to rely on young participants (less than 25 years old) and to focus on very short-term outcomes (typically measured on the day of the intervention). In this sense, our setting offers a valuable addition to the literature. We should stress, however, that we do not have random assignment of peers and hence the analysis in this section should be taken as suggestive.

We construct several measures of exposure to the outgroup at school. The teacher database contains information on all teachers in each school, including their demographic information and main fields of study. Since all examiners are teachers, merging it with parents’ files enables us to compute for each examiner in a given year: (1) the proportion of peers at school from a religious background; (2) the proportion of peers at school from a religious background who teach the same subject as the examiner; and (3) the proportion of peers at school from a religious background who have the same gender.

⁷ Thus, on average 19% of the children in the same zip-code as the examiner have a different level of religiosity. In particular, for religious examiners, 30% of the children in their neighbourhood are secular. See Appendix Table A15

Similarly, we compute a geographical measure of examiners' exposure to other religious outgroup each year in their neighborhood, using the proportion of religious/secular students within the examiners' zip code. We use students' and teachers' neighborhood zip codes received from the Ministry of Education which enable to characterize for each teacher's zip code in a given year the proportion of students who attended religious schools, and merge it with teachers' files for the relevant year. We provide the full regression results in the Appendix, and summarize the results here.

Start by looking at examiners who teach in segregated religious localities. Ninety percent of the Jewish settlements in the West Bank are such communities, and three percent of the examiners teach in one of them. In Appendix Table A14 (column 2) we augment equation (1) with interactions with an indicator for religious examiners who teach in a religious settlement. The results suggest that ingroup bias of examiners from religious communities in the West Bank is about four times larger than the mean effect of 0.01.

Next, we use four different definitions of exposure, measured in two environments: the neighborhood in which one lives and the school in which one works (Appendix Table A15 provides descriptive statistics and Tables A16 and A17 report the results.⁸ Specifically, we look at:

- a) neighbors within the examiner's home zip code;
- b) peers (other teachers) at school;
- c) peers (other teachers) at school who teach the same subject;
- d) peers (other teachers) at school with the same gender.

We then augment equation (1) with the proportion of neighbors or peers in the environment with a different level of religiosity, fully interacted with the religiosity variables (and in particular with $ReligStudent * ReligExaminer_{ij}$). The regressions include year and student fixed effects as well

⁸ In Appendix Table A16 exposure is measured as a dummy variable indicating an above-median proportion of neighbors or peers in the environment with a different level of religiosity, while in Appendix Table A17 exposure is measured in continuous terms, as the proportion of neighbors or peers with a different religious orientation.

as, importantly, examiner by environment (zip code or school) by questionnaire fixed effects. Thus for example, in case (a), the interaction picks up the variation in ingroup bias for a given examiner living in the same neighborhood, whose neighborhood's religious composition changed over time.

The estimates for male examiners suggest that ingroup bias declines sharply when examiners encounter a higher proportion of the outgroup in their neighborhood. Ingroup bias is positive and quite large (estimated at 0.064, $se=0.021$) when the examiner is exposed to below-median proportion of neighbors with different religiosity, but drops to zero when the examiner is highly exposed to the other group in the neighborhood. Male ingroup bias is also associated with changes in exposure to "others" at work, especially to teachers who teach the same subject or are of the same gender.

For female examiners, the estimates suggest an interesting pattern. The main ingroup bias in all four cases is small and not significantly different from zero. However, ingroup bias appears to emerge among female examiners when they are in the minority in terms of religiosity at school, and in particular among female teachers at school. This is inconsistent with a simple version of the contact hypothesis that ignores the importance of the conditions under which contact takes place.

6. Conclusions

While secularization—and its opposite, resacralization—have drawn enormous attention, the economic effects of religion-based discrimination have gone largely unnoticed. Using data from Israel's high-stakes matriculation exams we are able to identify the level of religiosity of both students and examiners, and thus study discrimination across religious and secular members of the same ethno-religious group. This allows us to disentangle religion-based from ethnic discrimination.

We have five main findings. First, we document the existence of ingroup bias in grading decisions. This bias is detectable among professional graders who are making highly

consequential decisions. Second, the bias is, overall, very small, amounting to about 0.01 of a standard deviation in the grade. Third, looking at the sources of the bias, we find that it is almost entirely driven by male examiners: female examiners (who constitute over 80% of the examiners) show little if any bias. Fourth, using bunching in the grading distribution we find evidence that bias, at least at the top of the distribution, is largely driven by male religious examiners. Male religious examiners are six to ten percentage points more likely to bump a grade to 100 when the exam is written by a religious student, while male secular examiners are between one and three percentage points less likely to do so when grading a religious student. Finally, we find suggestive evidence that contact across religious and secular groups may attenuate these biases. While our setting offers a rather unique opportunity to study religion-based discrimination using large administrative datasets, the basic idea could be replicated in an experimental setting by randomly revealing to examiners the religiosity of some of the students, using culturally-relevant cues. This could be done in different countries, at different phases of the secularization-resacralization process.

References

- Angrist, J., & Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American economic review*, 99(4), 1384-1414.
- Anwar, S., P. Bayer, and R. Hjalmarsson (2012). "The Impact of Jury Race in Criminal Trials." *Quarterly Journal of Economics*, 127(2), 1017-1055.
- Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup Favouritism in Cooperation: a Meta-Analysis. *Psychological Bulletin*, 140(6), 1556.
- Barro, R. J. and R. M. McCleary (2003). "Religion and Economic Growth across Countries." *American Sociological Review*, 68(5), 760-781.
- Becker, S. O., Rubin, J., & Woessmann, L. (2020). "Religion in Economic History: A Survey." *CEPR Discussion Paper* No. DP14894.
- Bertrand, M. and E. Duflo (2017). "Field Experiments on Discrimination." In E. Duflo and A. Banerjee (eds.), *Handbook of Economic Field Experiments*, Volume 1 (pp. 309-393). Amsterdam: North Holland.

- Bryan, G., J. J. Choi, and D. Karlan (2018). "Randomizing Religion: The Impact of Protestant Evangelism on Economic Outcomes." *NBER Working Paper* 24278.
- Burgess, S. and E. Greaves (2013). "Test Scores, Subjective Assessment, and Stereotyping of Ethnic Minorities." *Journal of Labor Economics*, 31, 535–576.
- Charles, K. K. and J. Guryan (2013). "Taste-based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots." *The Economic Journal*, 123(572), 417–432.
- Diamond, R. and P. Persson (2016). "The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests." NBER Working Paper 22207.
- Feld, J., N. Salamanca, and D. S. Hamermesh (2016). "Endophilia or Exophobia: Beyond Discrimination." *The Economic Journal*, 126(594), 1503–1527.
- Gruber, J. and D. Hungerman (2008). "The Church vs the Mall: What Happens When Religion Faces Increased Secular Competition?" *Quarterly Journal of Economics*, 123(2), 831–862.
- Iyer, S. (2016). "The New Economics of Religion." *Journal of Economic Literature*, 54(2), 395–441.
- Lavy, V. (2008). "Do Gender Stereotypes Reduce Girls' or Boys' Human Capital Outcomes? Evidence from a Natural Experiment." *Journal of Public Economics*, 92, 2083–2105.
- Lavy, V. and E. Sand (2018). "On the Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Stereotypical Biases." *Journal of Public Economics*, 167, 263–279.
- McDonald, M. M., Navarrete, C. D., & Van Vugt, M. (2012). Evolution and the psychology of intergroup conflict: The male warrior hypothesis. *Phil. Trans. R. Soc. B*, 367(1589), 670–679.
- Paluck, E. L., Green, S. A., & Green, D. P. (2019). "The contact hypothesis re-evaluated". *Behavioural Public Policy*, 3(2), 129–158.
- Pettigrew, T. F. and L. R. Tropp (2006). "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology*, 90, 751–783.
- Shayo, M. and A. Zussman (2011). "Judicial Ingroup Bias in the Shadow of Terrorism." *Quarterly Journal of Economics*, 126(3), 1447–1484.
- Sidanius, J., S. Levin, J. Liu, and F. Pratto (2000). "Social Dominance Orientation, Anti-egalitarianism and the Political Psychology of Gender: An Extension and Cross-cultural Replication." *European Journal of Social Psychology*, 30(1), 41–67.
- Terrier C. (2016), "Boys Lag Behind: How Teachers' Gender Biases Affect Student Achievement." *SEII Working Paper* 2016.07.
- Van Vugt, M., Cremer, D. D., & Janssen, D. P. (2007). Gender differences in cooperation and competition: The male-warrior hypothesis. *Psychological science*, 18(1), 19–23.

Table 1: Religion-Based Discrimination in Test Scores

	Religious Examiners	Secular Examiners	All Examiners			Male Examiners	Female Examiners
	Questionnaire and Year Fixed Effects	Questionnaire and Year Fixed Effects	Questionnaire and Year Fixed Effects	Questionnaire, Year and Student Fixed Effects	Booklet Fixed Effects	Questionnaire, Year and Student Fixed Effects	Questionnaire, Year and Student Fixed Effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Religious Student	-0.041 (0.010)	-0.051 (0.007)	-0.051 (0.006)				
Religious Examiner			0.019 (0.007)	0.011 (0.005)	0.014 (0.004)	0.017 (0.013)	0.011 (0.0060)
Religious Student x Religious Examiner			0.011 (0.012)	0.010 (0.006)	0.008 (0.003)	0.030 (0.015)	0.010 (0.006)
Number of Observations	1,201,625	2,388,491	3,590,116	3,590,116	3,590,116	508,324	3,081,792
Number of Examiners	715	1,400	2508	2508	2508	431	2,064
Proportion of Examiners	0.338	0.662	1.000	1.000	1.000	0.173	0.827

Notes: The first two columns of the table present the difference in grades given to religious and secular students, separately by religious (column 1) and secular examiners (column 2). The estimates of the religious student indicator are from a specification that includes questionnaire and year fixed effects. The next four columns present the difference-in-differences ingroup bias estimates, from different specifications: in column 3 the specification includes only questionnaire and year fixed effects; in column 4 the specification includes also student fixed effects; and the last specification includes only booklet fixed effects. The next two columns present the difference-in-differences ingroup bias estimates from the preferred specification that includes questionnaire, year and student fixed effects, separately for male and female examiners. The number of observations is twice the number of booklets, since each booklet appears twice (once for each examiner). The numbers and proportions of examiners by gender and religiosity are presented in the last two columns. Dependent variables are standardized scores. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table 2: Ingroup Bias in Related Exam Outcomes

	Average External Exam Grade	Internal Exam Grade: Placebo Test	Average Final Exam Grade	Probability of Passing the Exam		
				All Students	Students with Low Parental Education	Students with High Parental Education
	(1)	(2)	(3)	(4)	(5)	(6)
Proportion of Religious Examiners	0.005 (0.002)	0.000 (0.002)	0.003 (0.002)	-0.002 (0.001)	-0.003 (0.002)	-0.001 (0.001)
Religious Student x Proportion of Religious Examiners	0.020 (0.005)	-0.002 (0.005)	0.010 (0.004)	0.005 (0.002)	0.009 (0.004)	0.001 (0.002)
Number of Observations	1,565,252	1,535,550	1,535,550	1,535,556	627,818	883,892

Notes: The table presents the estimated effect of ingroup bias of examiners on additional outcomes: 1) the average external exam grade (the average of the two examiners' normalized scores); 2) the normalized internal exams, which are exams examined by students' school teachers; 3) the final exam score (the average of the external and internal exams' normalized scores); 4) probability of passing the exam (if final grade ≥ 55); 5) probability of passing the exam from a subsample of students with low parental education (low parental education is equal to one if both parents have 12 or less years of schooling); 6) and the probability of passing the exam from a subsample of students with high parental education. The proportion of religious examiners is measured in each exam booklet. The number of observations is the number of booklets, since each booklet appears only once. All columns present the results from separated regressions based on the preferred specification (which includes year, questionnaire, and student fixed effects). Standard errors are corrected for clustering at the student level and are presented in parentheses.

Table 3: Ingroup Bias in the Probability of Passing the Exam, by Examiner Gender, Religiosity and Test Score Range

	All Examiners				Male Examiners				Female Examiners			
	[60,50]	[54,60]	[54,57]	[54,56]	[60,50]	[54,60]	[54,57]	[54,56]	[60,50]	[54,60]	[54,57]	[54,56]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. All Examiners												
Religious Student	-0.009 (0.002)	-0.001 (0.002)	0.001 (0.004)	-0.001 (0.005)	-0.012 (0.007)	-0.006 (0.060)	-0.007 (0.011)	-0.016 (0.014)	-0.008 (0.003)	0.000 (0.002)	0.003 (0.004)	0.002 (0.006)
Religious Examiner	-0.008 (0.003)	0.002 (0.003)	0.001 (0.007)	-0.004 (0.009)	-0.017 (0.011)	-0.010 (0.012)	-0.028 (0.022)	-0.054 (0.028)	-0.007 (0.004)	0.004 (0.004)	0.005 (0.007)	0.003 (0.009)
Religious Student x Religious Examiner	0.006 (0.004)	0.007 (0.004)	0.010 (0.007)	0.009 (0.009)	0.012 (0.012)	0.017 (0.010)	0.032 (0.018)	0.043 (0.026)	0.005 (0.004)	0.005 (0.004)	0.006 (0.007)	0.002 (0.010)
Number of Observations	371,094	255,779	127,998	84,110	51,394	42,279	18,070	11,722	319,700	220,236	109,028	72,388
B. Secular Examiners												
Religious Student	-0.009 (0.002)	-0.001 (0.002)	0.001 (0.004)	-0.001 (0.005)	-0.012 (0.007)	-0.006 (0.006)	-0.007 (0.011)	-0.016 (0.014)	-0.008 (0.003)	0.003 (0.004)	0.003 (0.004)	0.002 (0.006)
Number of Observations	250,814	173,779	87,446	57,752	33,476	23,929	11,996	7,862	217,338	150,487	75,450	49,890
C. Religious Examiners												
Religious Student	-0.003 (0.003)	0.006 (0.003)	0.011 (0.006)	0.008 (0.008)	0.000 (0.010)	0.011 (0.008)	0.024 (0.015)	0.027 (0.022)	-0.003 (0.004)	0.005 (0.003)	0.009 (0.006)	0.004 (0.008)
Number of Observations	120,280	82,000	40,552	26,358	17,918	12,251	6,074	3,860	102,362	64,740	34,478	22,498

Notes: The dependent variable is the probability of passing the exam (if score \geq 55). The coefficients in each column are from separated regressions that include questionnaire fixed effects, for four different subsamples: in the first column the subsample includes all tests with scores between 50 and 60; in the second column the subsample includes all tests with scores between 54 and 60; in the third column the subsample includes all tests with scores between 54 and 57; and in the last column the subsample includes all tests with scores between 54 and 56. Panel A includes all examiners and Panel B (Panel C) includes only secular (religious) examiners. Standard errors are corrected for clustering at the examiner level and are presented in parentheses.

Table 4: Ingroup Bias in the Probability of Scoring 100, by Examiner Gender, Religiosity and Test Score Range

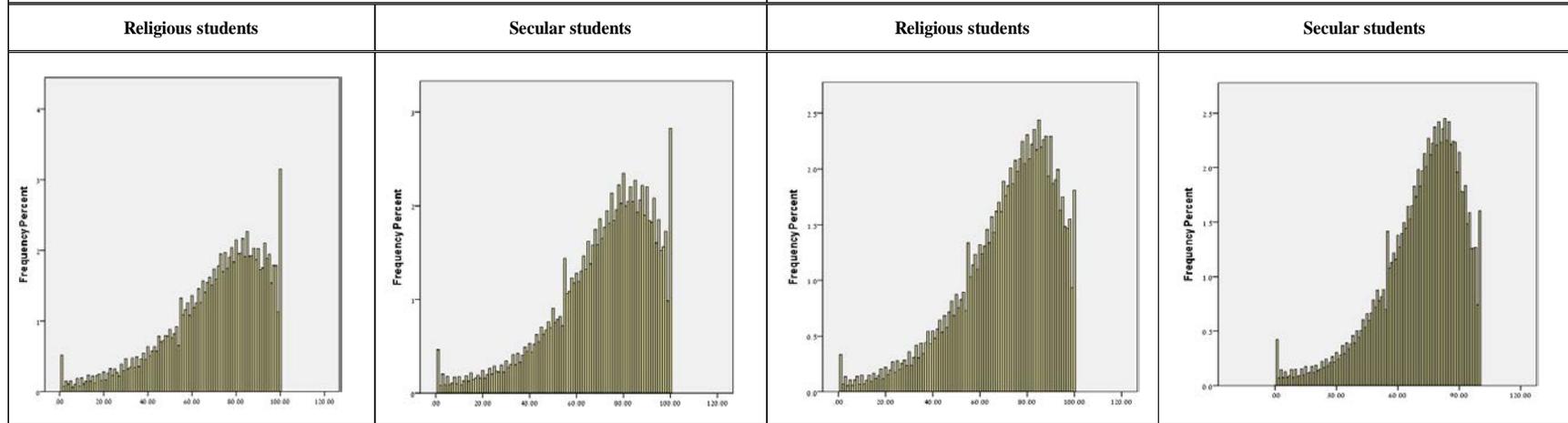
	All Examiners				Male Examiners				Female Examiners			
	[90,100]	[95,100]	[98,100]	[99,100]	[90,100]	[95,100]	[98,100]	[99,100]	[90,100]	[95,100]	[98,100]	[99,100]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. All Examiners												
Religious Student	-0.001 (0.003)	-0.008 (0.006)	-0.023 (0.011)	-0.002 (0.013)	-0.006 (0.005)	-0.011 (0.010)	-0.035 (0.017)	-0.017 (0.021)	0.001 (0.004)	-0.006 (0.007)	-0.019 (0.013)	0.002 (0.014)
Religious Examiner	0.000 (0.005)	-0.005 (0.009)	-0.022 (0.016)	-0.025 (0.018)	-0.001 (0.016)	-0.015 (0.029)	-0.050 (0.049)	-0.070 (0.053)	-0.001 (0.005)	-0.004 (0.009)	-0.016 (0.015)	-0.012 (0.018)
Religious Student x Religious Examiner	0.005 (0.005)	0.009 (0.009)	0.026 (0.016)	0.025 (0.018)	0.029 (0.011)	0.046 (0.019)	0.098 (0.032)	0.109 (0.036)	0.001 (0.005)	0.000 (0.009)	0.006 (0.018)	0.000 (0.019)
Number of Obs.	557,641	243,970	105,919	68,332	89,101	42,158	20,001	13,894	468,540	201,812	85,918	54,438
B. Secular Examiners												
Religious Student	-0.001 (0.003)	-0.008 (0.006)	-0.023 (0.011)	-0.002 (0.013)	-0.006 (0.005)	-0.011 (0.010)	-0.034 (0.017)	-0.017 (0.021)	0.001 (0.004)	-0.006 (0.007)	-0.019 (0.013)	0.002 (0.014)
Number of Obs.	361,929	156,690	67,505	43,667	55,150	25,384	11,863	8,233	306,779	131,306	55,642	35,434
C. Religious Examiners												
Religious Student	0.004 (0.003)	0.000 (0.006)	0.003 (0.012)	0.023 (0.013)	0.023 (0.009)	0.035 (0.016)	0.065 (0.027)	0.096 (0.031)	0.001 (0.003)	-0.007 (0.007)	-0.013 (0.012)	0.000 (0.013)
Number of Obs.	195,712	87,280	38,414	24,665	33,951	16,774	8,138	5,661	161,761	70,506	30,276	19,004

Notes: The dependent variable is the probability of scoring 100 on the exam. The coefficients in each column are from separated regressions that include questionnaire fixed effects, for four different subsamples: in the first column the subsample includes all tests with scores between 90 and 100; in the second column the subsample includes all tests with scores between 95 and 100; in the third column the subsample includes all tests with scores between 98 and 100; and in the last column the subsample includes all tests with scores between 99 and 100. Panel A includes all examiners and Panel B (Panel C) includes only secular (religious) examiners. Standard errors are corrected for clustering at the examiner level and are presented in parentheses.

Figures: The Distributions of Scores, by Examiners' Gender and Religiosity and Students' Religiosity

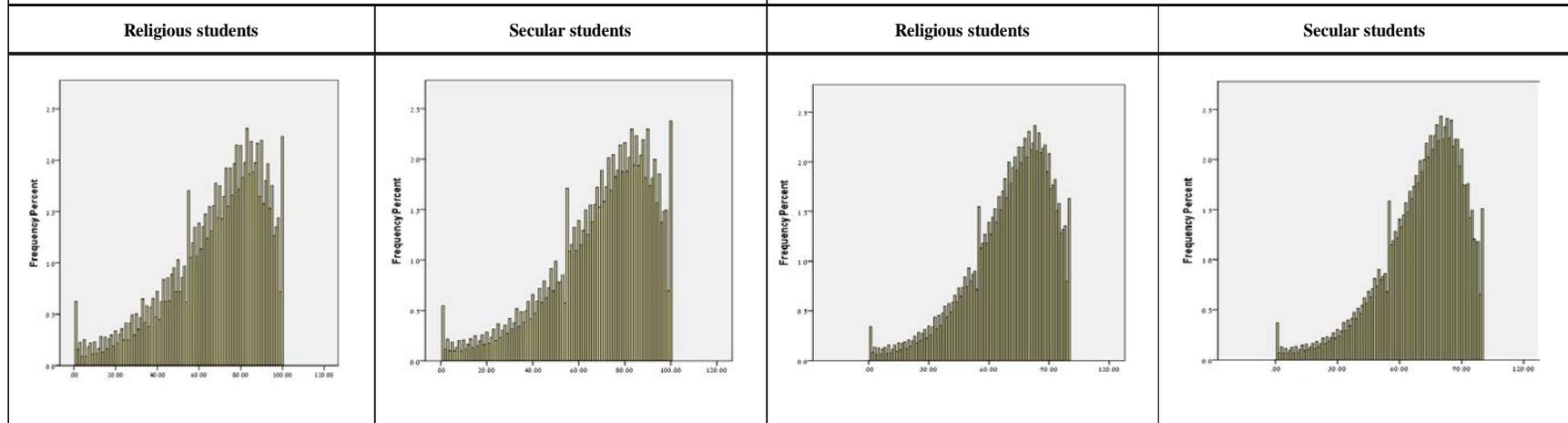
A. Religious Male Examiners

B. Religious Female Examiners



C. Secular Male Examiners

D. Secular Female Examiners



Appendix for Online Publication

The Surprisingly Small Effects of Religion-Based Discrimination in Education

Victor Lavy, Edith Sand and Moses Shayo

November 2020

Contents

Construction and Description of the Database.....	3
Revealing Religiosity	5
Table A1: Balancing Tests for Writing BS"D.....	6
Table A2: Balancing Tests for the Assignments of Students' Booklets to Examiners with Missing Values of Religious Status.....	7
Table A3: Summary Statistics of Students' Characteristics, by Students' Religious Status	8
Table A4: Summary Statistics of Tests Scores, by Students' Religious Status.....	9
Table A5: Summary Statistics of Examiners' Characteristics, by Gender and Religiosity	10
Table A6: Summary Statistics of Tests Examination	11
Table A7: Number of Booklets by Religious Status and Gender of Examiners	12
Table A8: Balancing Tests for the Assignments of Students' Booklets to Examiners, by Examiners' Gender	13
Table A9: Estimated In-Group Biases of Examiners on Raw Test Scores, by Examiners' Gender.....	14
Table A10: Estimated In-Group Biases of Examiners on Test Scores, According to a Stricter Definition of Examiners' Religiosity Status	15
Table A11: Sensitivity of the Results to Students' Characteristics.....	16
Table A12: Estimated In-Group Biases of Examiners in STEM and Non-STEM Subjects, by Examiners' Gender	17
Table A13: Estimated In-Group Biases of Examiners on Test Scores, by Examiners' Characteristics....	18
Table A14: Estimated In-Group Biases of Examiners on Test Scores, by Type of Religious Examiners..	19
Table A15: Summary Statistics of Examiners' Exposure to Different Religiosity Environmental Measures, by Religiosity and Gender of Examiners.....	20
Table A16: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Discrete Measures of Exposure, by Examiners' Gender	21
Table A16: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Discrete Measures of Exposure, by Examiners' Gender - Continued	22
Table A17: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Continuous Measures of Exposure, by Examiners' Gender	23
Table A17: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Continuous Measures of Exposure, by Examiners' Gender - Continued	24
Figure A1: Sample of Religious Students' Booklets.....	25

Construction and Description of the Database

The data used in this study includes all matriculation questionnaires taken in the summer session by Jewish students in both the religious and secular state education system in the school years 2010–2015.¹ Since we do not have information on the matriculation exams' language, we exclude Arab students who attended Arab schools and foreign-born students as their exam booklets were most likely not written in Hebrew. We start with the matriculation test scores database. Each matriculation test score record contains student, school, and class identifiers, as well as the grade, questionnaire number, number of credit units, scores given by the first and second examiners, and the school-level ("internal") score. Importantly, we also have data on both examiners' identifiers. Next, we merge the matriculation exam record of each student with the student database of the same year to obtain student characteristics (grades, class and school assignment and school zip code, gender, ethnicity based on parents' country of birth², number of siblings, and parents' education). Student religiosity was determined according to their schools' religious orientation by merging the data with the school file (containing each school's location, religious orientation, and whether it is a gender-segregated school).

A crucial requirement for the analysis was obtaining information on examiners. The fact that examiners have to teach the subject of the exam in high school for several years before grading matriculation exams enables us to obtain information on them from teachers' files for the years 2000–2015. The information on each examiner (main field of instruction, main school assignment, gender, number of children, age, education and ethnicity, school assignment and school zip code) is obtained from the teacher database of the relevant year or earlier (in case the examiner did not teach in a certain year) and merged with the school database of the same year in order to add schools' religious orientation.

To determine examiner religiosity, we constructed a new database that contains each parent who had a child enrolled in high school during 1998–2016. This new parent database was obtained by merging students' files (which contain parents' identifiers) for the years 1998–2016 with the same year's school databases containing schools' religious

¹ We have data on questionnaires given in the summer session only. We define matriculation questionnaires as jointly taken by both secular and religious sectors, if the proportion of religious students that take the questionnaire is in the range [0.1, 0.9].

² Parents' country of birth is in general defined by fathers' country of birth. In case of missing values or Israeli-born fathers it is defined by mothers' country of birth.

orientation. Parents were defined as religious if at least one of their children attended a religious school.³ Since we have students' files for many years (1998–2016) we are able to determine the level of religiosity of most of the examiners in our sample according to this definition (about 85% of the examiners and 87% of the graded exam booklets).

Finally, we also develop several measures of examiners' exposure to different environments. We construct several measures of exposure to the outgroup at school. These variables were constructed at the examiner level in each year using information on examiners' peers at school from teachers' files in each year together with the parents' files. The teacher database contains information on all teachers in each school, including their demographic information and main fields of study. Therefore, merging it with parents' files enables us to compute for each teacher in a given year: (1) the proportion of peers at school from a religious background; (2) the proportion of peers at school from a religious background who teach the same subject; and (3) the proportion of peers at school from a religious background who have the same gender.

Similarly, we also compute a geographical measure of examiners' exposure to the religious outgroup each year in their neighborhood, using the proportion of religious/secular students within the examiners' zip code. We use students' and teachers' neighborhood zip codes received from the Ministry of Education which enable to characterize for each teacher's zip code in a given year the proportion of students who attended religious schools, and merge it with teachers' files for the relevant year.

Our final dataset thus consists of panel data for six years of matriculation exams between the years 2010–2015. It includes information on the matriculation exam (student, school, class, both examiners identifiers, questionnaire number, number of credits, scores given by the first and second examiners, and the "internal" exam score); the student (grades, class, and school assignment and school zip code, gender, ethnicity, number of siblings, and parents' education); the school (location, religious orientation, and whether it is a gender-segregated school); and the two examiners of each exam booklet (main field of instruction, gender, age, education and ethnicity, main school's characteristics, and peers' and neighbors' religious orientation).

³ In the Appendix we also report results using a stricter definition of religiosity, where a parent is defined as religious if all her children attended religious school. When using this stricter definition, we obtain a marginally higher ingroup bias.

Revealing Religiosity

The external exam booklets do not reveal a student's identity to the grader: they only include the student's ID number and school code. Nonetheless, the examiner can, in principle, look up the school code and find out whether the school is religious or not. Importantly, even if the examiner does not look up the school code, religious Jews write a special inscription—*BS"D*—at the top of every written document. Thus, the level of religiosity of the student is effectively revealed to the examiners. Appendix Figure A1 presents examples of first pages of religious students' notebooks with the *BS"D* (בס"ד) notation at the top of each page. The pages include Hebrew, math/science and English paragraphs.

To check the prevalence of writing *BS"D* on exam booklets we were allowed to randomly sample 442 exam booklets. The sample contains 199 booklets from a 2-credit Hebrew questionnaire exam from 2015 (100 students from religious schools and 99 students from secular schools) and 243 exam booklets from a 3-credit mathematics questionnaire exam from 2014 (119 students from religious schools and 124 students from secular schools). As expected, in 83% of the cases the religiosity of students' schools coincides with a religious *BS"D* notation (86% in math and 80% in Hebrew). The inconsistent cases are mostly due to students from religious schools who do not write *BS"D* (26% in the math sample and 39% in Hebrew), while very few students from secular schools wrote *BS"D* (3% in math and 2% in Hebrew). Appendix Table A1 presents the coefficients of balancing tests for writing *BS'D*. The dependent variable in each regression is the characteristic of the student and the explanatory variable is a dummy for religious student who wrote *BS'D* (the regression includes questionnaire FE). The first column includes all students and the second column includes religious students only. Overall, the estimates indicate that writing *BS"D* is highly correlated with the religiosity of students (first column) and that writing *BS"D* among religious students (second column) is more prevalent among female students, among students with more siblings, and among students with low parental education.

As noted above, all the exam booklets that are distributed in a specific classroom are assigned in batch to two randomly assigned examiners. Therefore, even if the examiner does not look up the school, and even if not all religious students include the *BS"D* notation, as long as most booklets from a given classroom bear the *BS"D* inscription, the examiner can easily infer that the few students in the room who did not write this inscription are also from a religious school.

Table A1: Balancing Tests for Writing BS"D

	All Students	Religious Students
	(1)	(2)
Gender (Boy = 1)	-0.175 (0.049)	-0.256 (0.071)
Number of siblings	1.119 (0.161)	0.422 (0.235)
Father's years of schooling	-0.854 (0.502)	-2.065 (0.751)
Mother's years of schooling	-0.961 (0.477)	-1.206 (0.758)
Ethnicity Asia/Africa	0.062 (0.036)	0.031 (0.053)
Ethnicity Europe/America	0.008 (0.034)	-0.099 (0.058)
Ethnicity Israel	-0.014 (0.049)	0.097 (0.073)
Ethnicity Former Soviet Union	-0.075 (0.029)	-0.038 (0.043)
Ethnicity Ethiopia	0.018 (0.018)	0.010 (0.026)
Religiosity	0.722 (0.029)	1.000 (0.000)
Gender separated school	0.619 (0.039)	(0.056) (0.059)
External Grade	0.158 (0.096)	0.070 (0.134)
Internal Grade	0.216 (0.103)	0.039 (0.143)
N	443	219

Notes: The estimates in each column are each from separate regressions which includes questionnaire FE. The dependent variable in each regression is the characteristic of the student and the explanatory variable is a dummy for writing BS"D. The first column includes the sample of all students and the second column includes religious students only. Standard errors are presented in parentheses.

Table A2: Balancing Tests for the Assignments of Students' Booklets to Examiners with Missing Values of Religious Status

Gender (Boy = 1)	0.005 (0.004)
Number of siblings	-0.010 (0.005)
Father's years of schooling	-0.027 (0.003)
Mother's years of schooling	-0.011 (0.031)
Ethnicity Asia/Africa	0.000 (0.001)
Ethnicity Europe/America	-0.000 (0.001)
Ethnicity Israel	0.001 (0.002)
Ethnicity Former Soviet Union	-0.000 (0.002)
Ethnicity Ethiopia	-0.000 (0.000)
Religious Student	-0.002 (0.003)
N	4,133,100

Notes: The dependent variable in each regression is the characteristic of the student and the explanatory variable is a dummy for missing value of religious examiner. Each regression controls for questionnaire and year fixed effects. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A3: Summary Statistics of Students' Characteristics, by Students' Religious Status

	All Students	Religious Students	Secular Students
	(1)	(2)	(3)
Proportion of Boys	0.472 (0.499)	0.376 (0.484)	0.492 (0.499)
Mean Father's Education	12.525 (4.693)	12.568 (5.339)	12.402 (4.536)
Mean Mother's Education	12.899 (4.208)	12.134 (5.173)	13.066 (3.893)
Mean Number of Siblings	1.341 (1.475)	2.250 (2.051)	0.943 (0.978)
Proportion of Asian/African Ethnicity	0.123 (0.329)	0.152 (0.359)	0.112 (0.316)
Proportion of European/American Ethnicity	0.104 (0.305)	0.140 (0.347)	0.092 (0.288)
Proportion of Israeli Ethnicity	0.641 (0.480)	0.622 (0.484)	0.646 (0.478)
Proportion of Former Soviet Union	0.112 (0.315)	0.056 (0.232)	0.131 (0.337)
Proportion of Religious Students	0.257 (0.437)	1.000 (0.000)	0.000 (0.000)
Number of Students	423,002	108,594	314,408

Notes: The sample includes students in Jewish schools who were born in Israel and took at least one matriculation test in an identical questionnaire for both the religious and secular sectors. Religious students are defined by the religiousness of the students' school (dummy=1 if the school is a religious school). Standard deviations are reported in parentheses.

Table A4: Summary Statistics of Tests Scores, by Students' Religious Status

	All Students	Religious Students	Secular Students
	(1)	(2)	(3)
External Exam Grade	70.378 (20.150)	70.050 (20.622)	70.470 (20.015)
Internal Exam Grade	80.210 (14.823)	81.870 (14.690)	79.740 (14.827)
Final Grade in the Exam (Average of the External and Internal Exams Grades)	75.410 (16.030)	76.116 (16.079)	75.216 (16.011)
Probability of Passing the Exam	0.892 (0.901)	0.890 (0.313)	0.892 (0.896)
Probability of Passing the Exam (Students with High Parental Education)	0.935 (0.246)	0.932 (0.252)	0.936 (0.245)
Probability of Passing the Exam (Students with Low Parental Education)	0.833 (0.372)	0.825 (0.379)	0.835 (0.371)
Number of Observations	2,024,537	441,953	1,582,584

Notes: External exam grades are the mean of the two examiners' scores; Internal exam grades are tests examined by the teacher of the student in each subject; Final grades are the average of the external and internal exam grades; the probability of passing the exam is the probability of receiving a grade higher or equal to 55. Low parental education is equal one if both parents have 12 or less years of schooling. Standard deviations are reported in parentheses.

Table A5: Summary Statistics of Examiners' Characteristics, by Gender and Religiosity

	All Examiners (1)	Religious Examiners (2)	Secular Examiners (3)	Male Examiners (4)	Female Examiners (5)
Proportion Male	0.173 (0.378)	0.167 (0.374)	0.175 (0.378)	1.000 (0.000)	0.000 (0.000)
Proportion Science	0.478 (0.500)	0.477 (0.499)	0.475 (0.499)	0.650 (0.478)	0.440 (0.497)
Proportion Religious	0.338 (0.473)	1.000 (0.000)	0.000 (0.000)	0.336 (0.473)	0.338 (0.473)
Proportion Ultra-Orthodox	0.111 (0.315)	0.374 (0.484)	0.000 (0.000)	0.055 (0.228)	0.122 (0.327)
Proportion who Teach in Schools Located in Segregated Religious Areas	0.030 (0.169)	0.128 (0.334)	0.000 (0.000)	0.026 (0.159)	0.030 (0.170)
Age	51.880 (9.741)	49.906 (10.574)	51.374 (9.402)	54.832 (10.509)	51.260 (9.460)
Proportion Highly Educated	0.656 (0.464)	0.571 (0.500)	0.670 (0.470)	0.689 (0.456)	0.649 (0.466)
Proportion of Asia/Africa Ethnicity	0.050 (0.218)	0.050 (0.218)	0.053 (0.223)	0.082 (0.275)	0.044 (0.205)
Proportion of Europe/America Ethnicity	0.120 (0.325)	0.163 (0.369)	0.106 (0.308)	0.119 (0.324)	0.120 (0.325)
Proportion of Former Soviet Union	0.108 (0.310)	0.057 (0.232)	0.126 (0.332)	0.159 (0.366)	0.098 (0.297)
Proportion of Israel Ethnicity	0.720 (0.449)	0.728 (0.445)	0.713 (0.452)	0.638 (0.481)	0.736 (0.440)
Number of Examiners	2,508	715	1,400	431	2,064

Notes: Religious examiners are defined by the degree of religiosity of their children school (dummy==1 if the school is a religious school). Ultra- Orthodox religious examiners are also defined by the degree of religiosity of their children school (dummy==1 if the school is an Ultra-Orthodox religious school). High educated examiners are examiners with a M.A. or a Ph.D. Note some examiners have missing values for religiosity or for gender. Standard deviations are reported in parentheses.

Table A6: Summary Statistics of Tests Examination

Number of Questionnaires Examined by each Examiner	2.230 (1.383)
Number of Booklets Examined by each Examiner	1650 (1443)
Number of Yearly Booklets per School Examined by each Examiner	12.310 (5.332)
Number of Booklets per Student	4.885 (2.778)
Total Number of Booklets	2,066,619
Number of Schools	1,003

Notes: Standard deviations are reported in parentheses.

Table A7: Number of Booklets by Religious Status and Gender of Examiners

	All Examiners			Male Examiners			Female Examiners		
	All Examiners	Secular Examiners	Religious Examiners	All Examiners	Secular Examiners	Religious Examiners	All Examiners	Secular Examiners	Religious Examiners
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Number of Booklets of Secular Students	3,234,045	1,876,601	927,710	496,996	261,299	140,996	2,710,385	1,615,302	786,714
Number of Booklets of Religious Students	899,121	511,944	273,918	130,800	65,596	40,436	761,315	446,348	233,482

Notes: Each booklet appears twice (once for each examiner).

Table A8: Balancing Tests for the Assignments of Students' Booklets to Examiners, by Examiners' Gender

	All Examiners	Male Examiners	Female Examiners
	(1)	(2)	(3)
Gender (Boy = 1)	-0.002 (0.001)	-0.002 (0.003)	-0.002 (0.001)
Number of children	-0.001 (0.003)	0.006 (0.008)	-0.001 (0.004)
Father's years of schooling	0.009 (0.013)	-0.007 (0.026)	0.010 (0.014)
Mother's years of schooling	0.001 (0.012)	0.000 (0.025)	0.001 (0.013)
Ethnicity Asia/Africa	0.000 (0.000)	-0.000 (0.001)	0.000 (0.001)
Ethnicity Europe/America	0.001 (0.001)	-0.001 (0.001)	0.001 (0.001)
Ethnicity Israel	-0.001 (0.001)	0.001 (0.002)	-0.002 (0.001)
Ethnicity Former Soviet Union	0.001 (0.001)	0.000 (0.002)	0.001 (0.001)
Ethnicity Ethiopia	0.000 (0.000)	-0.000 (0.001)	0.000 (0.000)
Religious Student	-0.002 (0.002)	-0.002 (0.005)	-0.002 (0.002)
N	3,590,116	508,324	3,081,792

Notes: The dependent variable in each regression is the characteristic of the student and the explanatory variable is a dummy for religious examiner. Column 1 includes all examiners, column 2 includes only male examiners, and column 3 include only female examiners. Each regression controls for questionnaire and year fixed effects. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A9: Estimated In-Group Biases of Examiners on Raw Test Scores, by Examiners' Gender

	All Examiners Questionnaires, Year and Student Fixed Effects	Male Examiners Questionnaires, Year and Student Fixed Effects	Female Examiners Questionnaires, Year and Student Fixed Effects
	(4)	(11)	(18)
Religious Examiner	0.204 (0.098)	0.4070 (0.256)	0.192 (0.106)
Religious Student x Religious Examiner	0.235 (0.115)	0.776 (0.309)	0.187 (0.124)
Number of Observations	3,590,116	508,324	3,081,792

Notes: See Table 1. Dependent variables are raw scores (not normalized). Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A10: Estimated In-Group Biases of Examiners on Test Scores, According to a Stricter Definition of Examiners' Religiosity Status

	Religious Examiners	Secular Examiners	All Examiners			Male Examiners	Female Examiners
	Questionnaire and Year Fixed Effects	Questionnaire and Year Fixed Effects	Questionnaire and Year Fixed Effects	Questionnaires, Year and Student Fixed Effects	Script Fixed Effects	Questionnaires, Year and Student Fixed Effects	Questionnaires, Year and Student Fixed Effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Religious Student	-0.032 (0.012)	-0.052 (0.007)	-0.052 (0.006)				
Religious Examiner			0.015 (0.008)	0.005 (0.006)	0.013 (0.005)	0.009 (0.015)	0.007 (0.006)
Religious Student x Religious Examiner			0.017 (0.014)	0.015 (0.007)	0.010 (0.003)	0.036 (0.019)	0.013 (0.007)
Number of Observations	790,439	2,390,827	3,181,266	3,181,266	3,181,266	444,017	2,733,046

Notes: This table replicates the results presented in Table 1 according to a stricter definition of examiners' religiosity that includes in the analysis only examiners who send their children to either religious or secular schools. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A11: Sensitivity of the Results to Students' Characteristics

	Boy	High Educated Mother	High Educated Father	High Number of Siblings	Israeli Ethnicity	Europe/America Ethnicity	Asia/Africa Ethnicity	Former Soviet Union	All Characteristics
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Religious Examiner	0.006 (0.006)	0.007 (0.008)	0.010 (0.008)	0.011 (0.005)	0.013 (0.006)	0.010 (0.005)	0.012 (0.005)	0.009 (0.005)	0.015 (0.014)
Religious Student x Religious Examiner	0.011 (0.006)	0.010 (0.006)	0.010 (0.006)	0.011 (0.006)	0.010 (0.006)	0.010 (0.006)	0.011 (0.006)	0.011 (0.006)	0.012 (0.006)
Student Characteristic x Religious Examiner	0.009 (0.008)	0.000 (0.001)	0.000 (0.000)	-0.001 (0.001)	-0.003 (0.003)	0.006 (0.006)	-0.010 (0.004)	0.009 (0.005)	
Number of Observations	3,590,116	3,547,780	3,541,390	3,551,430	3,590,116	3,590,116	3,590,116	3,590,116	3,496,361

Notes: The table presents the sensitivity of the ingroup bias estimate to students' characteristics. All columns present the results from separated regressions based on the preferred specification (which includes year, questionnaire, and student fixed effects), where each regression additionally includes the interaction between the dummy for religious examiner and a different student characteristic. Standard errors are corrected for clustering at the examiner level and are presented in parentheses.

Table A12: Estimated In-Group Biases of Examiners in STEM and Non-STEM Subjects, by Examiners' Gender

	All Examiners	Male Examiners	Female Examiners
	(1)	(2)	(3)
A. STEM Test Scores			
Religious Examiner	0.007 (0.007)	0.019 (0.016)	0.004 (0.008)
Religious Student x Religious Examiner	0.012 (0.007)	0.033 (0.018)	0.013 (0.008)
Number of Observations	1,652,315	320,764	1,331,551
B. Non-STEM Test Scores			
Religious Examiner	0.01 (0.007)	-0.01 (0.032)	0.02 (0.007)
Religious Student x Religious Examiner	0.010 (0.008)	0.048 (0.052)	0.006 (0.009)
Number of Observations	1,937,801	187,560	1,750,241

Notes: The table presents the in-group bias estimates of examiners, separately for STEM (Panel A) and non-STEM (Panel B) subjects. All columns present the results from separated regressions based on the preferred specification (that include year, questionnaire and student fixed effects). In the first column all examiners are included; in the second and third column the sample is stratified according to examiners' gender. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A13: Estimated In-Group Biases of Examiners on Test Scores, by Examiners' Characteristics

	Examiners' Subject of Instruction	Examiners' Age	Examiners' Education
	(1)	(2)	(3)
Religious Student x Religious Examiners	0.013 (0.009)	0.019 (0.008)	0.001 (0.010)
Religious Student x Religious Examiners x Dummy for Science/ Older/ High Educated Examiners	-0.005 (0.012)	-0.019 (0.011)	0.014 (0.012)
Number of Observations	3,590,116	3,590,116	3,590,116

Notes: All columns present the results from separated regressions based on the preferred specification (that include year, questionnaire and student fixed effects), where each regression include additionally the interaction with a different characteristic of examiners. Examiners' age is a dummy variable that equals one if the examiner is older than 51 years old, (the median of the distribution). Examiners' education is a dummy variable that equals one if the examiner is with a M.A. or a PhD. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A14: Estimated In-Group Biases of Examiners on Test Scores, by Type of Religious Examiners

	Type of Religiosity (Dummy for Ultra-Orthodox Jew)	Examiners' School is Located in Segregated Religious Areas (Dummy=1 if Segregated Religious Areas)
	(1)	(2)
Religious Student x Religious Examiner	0.017 (0.007)	0.005 (0.006)
Religious Student x Religious Examiner x Dummy for Orthodox/Religious Examiners Teaching in Segregated Religious Areas	-0.020 (0.010)	0.038 (0.016)
Number of Observations	3,590,116	3,550,108

Notes: All columns present the results from separated regressions based on the preferred specification (that include year, questionnaire and student fixed effects), where each regression include additionally the interaction with a dummy for Ultra-Orthodox/religious examiners teaching in segregated religious areas. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A15: Summary Statistics of Examiners' Exposure to Different Religiosity Environmental Measures, by Religiosity and Gender of Examiners

		All Examiners			Male Examiners			Female Examiners		
		All	Secular	Religious	All	Secular	Religious	All	Secular	Religious
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Exposure to Neighbors with a Different Religious Orientation	Mean	0.189	0.136	0.296	0.230	0.154	0.363	0.183	0.133	0.284
	Std. Deviation	(0.243)	(0.164)	(0.327)	(0.268)	(0.179)	(0.337)	(0.238)	(0.161)	(0.324)
	Median	(0.080)	(0.080)	(0.120)	(0.120)	(0.090)	(0.340)	(0.080)	(0.080)	(0.090)
	Percent of Within Examiner Variation	<i>0.075</i>	<i>0.103</i>	<i>0.019</i>	<i>0.103</i>	<i>0.114</i>	<i>0.084</i>	<i>0.077</i>	<i>0.103</i>	<i>0.024</i>
Exposure to Peers at School with a Different Religious Orientation	Mean	0.185	0.152	0.253	0.219	0.165	0.316	0.180	0.150	0.242
	Std. Deviation	(0.236)	(0.148)	(0.341)	(0.268)	(0.186)	(0.353)	(0.230)	(0.141)	(0.338)
	Median	(0.100)	(0.110)	(0.060)	(0.110)	(0.110)	(0.120)	(0.100)	(0.110)	(0.050)
	Percent of Within Examiner Variation	<i>0.175</i>	<i>0.181</i>	<i>0.163</i>	<i>0.155</i>	<i>0.187</i>	<i>0.097</i>	<i>0.181</i>	<i>0.180</i>	<i>0.181</i>

Notes: The proportion of neighbours with a different religious status is based on the proportion of religious students in the examiner's zip code in each year. The proportion of peers at school with a different religious orientation is based on the proportion of peer teachers at school in each year. Standard deviations and medians are reported in parentheses, percentages of within examiner variation are reported in italic (variation in the dummy variable for higher than the median exposure to different religious environment, by religiosity status and gender).

Table A16: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Discrete Measures of Exposure, by Examiners' Gender

	All Examiners (1)	Male Examiners (2)	Female Examiners (3)
A. High Exposure to Neighbours with a Different Religious Orientation than that of the Examiner			
Religious Student x Religious Examiners	0.011 (0.008)	0.064 (0.021)	0.004 (0.009)
Religious Student x Religious Examiners x Dummy for Exposure to a High Proportion of Neighbours with a Different Religious Orientation	-0.007 (0.012)	-0.074 (0.030)	0.004 (0.013)
Observations	3,505,201	497,811	3,007,390
B. High Exposure to Peers at School with a Different Religious Orientation than that of the Examiner			
Religious Student x Religious Examiners	0.002 (0.008)	0.043 (0.021)	-0.005 (0.009)
Religious Student x Religious Examiners x Dummy for Exposure to a High Proportion of Peers at School with a Different Religious Orientation	0.011 (0.011)	-0.032 (0.029)	0.021 (0.012)
Observations	3,590,116	508,324	3,081,792

Notes: The coefficients in each column and panels are from separated regressions that includes a dummy for different types of exposure variables and its interactions with the variables of the main specification. Each regression includes additionally year and students fixed effects and examiner by questionnaire by zip code/school fixed effects. The proportion of neighbours with a different religious orientation is based on the proportion of religious students in the examiner's zip code in each year. The proportion of peers at school with a different religious orientation is based on the proportion of peer teachers at school in each year. The dummy variables for high exposure equal one if the proportion of neighbours or peers of the examiner is higher than the median of each group (by religiosity and gender). Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A16: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Discrete Measures of Exposure, by Examiners' Gender - Continued

	All Examiners (1)	Male Examiners (2)	Female Examiners (3)
C. High Exposure to Peers at School with a Different Religious Orientation than that of the Examiner but who Teach the Same Subject			
Religious Student x Religious Examiners	0.012 (0.007)	0.050 (0.020)	0.008 (0.007)
Religious Student x Religious Examiners x Dummy for Exposure to a High Proportion of Same Subject Peers at School with a Different Religious Orientation	-0.018 (0.012)	-0.050 (-0.032)	-0.014 (0.013)
Observations	3,485,422	498,185	2,987,237
D. High Exposure to Peers with a Different Religious Orientation than that of the Examiner but of the Same Gender at School			
Religious Student x Religious Examiners	0.002 (0.008)	0.052 (0.022)	-0.006 (0.009)
Religious Student x Religious Examiners x Dummy for Exposure to a High Proportion of Same Gender Peers at School with a Different Religious Orientation	0.010 (0.011)	-0.050 (0.030)	0.022 (0.012)
Observations	3,590,116	508,324	3,081,792

Notes: The coefficients in each column and panels are from separated regressions that include a dummy for religious student and a dummy for religious teacher and their interactions with the different types of exposure variables. Each regression includes additionally year and students fixed effects and examiner by questionnaire by zip code/school fixed effects. The proportion of neighbours with a different religious orientation is based on the proportion of religious students in the examiner's zip code in each year. The proportion of peers at school with a different religious orientation is based on the proportion of peer teachers at school in each year. The dummy variables equal one if the proportion of neighbours or peers of the examiner is higher than the median of each group (by religiosity and gender). Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A17: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Continuous Measures of Exposure, by Examiners' Gender

	All Examiners (1)	Male Examiners (2)	Female Examiners (3)
A. Exposure to Neighbors with a Different Religious Orientation than that of the Examiner			
Religious Student x Religious Examiners	0.016 (0.008)	0.073 (0.022)	0.009 (0.008)
Religious Student x Religious Examiners x Proportion of Neighbors with a Different Religious Orientation	-0.033 (0.028)	-0.153 (0.067)	-0.007 (0.030)
Observations	3,505,201	497,811	3,007,390
B. Exposure to Peers at School with a Different Religious Orientation than that of the Examiner			
Religious Student x Religious Examiners	0.008 (0.008)	0.034 (0.021)	0.005 (0.009)
Religious Student x Religious Examiners x Proportion of Peers with a Different Religious Orientation	0.000 (0.029)	-0.013 (0.060)	0.005 (0.033)
Observations	3,590,116	508,324	3,081,792

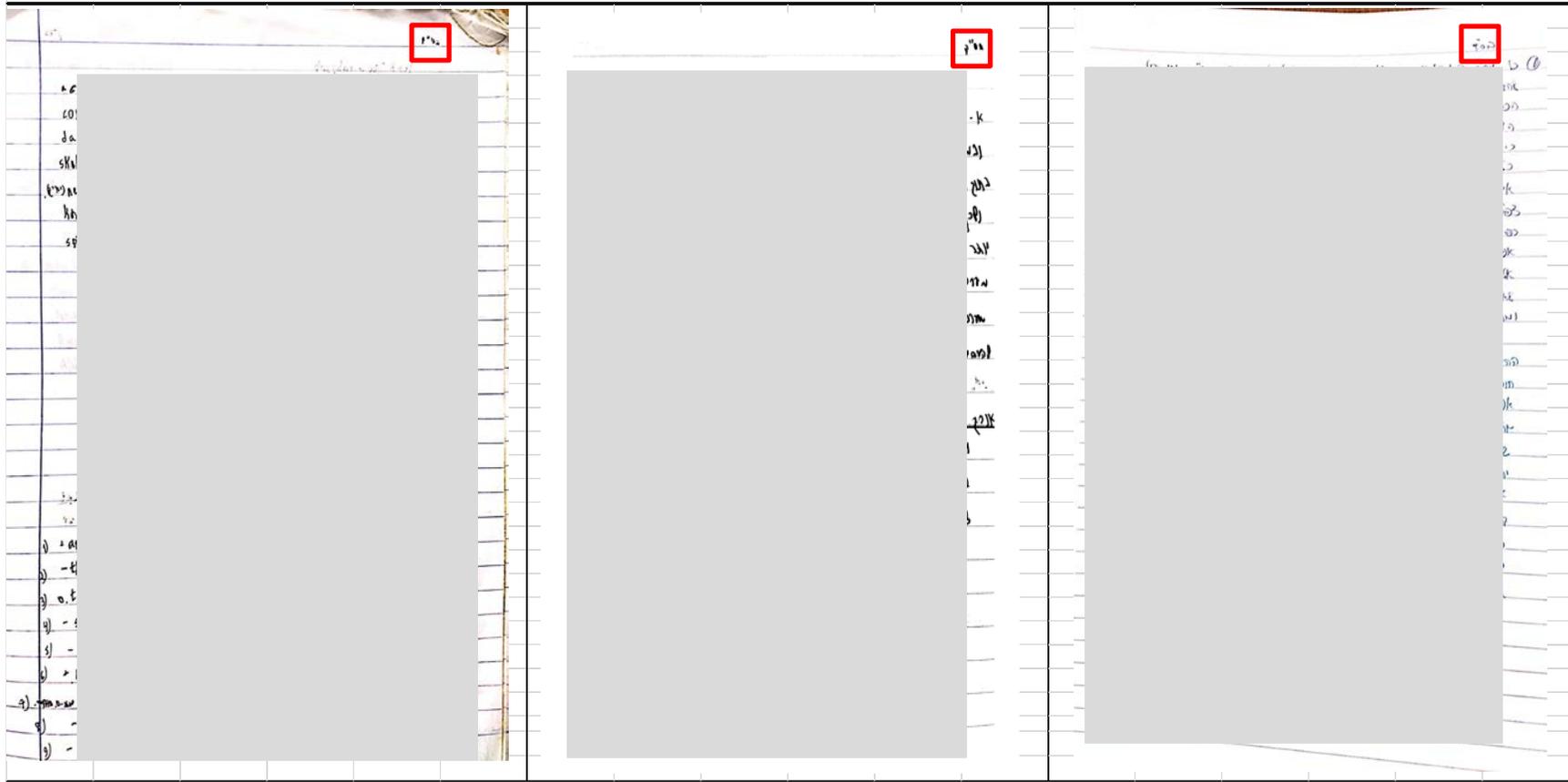
Notes: See Table A16. The table presents the estimated effects of exposure to a different religious environment on the in-group biases of examiners based on continuous variables of exposure to different religious environments instead of discrete variables as in Table 10. Standard errors are corrected for examiners clustering and are presented in parentheses.

Table A17: Estimated Effect of Exposure to a Different Religious Environment on In-Group Biases of Examiners- Continuous Measures of Exposure, by Examiners' Gender - Continued

	All Examiners (1)	Male Examiners (2)	Female Examiners (3)
C. Exposure to Peers at School with a Different Religious Orientation than that of the Examiner but who Teach the Same Subject			
Religious Student x Religious Examiners	0.012 (0.007)	0.045 (0.018)	0.007 (0.007)
Religious Student x Religious Examiners x Proportion of Same Subject Peers at School with a Different Religious Orientation	-0.025 (0.025)	-0.100 (0.067)	-0.006 (0.027)
Observations	3,590,116	498,185	2,987,237
D. Exposure to Peers with a Different Religious Orientation than that of the Examiner but of the Same Gender at School			
Religious Student x Religious Examiners	0.008 (0.008)	0.038 (0.021)	0.005 (0.008)
Religious Student x Religious Examiners x Proportion of Same Gender Peers at School with a Different Religious Orientation	0.001 (0.028)	-0.045 (0.061)	-0.088 (0.156)
Observations	3,590,116	508,324	3,081,792

Notes: See Table A16. The table presents the estimated effects of exposure to a different religious environment on the in-group biases of examiners based on continuous variables of exposure to different religious environments instead of discrete variables as in Table A16. Standard errors are corrected for examiners clustering and are presented in parentheses.

Figure A1: Sample of Religious Students' Booklets



Notes: The inscription is 'BS"D' (בס"ד in Hebrew)-- acronym for Besiyata Dishmaya, an Aramaic phrase, meaning "with the help of Heaven". Religious Jews write this notation at the top of every page in a written document as a reminder to them that all comes from God.