Do Children Perform Better in Religious Schools? Evidence From Population Data

By DENI MAZREKAJ AND CHRISTIAAN MONDEN*

Religious schools enjoy a high academic reputation among the parents in many societies. Previous studies that assessed the effect of religious schools were conducted in countries where religious schools are private or where they charge fees and set admission criteria. As a result, the effect of religious schooling could not be separated from the effect of private schooling and selection. We contribute to the literature by using Dutch data that include the entire population of children born between 1998 and 2007. In the Netherlands, both public and religious schools are publicly funded, schooling is free of charge and admission is independent of the child's religious or ideological character. Using a range of models including fixed effects models, coarsened exact matching, and treatment effect bounds, we compare school outcomes of children in religious versus public schools. Our results indicate that children in religious schools outperform children in public schools in primary education. The benefits of religious schooling were largest for children in Orthodox Protestant, Islamic and Hindu schools, which mostly attract children from a disadvantaged socioeconomic background. However, the influence of religious schooling fades out by the end of secondary education.

*Deni Mazrekaj: Department of Sociology, University of Oxford, Park End Street 42-43, OX1 1JD Oxford, United Kingdom (e-mail: deni.mazrekaj@sociology.ox.ac.uk); Christiaan Monden: Department of Sociology, University of Oxford, Park End Street 42-43, OX1 1JD Oxford, United Kingdom (e-mail: christiaan.monden@sociology.ox.ac.uk). The data for this study are protected by a confidentiality agreement and we are precluded from sharing the data with others. Interested readers can consult the online supplement for information on how to obtain access to the data. We would be happy to provide assistance and Stata code to replicate the results of this study. The project leading to this research received funding from the European Research Council [grant number 681546: FAMSIZEMATTERS]. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

Keywords: Religious Schools, Population Data, School Outcomes, Coarsened Exact Matching, Treatment Effect Bounds.

1. Introduction

Whether state should subsidize private education is a highly debated issue in the United States. On June 30, 2020, the U.S. Supreme Court overturned the prohibition of state-funded scholarships and school vouchers for religious schools, narrowing the separation between church and state (Espinoza et al. v. Montana Department of Revenue et al. 2020). When this issue was being contested, the parents argued that they should be able to direct the religious upbringing of their children by enrolling them into private religious schools. Often, these private religious schools enjoy a high academic reputation, although this is not necessarily the case, especially for smaller religious denominations. For instance, in 2015, the New York City Department of Education opened an investigation into ultra-Orthodox Jewish private schools (*yeshivas*), based on concerns raised by parents, former students, and teachers that these schools do not provide adequate math and science education to the students (Deen 2018, Taylor 2015). Similar concerns have been raised in Europe, mainly about the Islamic schools (Richardson 2014, RTL Nieuws 2019).

Since the seminal works by Coleman, Hoffer, and Kilgore (1981, 1982), many studies have examined the influence of religious schools—almost exclusively Catholic schools—on children's educational outcomes (see Jeynes (2012) for a meta-analysis of over 90 studies). By and large, these early studies suggest that children in religious schools outperform children in public schools. However, most studies were conducted in countries where religious schools are all private, e.g. United States and Australia, or where religious schools charge fees and set admission criteria such as church attendance or a recommendation from a local minister, e.g. United Kingdom. As a result, the effect of religious schooling is intertwined with the effects of private schooling and selection. A more recent literature that used more rigorous methods concluded that the seeming religious school advantage occurs entirely through sorting of advantaged children into religious schools (Altonji, Elder and Taber 2005b, Elder and Jepsen 2014, Gibbons and Silva 2011, Gihleb and Giuntella 2017, Nghiem, et al. 2015).

In this study, we shed light on the consequences of being enrolled in religious schools by leveraging unique administrative population data from the Netherlands. Dutch data are particularly suited to empirically assess children's outcomes in religious schools as we can examine how things have played out in a nation in which the state has been funding religious schools already since 1917. Specifically, the distinction between public and private schools in the Netherlands is purely administrative; both public and private schools receive financial support from the state. School funding is mainly driven by the share of disadvantaged children per school, irrespective of the

school's religious or ideological character. Moreover, primary and secondary education is free of charge and parents may freely choose a school for their children. Our data include the entire population of children born between 1998 and 2007, and their educational performance can be tracked until 2019. We compare scores on a high-stakes standardized test in the final year of primary education between children in public schools and children in religious schools using a range of models: fixed effects models, multilevel models, coarsened exact matching, cousin fixed effects, and treatment effect bounds. For about one third of the children, we also analyze diploma attainment in secondary education and enrolment into tertiary education.[1] In addition to Catholic schools, we include mainstream Protestant, Orthodox Protestant, Jewish, Islamic and Hindu schools. Observing schools other than Catholic schools is useful because different school denominations differ in religiosity and may attract a different population of students. In the Netherlands, Catholic, mainstream Protestant and Hindu schools are typically less pronounced in their religiosity than Orthodox Protestant, Jewish and Islamic schools. Moreover, children in Catholic and mainstream Protestant schools largely have a similar socioeconomic status as children in public schools, whereas children in Jewish schools tend to have wealthier and more educated parents. On the other hand, children in Orthodox Protestant (Evangelical, Baptist, Reformatory, and Remonstrant), Islamic and Hindu schools mainly attract children from a disadvantaged socioeconomic background.

Our results indicate that children in religious schools outperformed children in public schools on standardized tests in primary education. The benefits of religious schooling were largest for children in Orthodox Protestant, Islamic and Hindu schools, who scored about 0.1 to 0.25 standard deviations higher than children in public schools. We found no heterogeneity by gender. These results could not be explained by sorting of children into religious schools as children in Orthodox Protestant, Islamic and Hindu schools had lowest socioeconomic status in the population. Our results mostly support the social capital model and discrimination theory. Specifically, children in religious schools may benefit from the social capital provided by the larger religious communities. In addition, religious schools may serve a protective role against discrimination for minority children. In the longer run, the influence of religious schooling faded out. Once we accounted for selection effects, children in religious schools were just as likely as children in public schools to graduate from high school and to enrol into tertiary education.

[1] The other children were too young to have graduated from high school.

2. Background

2.1. Previous Empirical Evidence

A sizeable literature has studied the association between religious schools and children's academic outcomes already since the early 1960s (Bressler and Westoff 1963, Lenski 1961) and popularized in the works by Coleman, Hoffer, and Kilgore (1981, 1982). Overall, this literature suggests that children in religious schools outperform children in public schools in both primary and secondary education. In a meta-analysis of 54 studies examining the effectiveness of religious schools, Jeynes (2012) found a religious school advantage of about .15 standard deviations in primary education and a slightly higher advantage of about .17 standard deviations in secondary education.[2] Religious schools were particularly beneficial for ethnic minority children who had .39 standard deviations higher achievement in religious schools rather than public schools.

Although the consensus appears to be that children in religious schools outperform children in public schools, the large majority of the studies were conducted in countries where religious schools are all private. For instance, from the aforementioned 54 studies included in the meta-analyis by Jeynes (2012), 80% were conducted in the United States, 9% in Australia, and another 6% in other countries where religious schools are privately funded. This is problematic because private schooling may exert an independent positive effect on children's outcomes (Carbonaro and Covay 2010, Epple, Romano and Urquiola 2017). Therefore, the effects of private schooling and religious denomination are intertwined. Other studies have used data from the United Kingdom where religious schools are publicly funded (Gibbons and Silva 2011, Sullivan, et al. 2018). However, religious schools in the United Kingdom often charge fees and set admission criteria such as attendance at a local church or a recommendation from a local minister (Gibbons and Silva 2011). This may lead to selection effects because religious schools with fees and admission criteria may attract children with different observed and unobserved characteristics than public schools with no fees and admission criteria. Thus, the population of children in the two types of schools may be different due to sorting of children into different schools and a comparison of outcomes is not straightforward.

Several attempts have been made to separate the effect of religious denomination from the effects of private schooling and selection. Early studies used an instrumental variables framework with instruments such as family religion (Evans and Schwab 1995, Neal 1997, Noell 1982), the

[2] These effect sizes were calculated after controlling for socioeconomic status and when using standardized tests. Without these restrictions, the effect sizes were about .26 standard deviations.

geographical density of religious schools or distance to religious schools (Figlio and Stone 1999, Grogger and Neal 2000, Neal 1997), or interactions of these instruments (Sander 1996). However, the analysis in Altonji, Elder, and Taber (2005a) shows that these instruments provide biased results due to violation of the exclusion restriction. Namely, family religion may influence children's school outcomes independently of the school denomination[3] (see Pearce, Uecker, and Lundquist Denton (2019) for an overview of the literature), and parents may decide to live near religious schools because they want their children to go to religious schools, especially if these schools are private (Bibler and Billings 2020, Candipan 2020, Pearman and Swain 2017). A more recent study used the decline in the number of Catholic sisters between 1966 and 1980 in the United States as an instrument (Gihleb and Giuntella 2017). This decline occurred as a result of the Second Vatican Council and was followed by a drop in religious teachers, forcing many Catholic schools to close. Although innovative, it is unclear whether the decline in the number of nuns affected solely the number of Catholic schools or also the quality. If the latter is true, the exclusion restriction would be violated.

Given the difficulty in finding appropriate instruments, other studies have used alternative methods. In the United States, Jepsen (2003) used value-added models to control for children's background, whereas Altonji, Elder, and Taber (2005b), Elder and Jepsen (2014), and Gihleb and Giuntella (2017) used a bounding approach in which they inferred the degree of selection bias from the selection on observable characteristics. Nghiem and colleagues (2015) applied both these methods on Australian data. Similary, Gibbons and Silva (2011) used both a value-added and a bounding approach in England, but also conducted two additional analyses. First, they compared children in religious primary schools with children in public primary schools while controlling for children's secondary school enrolment. The assumption is that the choice of high school (either a public or a religious school) conveys information about selection into a primary school. Their second analysis includes a bounding analysis in which they (1) construct a lower bound by comparing children that went to a religious primary school but a public high school, with children that went to a public primary school but a religious high school, and (2) construct an upper bound by comparing children that went to a religious school with children that went to a public school in both primary and secondary education. All the studies using a value-added model and/or a bounding analysis suggest that children in public

[3] We will refer to the public versus religious dichotomy as "school denomination" (Gibbons and Silva 2011) to stress that we are primary interested in the religious character of schools. Another commonly used term is "school sector" (Coleman, Hoffer and Kilgore 1982), but may refer more to the public versus private distinction.

schools perform just as well as children in religious schools. These studies argue that the seeming religious school advantage occurs through sorting of children into religious schools according to pre-existing characteristics and preferences.

An important limitation of the studies that used more advanced approaches to account for selection bias is that these studies solely focused on Catholic schools. Although Catholic schools are often most represented among the private schools, a large share of children attend schools with other denominations. For instance, of the 4.4 million children enrolled in private religious schools in the United States in 2015, only 47% were enrolled in Catholic schools (National Center for Education Statistics 2019). Thus, most children were not included in the analyses. In addition, Catholic schools are most prone to selection bias as these schools enjoy a high academic reputation among the parents and mostly attract children from a high socioeconomic background. Other religious schools, such as Islamic schools for instance, typically comprise of children from a disadvantaged background (De Witte and Van Klaveren 2014) and the selection bias is likely to operate in the opposite direction than the selection bias on Catholic schools. As a result, a positive influence of Islamic schools could be considered a better proxy of the true positive effect of religious schools than a positive influence of Catholic schools.

Several studies have analyzed countries in which religious schooling is not privately funded and children may choose schools freely, such as Belgium and the Netherlands. Using Belgian (Flemish) data on 1,795 children, Brutsaert (1998) found that children in Catholic schools outperform children in public schools. However, the study used children's self-reported grade point average scores in Grade 6, rather than standardized tests. These scores are subjective and may be prone to bias. Building on the Brutsaert's (1998) study, Agirdag, Driessen, and Merry (2017) analysed 5,069 Flemish children and found no support for the Catholic school advantage. In terms of prior work, Driessen, Agirdag, and Merry's (2016) paper comes closest to our study. Using multi-level models on 27,457 children in grades 2, 5, and 8 in the Netherlands, they found no substantial differences in outcomes between public schools on the one hand and Catholic and Protestant schools on the other. However, their results indicated that Islamic schools had a large value-added potential. Children in Islamic schools had low raw scores, but after controlling for child and school characteristics, children in Islamic schools performed just as well as children in public schools. We build on this prior work by (1) including the entire population of children in the Netherlands born between 1998 and 2007 and enrolled in schools with all denominations (also Jewish and Hindu schools), (2) using a range of

models including a novel bounding analysis by Oster (2019), and (3) also analysing outcomes beyond primary education such as diploma attainment and tertiary education enrolment.

2.2. Theoretical Mechanisms

Several mechanisms have been proposed to explain the apparent religious school advantage. The main sociological argument draws on Coleman's social capital theory (Coleman 1988, Coleman, Hoffer and Kilgore 1981), with social capital being characterized by relations among individuals. Religious schools may be embedded in "functional communities" with considerable reciprocal trust among individuals. Parents may know their children's friends, parents, teachers, and school principals, creating a type of unity that facilitates learning. Being part of a religious community, teachers may for instance offer remedial courses to children after school hours, or assign more homework than teachers in public schools (Coleman and Hoffer 1987). In the Netherlands, there appears to be a discrepancy in the extent to which people from different religions form communities. Specifically, communities are not particularly formed by Catholic, mainstream Protestant, and Hindu individuals, whereas Orthodox Protestant, Islamic, and Jewish groups are likely to form closed communities.

Language of instruction may also be closely related to the distribution of social capital. In religious minority schools such as Islamic schools where children may need help in understanding course material in the country's official language, teachers may explain difficult concepts in children's native language. Nonetheless, Islamic schools in the Netherlands mostly attract Moroccan and Turkish children, who do not speak the same language and who seldom have Arabic as their first language. Thus, teachers may be able to help only a proportion of students in their native language. Public schools may lack this community trustworthiness and the relationship between parents and teachers may be detrimental to disadvantaged children. For instance, McCrory Calarco (2020) found that U.S. public schools tend to appease privileged families. Teachers in public schools selectively enforced rules, disproportionately punishing and negatively evaluating children from a disadvantaged background, while giving leeway to children of advantaged White "helicopter parents." Thus, there may exist a favourable discrepancy in social capital for religious rather than public school children as reflected by the religious community network and support. On the other hand, the downside of this social capital may be that excessive monitoring represses creativity and renders children unequipped for the disorder of the larger society (Portes 1998). As a result, children in religious schools may actually perform worse on standardized tests (Morgan and Sørensen 1999). A testable implication of

the social capital theory is that children from separated or single parents especially benefit from religious schools. As these children are likely to have a lower social capital at home than children with married parents, the social capital resulting from religious schooling may compensate for the lack of social capital at home. A similar implication holds for children who moved neighbourhoods or schools. Moreover, children's social capital may also be related to the number of relatives that are enrolled in the same school. Thus, we speculate that children with separated and single parents as well as children who moved neighbourhoods or schools will disproportionately benefit from religious schools compared to children with married parents. We also expect that children in religious schools will have more relatives going to the same school.

Religious schools may also lead to better school outcomes if they have more resources to spend on school inputs such as smaller classes and better teachers (Jackson, Johnson and Persico 2016). However, this mechanism mainly applies to countries such as the United States or Australia where religious schools are private. In the Netherlands, all schools receive the same funding from the state based on the proportion of disadvantaged children, irrespective of the school's religious or ideological character. The teacher-to-pupil ratio as reflected in class size is therefore unlikely to drive any potential religious school advantage. On the other hand, quality teachers who motivate children, instill better study habits, and produce high school outcomes may still sort to religious schools. However, this argument would primarly hold for Catholic schools, and not necessarily for religious schools with a large share of disadvantaged minority children such as Orthodox Protestant, Islamic and Hindu schools. As quality teachers are likely to flock towards schools with a large share of advantaged children (Goldhaber, Quince and Theobald 2017), we do not expect the school resources hypothesis to hold much weight in our setting.

A final reason to expect that children in religious schools may perform better than children in public schools is given by discrimination theory. Religious schools may be offering more equal educational opportunities for children of different race and ethic origin than public schools (Coleman, Hoffer and Kilgore 1982). Said otherwise, religious schools may segregate children less by race and ethnic origin that public schools, fostering positive peer effects that increase school outcomes (Billings, Deming and Rockoff 2014). This argument is supported by Jeynes's (2012) meta-analysis that found that African American and Latino children in religious schools rather than public schools had a .39 standard deviations higher achievement. The inclusion of Islamic schools in our setting provides a unique testable implication of the discrimination theory: Islamic terrorist attacks may have changed the association between Islamic school enrolment and children's school outcomes. If discrimination

theory drives religious school advantage, we would expect a more positive association after terrorist attacks than before. The intuition behind this is that Islamic children may be more discriminated against in public schools after terrorist attacks. For instance, after the 2017 Westminster attack in the United Kingdom, a councelling service for children *Childline* revealed that children as young as nine were being accused of being associated with IS and being threatened with violence (Bulman 2017). Some children even resorted to self-harm and skipped school to escape the bullying. Therefore, we would expect that after terrorist attacks, the protecting effect of religious schools would be higher and the association between religious schooling and achievement larger. Although large terrorist attacks were absent in the Netherlands in our study period (2010-2019), three coordinated suicide bombings by the Islamic State (IS) occurred in the neighboring country Belgium, two at the Brussels airport and one at a metro station in central Brussels in 2016 (Witte, Mekhennet and Birnbaum 2016). In total, 32 civilians were killed and more than 300 people were injured. We believe these terrorist attacks also had an impact in the Netherlands. There was wide media coverage of these events in the Netherlands and the Dutch prime minister increased anti-terrorist measures following the attacks (Nederlandse Omroep Stichting 2016). Five men were arrested for throwing a Molotov cocktail at a mosque in the Netherlands later that year, with about 30 people inside including children (Times NL 2016).

3. Data

3.1. Sample Construction

We use administrative records collected by Statistics Netherlands that cover the entire population of children born between 1998 and 2007. These children mainly started primary education at age 4, finishing at age 12. Each child is given a unique personal identification number in the education registers, allowing us to follow children through their entire educational career until 2019. Each year, we observe in which school a child was enrolled and the school's denomination. In the final year of primary education (from 2010 to 2019), we observe children's scores on a standardized test, the type of test, and the exact date of the test as well as teachers' recommendation on which track to follow in secondary education (see Educational Outcomes section for further explanation). We also observe whether and when children dropped out of school and whether they obtained a high school diploma. As we follow children annually, we can distill whether students enrolled in tertiary education by 2019.

The administrative records also include identifiers for children's parents, stepparents, adoptive parents, and grandparents as well as their addresses, allowing us to construct households of individuals living at the same address. As we observe parental and grandparental identifiers, we can further identify children's siblings and cousins. Moreover, the data include a variable for the highest education obtained for each individual, allowing us to observe highest education in the household. Linking the households with marital status and tax registers, we can further determine the marital status and income in a household. Finally, linking individuals to demographic registers, we observe the date of birth, sex, and birth country.

To compare school outcomes of children in religious versus public schools, we restrict the sample in three ways. First, we remove children who were not old enough to conduct the test in the final year of primary education (typically at age 12) and who were exempt from taking the test (e.g. children with a severe handicap, see Educational Outcomes section). Then, we remove children (1.74% of the sample) who went to a heterogeneous group of atypical schools such as Steiner and Humanist schools as these schools do not have a clear religious denomination. Finally, we remove children with missing information on any of the variables (2.92% of the sample). The final sample includes 1,559,678 children born between 1998 and 2007.

3.2. Children in Religious Schools

The variable of interest is the school denomination. This variable includes seven categories: *public school (no denomination)*, *Catholic school*, *mainstream Protestant school*, *Orthodox Protestant school*, *Islamic school*, *Jewish school*, and *Hindu school*. We make a distinction between mainstream Protestant and Orthodox Protestant schools as these schools may be very different in practice. As Orthodox Protestant schools, we consider Evangelical, Baptist, Reformatory, and Remonstrant schools. The language of instruction in all schools is Dutch and the curriculum is analogous for all schools as set by the Dutch Ministry of Education.

We consider school denomination of the primary school in which the child took the standardized test. Note, however, that this is the school in the final year of primary education. It is possible, therefore, that a child first went to a public school and then changed to a religious school in the last year, or vice-a-versa. This would obscure the effect of religious schooling as treatment and control groups would be contaminated. We observe that 84% of the children did not switch school denominations and spent the entire six years of primary schooling in the same school denomination. In **Table S1** in the online supplement, we repeat the analyses solely for these children. We also

conduct the analyses for children who not only went to the same school denomination throughout primary schooling, but in addition went to one particular school only. Alternatively, we also use control variables for school denomination switching, and for school switching. **Table S1** shows that these results are very similar to the main results.

We observe a total of 1,559,678 children; of which 35.62% (555,572 children) are in public schools, 35.14% (548,025 children) in Catholic schools, 24.64% (384,280 children) in mainstream Protestant schools, 3.83% (59,738 children) in Orthodox Protestant schools, 0.02% (361 children) in Jewish schools, 0.62% (9,737 children) in Islamic schools, and 0.13% (1,965 children) in Hindu schools. Although we include several hundreds of children in Jewish schools, these children went to only a few different schools.[4] Therefore, these results should be interpreted with caution.

3.3. Educational Outcomes

The main outcome of interest is the score on a high-stakes standardized test in primary school. Primary education in the Netherlands is compulsory from the age of five and lasts seven years, until age 12. Most parents, however, enrol their children into a primary school at the age of four, leading to effectively eight years of primary schooling. In the final year, children take a test that is considered as high-stakes by children and parents, and is also used for school evaluations (Scheerens, et al. 2012).[5] It is an objective measure of children's achievement graded by a computer, and all children within a school conduct the same test. Although schools may choose from several (often four to six) standardized tests prescribed by the Dutch Ministry of Education, most schools use the Central End Test (successor to the Central Institute of Test Development - CITO test), which includes 200-240 multiple choice questions testing children's knowledge of the Dutch language, mathematics, world orientation (comprising geography, biology, and history), and information processing. All tests are set in Dutch. In our sample, 79% of the children took the Central End Test. However, these tests are highly comparable and we use all tests in our analyses to maximize sample size. To make test scores comparable across different tests and years, we construct z-scores by test and year, and interpret test

[4] For confidentiality reasons, Statistics Netherlands prohibits specifying the exact number of schools if the number is lower than eight.

[5] It is possible for children to repeat a grade in primary education and conduct the test several times. Each year, the test is similar in difficulty but the exact exercises change. For children that took the test more than once (2.5% of the sample), we consider the latest test. Our results are robust to excluding these children from the analyses (see **Table S2** in the online supplement).

scores in standard deviations as common in the literature. Nonetheless, we also conduct the analyses using only the Central End Test to exclude the possibility that tests vary in difficulty and religious schools choose easier tests. The results are displayed in **Table S2** in the online supplement and are very similar to the main results.

Before 2015, schools could choose whether to participate in standardized testing, although in practice, almost all schools participated. For instance, in 2010 (the first test year in our data), 92% of schools participated (Central Institute of Test Development 2010). Starting from 2015, standardized tests have been made mandatory for all children, with two exceptions: (1) children with a severe handicap that prevents them from learning effectively, and (2) children who have not lived in the Netherlands in the last subsequent four years prior to taking the test and are lagging behind in Dutch language. In our sample, we include all children that have taken a test between 2010 and 2019. Nonetheless, we also separately analyze children who took the test before and after 2015 to account for the possibility that poor-performing religious schools chose not to administer any tests. Similarly, we also conduct the analyses without children in special education and children who have not lived in the Netherlands in the last subsequent four years prior to taking the test in case religious schools included more of these children than public schools. Our results are robust to these additional sample restrictions (see **Table S2** in the online supplement).

Secondary education in the Netherlands is compulsory until age 18 or younger if a student obtains a high school diploma earlier. Students can choose between three main tracks: pre-vocational (VMBO), general (HAVO), and pre-university (VWO) track. Students enter a track based on the primary school teacher's recommendation and the standardized test score. Typically, high ability students enrol into the pre-university track, which is considered the most prestigious. In practice, teacher's recommendations closely follow standardized test scores. There is evidence, however, of some degree of teacher bias in which teachers give less favorable recommendations to children from a lower socioeconomic background, independently of achievement on the standardized test (Timmermans, et al. 2018). In **Table S3** in the online supplement, we repeat the analyses using teacher's recommendation as an outcome. We code this variable as an indicator with a value of 1 if the primary school teacher advised the pre-university track, and 0 otherwise. We do not show these results in the main analyses because teacher's recommendation is a subjective assessment of the child's achievement. Nonetheless, the conclusions remain the same when we replace standardized test scores by teacher's recommendations.

For the earlier cohorts—children born between 1998 and 2000—we also observe diploma attainment in secondary education and enrolment into tertiary education. Other children were too young to have graduated from high school. Diploma attainment is measured as an indicator with a value of 1 if the student obtained an upper-secondary diploma by 2018 according to the International Standard Classification of Education framework (i.e., qualification at ISCED 2011 level 3), and 0 if the student was a high school dropout. Enrolment into tertiary education is measured as an indicator with a value of 1 if the student was enrolled into tertiary education by 2019, and 0 otherwise.

3.4. Control Variables

At the child level, we control for the child's sex (1 is boy, 0 is girl), year of birth, and origin. The latter is constructed based on parental or grandparental birth country and includes three categories: *Dutch*, *Moroccan or Turkish*, and *Other*. For instance, if a child has at least one parent or grandparent born in Morocco, this child is categorized as of a Moroccan origin. We distinguish between Moroccan and Turkish origin on the one hand and Other origin on the other hand, because Morrocan and Turkish children are among the most common minority children in the data. The origin variable is especially necessary when considering the influence of religious minority schools such as Islamic schools, that primarily include ethnic minority children (Agirdag, Driessen and Merry 2017, De Witte and Van Klaveren 2014). Therefore, it is important to compare ethnic minority children in religious schools with ethnic minority children in public schools.

We further construct variables to proxy for a child's socioeconomic status. Although schools in the Netherlands all receive public funding and there are no fees, different schools may still attract children from different backgrounds. For instance, as in the United States and the United Kingdom, Catholic, mainstream Protestant and Jewish schools may attract children from a higher socioeconomic status than public schools (Darnell and Sherkat 1997, Gibbons and Silva 2011, Sander 2010), whereas the reverse may be true for Islamic and Hindu schools (Agirdag, Driessen and Merry 2017, De Witte and Van Klaveren 2014), reflecting the lower socioeconomic position of their communities. We include three variables: household education, household income, and municipality of residence. Parental education is an indicator for the highest education of the child's guardians. These may be biological parents, stepparents, adoptive parents or foster parents. Parental education is constructed based on the International Standard Classification of Education (ISCED) 2011 classification and includes four categories: *less than high school (ISCED 0-2)*, *high school diploma (ISCED 3-4)*, *bachelor degree (ISCED 5-6)*, and *master degree or PhD (ISCED 7-8)*. We also include a continuous measure for

household net income in euros based on the income of the child's guardians. We take the logarithm of household income to smoothen the income distribution and to interpret the coefficients in percentages. Finally, we include indicators for the municipality of residence (postcode). Gibbons and Silva (2011) have shown that some of the disparities between religious and public schools can be explained by differences in geographical setting.

As a last set of control variables, we include five socio-demographic characteristics. First, we construct three continuous variables: mother's age at birth, the number of children in the household, and birth order. Whereas birth order is identified based on the mother and father identifiers, we defined the number of children in the household as individuals under 18 years old, thus also potentially cousins. This is to proxy for the number of children in the household over which the household resources are divided. Then, we construct a categorical variable for the household structure. It includes five categories: *married parents*, *cohabiting parents*, *parent and stepparent*, and *single parent*. Individuals are considered married based on a either a marriage certificate or a registered/civil partnership certificate. We do not distinguish between the two as registered partnership is an almost perfect substitute for marriage in the Netherlands. Cohabitation is determined based on either a cohabitation agreement or by not being married, in a registered partnership, or single. Finally, we also include a continuous measure for the number of household transitions, defined as the number of movements between the four household structures.

Control variables are measured in the year of the standardized test. However, it is possible that religious schooling may affect some of the control variables, leading to post-treatment bias (Rosenbaum 1984), also known as collider bias (Elwert and Winship 2014) or bad control bias (Angrist and Pischke 2009). For instance, if the parents want their children to take the test in a religious school in the final year of primary education, they may decide to move closer to a religious school (Bibler and Billings 2020, Candipan 2020, Pearman and Swain 2017), affecting the municipality control variable. To exclude post-treatment bias, we also conduct the analyses with control variables measured at the start of primary education. These analyses are shown in **Table S3** in the online supplement and yield almost identical results as the main analyses. Further, Gibbons and Silva suggest controlling for the secondary school, essentially comparing children in religious versus public schools in primary education, but who went to the same school or school denomination in secondary education. We refrain from using this model in the main analyses as controlling for secondary school may lead to post-treatment bias. Nonetheless, the results in **Table S3** show that our main results are robust to this additional control variable.

4. Methods

4.1. Postcode Fixed Effects

We estimate school outcomes of children in religious versus public schools by a fixed effects model. We follow Gibbons and Silva (2011) and include fixed effects at the postcode level as follows:

$$(1) \qquad y_i = \alpha_0 + \beta_1 R_i + \boldsymbol{\theta X_i} + \delta_i + \varepsilon_i$$

where y_i is the outcome variable of child i (i.e., the standardized test score at the end of primary education, diploma attainment in secondary education, and tertiary education enrolment).[6] The variable of interest is R_i, representing a categorical variable for a child in religious versus public school. The parameter of interest is β_1, representing the difference in school outcomes between children in different types of schools. Depending on the specification, **Equation (**1 also includes control variables defined above as part of $\boldsymbol{X_i}$ and δ_i are the postcode fixed effects. In **Table S4**, we also conducted a multilevel model to quantify the size of the municipality effects at the cost of imposing the additional assumption of random sorting across municipalities. The main conclusions remained unchanged. We follow the advice in Cameron and Miller (2015) and cluster standard errors at the highest level, i.e. the postcode level to account for dependence of observations within municipalities. Nonetheless, our results are robust to clustering standard errors at the school level (see **Table S4** In the online supplement).

4.2. Matching Analysis

To improve on the covariate balance, we apply two matching techniques. First, we match on the propensity score as in Elder and Jepsen (2014). However, Iacus, King, and Porro (2012) found that coarsened exact matching (CEM) performs better than both ordinary least squares and propensity score matching in estimating causal effects. This technique matches each child in a religious school with a child in a public school with either the same observed characteristics or coarsened observed

[6] A fixed effects model may not be appropriate for binary outcome variables. Therefore, we repeat the analyses for diploma attainment and tertiary education enrolment using a conditional logit model. These analyses are presented in **Table S5** and lead to the same conclusions as the fixed effects models.

characteristics based on narrow categories (Blackwell, et al. 2009). Unlike propensity score matching, CEM matches characteristics on the entire distribution and not only the mean.

To perform CEM, we included all the categorical control variables and coarsened several continuous control variables. Household income was coarsened to deciles of the income distribution, mother's age at birth was coarsened to six categories (*younger than 25, 26 to 30, 31 to 35, 36 to 40, 41 to 45, older than 45*), and family size and birth order to three categories (*only child*, *two children*, *three or more children*; and *first-born*, *second-born*, *third or later born*, respectively). We further coarsened the number of family transitions to four categories: *no transitions*, *one transition, two transitions, three or more transitions*. Given that we include the entire population of children in the Netherlands, we searched for an exact match on all the covariates. The resuls are displayed in **Table S6** in the online supplement and are very similar to the main results. We follow the empirical rule of thumb by Vable and colleagues (2019), stating that OLS inferences are more unbiased than matching inferences if the estimates from OLS and matching are similar (i.e., confidence intervals overlap). Therefore, we do not present the matching results as part of our main analyses.

4.3. Cousin Fixed Effects

Initially, we attempted to use sibling fixed effects and compare children in religious schools with their brothers or sisters in public schools. However, most children attend the same school or at least the same school denomination as their siblings. Specifically, a sibling fixed effects model would calculate the coefficients on an effective sample of only about 300 children. Therefore, we exploited the grandparent identifiers in our data and we used cousin fixed effects instead. In essence, we compared school outcomes of children in religious schools with their cousins in public schools. On average, first cousins have about one-eight of their DNA in common (Weir, Anderson and Hepler 2006) and they my also share other family characteristics. Therefore, a cousin comparison may reduce the selection on unobserved characteristics (Geronimus, Korenman and Hillemeier 1994, Mazrekaj, De Witte and Cabus 2020). We present the results of cousin fixed effects in **Table S7** in the online supplement. It is reassuring to note that these analyses lead to very similar conclusions as the main analyses.

4.4. Bounding Analyses

To assess the causality of the potential religious school influence, we use Oster (2019) bounds. These bounds generalize Altonji and colleagues (2005b) bounds that have become the main method to analyze the causal influence of religious schools in the recent empirical literature (Elder and Jepsen 2014, Gibbons and Silva 2011, Gihleb and Giuntella 2017, Nghiem, et al. 2015). Both approaches rely on coefficient and R^2 movements before and after controlling for observed characteristics, to bound the treatment effect. Whereas Altonji and colleagues (2005b) assume that including the unobserved characteristics would produce an R^2 of 1, Oster (2019) allows the individual researchers to choose the maximum R^2. This is advantageous because a maximum R^2 of 1 may understate the robustness of the results, for instance, in the presence of measurement error in test scores. Indeed, Oster (2019) found that setting R_{max} of 1 rejects even the true associations as causal. Another advantage of the Oster (2019) bounds is that they allow us to calculate the selection-bias adjusted effect of religious schooling on school outcomes as follows:

$$(2) \qquad \beta^* \approx \tilde{\beta} - \delta[\dot{\beta} - \tilde{\beta}]\frac{R_{max}-\tilde{R}}{\tilde{R}-\dot{R}}$$

where β^* is the selection-bias adjusted effect of being enrolled in a religious school on school outcomes. Further, $\tilde{\beta}$ is the influence of religious schooling obtained from the regression of school outcomes on religious schooling and all the control variables (*controlled regression*), whereas $\dot{\beta}$ is the influence of religious schooling obtained from the regression of school outcomes on religious schooling without any control variables (*uncontrolled regression*). Correspondingly, $\tilde{R}$ is the R^2 (not adjusted R^2) from the controlled regression and $\dot{R}$ is the R^2 from the uncontrolled regression. The parameter R_{max} represents the maximum R^2 that can be achieved if we included all the unobserved characteristics in the model. We follow Oster's (2019) recommendation to set R_{max} at 1.3 times R^2 from the controlled regression as also used in Mazrekaj, De Witte, and Cabus (2020). Finally, δ is the selection ratio. The selection ratio represents the ratio of the effect of unobserved to observed characteristics. For instance, a selection ratio of 1 would mean that the effect of unobserved characteristics is as big as the effect of observed characteristics. We can manually specify this ratio to test the robustness of our results, i.e. we can calculate how big the effect of unobserved characteristics would have to be in relation to observed characteristics (control variables included in

the regression) to explain away the result. In the analyses, we will first calculate the selection-bias adjusted effect (β^*) when the selection on unobserved characteristics is just as important as the selection on observed characteristics, i.e. when the selection ratio is 1. Then, we will calculate the selection ratio that is necessary to explain away the results. For this purpose, we will bootstrap the standard errors using 1,000 replications. It should be noted that this bounding analysis assumes that selection on observed characteristics is informative about the selection on unobserved characteristics. We speculate that this is likely to be the case in our setting as we include some of the most commonly used determinants in the education production function such as gender, origin, household education, and household income.

5. Results

We start this section by comparing children in religious schools and children in public schools on the main control variables. Then, we estimate how children in religious schools perform on the standardized test at the end of primary education using the postcode fixed effects model. In addition, we present heterogeneity analyses by children's sex, origin, and family structure. We continue the section with an assessment of how children in religious schools perform beyong primary education, by analyzing diploma attainment and tertiary education enrolment for cohorts born between 1998 and 2000. We end the section with a bounding analysis to assess the extent of causality of our findings.

5.1. Sample Characteristics

Table 1 presents descriptive statistics by school denomination. In total, we observe 1,559,678 children in 7,108 schools. Contrary to prior work from the United States and the United Kingdom (Elder and Jepsen 2014, Gibbons and Silva 2011), children in Catholic schools appear to have a very similar socioeconomic status as children in public schools. Although children in Catholic schools are more likely to be of Dutch origin and to be in more intact families; their household income, mother's age at birth, household size, and household education appear to be very similar. This small rather than large difference as in the United States is likely a result of the absence of admission fees as well as religious and public schools all receiving public funding. We further observe that mainstream Protestant schools attract a very similar population of children as Catholic schools. The differences are much more pronounced when we observe children in Orthodox Protestant schools. These children are almost exclusively Dutch (98%) and live in large married families with both parents. Both their household education and income are lower than that of children in public schools. Children with the

highest socioeconomic status are children in Jewish schools. These children have the highest educated parents and enjoy the highest household income. In comparison to public schools, children in Jewish schools are also more likely to be in intact families, albeit these families are larger.

A unique feature of our data is that we also observe children with a significantly lower socioeconomic status than children in public schools, namely children in Islamic and Hindu schools. Children in Islamic schools are mostly of Moroccan or Turkish origin. They live in large, mostly married families with both parents present and their mothers gave birth at a relatively young age. In comparison to children in public schools, their parents are mostly high school dropouts with a very low household income. Specifically, children in Islamic schools have to do with only half of the income of children in public schools. A very similar situation is found for children in Hindu schools, who are mostly of Surinamese origin. Although their household income and education is higher than that of children in Islamic schools, it is still significantly lower than that of children in public schools. Unlike the children in Islamic schools, however, children in Hindu schools are much more likely to experience a family transition than children in public schools and to live with a single parent.

In **Table S7**, we dig deeper into the socioeconomic status of children in Islamic and Hindu schools. In particular, we asked whether these socioeconomic status differences also held if we controlled for children's origin. Thus, in **Table S7**, we reconstruct **Table 1** but solely for children of Moroccan, Turkish, and Surinamese origin. This allows us to compare for instance a Moroccan child in an Islamic school with a Moroccan child in a public or a Catholic school on observed characteristics.[7] Based on observed characteristics such as household education and income, we find that Moroccan, Turkish, and Surinamese children in Islamic and Hindu schools still have a lower socioeconomic status than Moroccan, Turkish, and Surinamese children in all other school denominations. Thus, Islamic and Hindu schools attract children with the lowest socioeconomic status in the Netherlands.

[7] Note that some schools do not include any Moroccan or Turkish children.

TABLE 1: DESCRIPTIVE STATISTICS

	All children	Public	Catholic	Mainstream Protestant	Orthodox Protestant	Jewish	Islamic	Hindu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gender (boy = 1)	.50	.50	.50	.50	.50	.52#	.47#	.49
Origin								
Dutch	.84	.80	.87#	.87#	.98#	.67#	.03#	.02#
Moroccan or Turkish	.07	.09	.06#	.05#	.00#	.01#	.79#	.01#
Other	.09	.10	.07#	.08#	.02#	.32#	.18#	.97#
Household education								
Less than high school	.15	.16	.14#	.15	.14#	.07#	.53#	.31#
High school diploma	.46	.43	.47#	.49#	.52#	.34#	.33#	.50#
Bachelor degree	.23	.22	.24#	.23	.23	.23	.10#	.14#
Master degree or Ph.D.	.16	.19	.15#	.14#	.11#	.35#	.05#	.05#
Household income	77,676	77,699	79,102	77,318	73,677#	95,080#	38,874#	54,410#
Mother's age at birth	30.95	31.09	31.16	30.74	29.52#	32.35#	28.11#	29.35#
Household structure								
Married parents	.61	.55	.61#	.64#	.83#	.70#	.66#	.38#
Cohabiting parents	.18	.19	.18	.16#	.11#	.10#	.16#	.17#
Parent and stepparent	.06	.07	.06	.06	.02#	.01#	.02#	.06
Single parent	.16	.18	.15#	.14#	.04#	.19	.15#	.38#
Family transitions	.46	.55	.44#	.42#	.14#	.49#	.51#	1.12#
Household size	4.45	4.38	4.32	4.48	6.04#	4.94#	5.50#	4.16
Birth order	1.77	1.71	1.70	1.79	2.74#	2.24#	2.32#	1.71
Number of schools	7,108	2,704	2,154	1,886	305	< 8*	50	< 8*
Number of children	1,559,678	555,572	548,025	384,280	59,738	361	9,737	1,965

Note:

The coefficient is significantly different from the baseline coefficient of children in public schools in column 2 at the 5 percent level using a two-tailed z test for continuous variables and a two-tailed test of proportions for categorical variables.

* For confidentiality reasons, Statistics Netherlands prohibits specifying the exact number of schools if the number is lower than eight.

5.2. Children in Religious Schools and Achievement in Primary Education

Table 2 presents the influence of religious schools on standardized test scores in primary education, interpreted in standard deviations. When we do not include any control variables other than the year of birth, column 1 coefficients largely represent the socioeconomic differences between children in different schools. Namely, children in Catholic, mainstream Protestant, and Jewish schools perform better on the standardized test than children in public schools, whereas the opposite is true for Islamic and Hindu schools. Interestingly, however, children in Orthodox Protestant schools perform .102 standard deviations higher on the test, although their household income is slightly lower than that in public schools. This may be a consequence of these children being more likely to live in intact families.

In column 2, we control for the observed characteristics and postcode fixed effects to account for geographical variation. **Table S4** in the online supplement includes a random-effects model to measure the size of municipality effects. The proportion of variance in the test scores accounted for by municipality is about 9 percent. Column 2 in **Table 2** shows that once we control for observed characteristics, children in all religious schools but Jewish schools outperform children in public schools on the standardized tests. Especially children in Islamic and Hindu schools appear to perform well, with large effect sizes of about 0.2 standard deviations. It is useful to estimate when exactly the coefficient changes from a negative influence to a positive influence of religious schooling. **Table S8** in the online supplement shows that solely including either household education or origin is sufficient to reverse the coefficient. Thus, the negative influence was mostly due to the low socioeconomic status of children enrolled in Islamic and Hindu schools. On the other hand, children in Jewish schools perform just as well on the test than children in public schools.

TABLE 2: THE INFLUENCE OF RELIGIOUS SCHOOLS ON CHILDREN'S STANDARDIZED TEST SCORES AT THE END OF PRIMARY EDUCATION (POSTCODE FIXED EFFECTS MODEL)

	(1)	(2)
School denomination (ref: public)		
Catholic	.052***	.028***
	(.011)	(.077)
Mainstream Protestant	.031**	.032***
	(.011)	(.006)
Orthodox Protestant	.102***	.147***
	(.021)	(.015)
Jewish	.182***	-.039
	(.063)	(.032)
Islamic	-.252***	.242***
	(.031)	(.026)
Hindu	-.056	.190***
	(.029)	(.053)
Gender (ref: female)		-.011***
		(.003)
Origin (ref: Dutch)		
Moroccan/Turkish		-.231***
		(.013)
Not Western		-071**
		(.026)
Household education (ref: less than high school)		
High school diploma		.277***
		(.005)
Bachelor degree		.598***
		(.006)
Master degree or Ph.D.		.872***
		(.009)
Log household income		.033***
		(.003)
Mother's age at birth		.017***
		(.000)
Household structure (ref: married parents)		
Cohabiting parents		-.086***
		(.003)
Parent and stepparent		-.159***
		(.006)
Single parent		-.109***
		(.010)
Family transitions		-.020***
		(.002)
Household size		.000
		(.000)
Birth order		-.066***
		(.002)
Fixed effects:		
Birth year	Yes	Yes
Postcode	No	Yes
Number of children	1,559,678	1,559,678
$\mathbf{R^2}$	.005	.140

Note: Standard errors clustered at the municipality (postcode) level are in parentheses.
* p < .05; ** p < .01; *** p < .001 (two-tailed t-tests).

5.3. Heterogenous Achievement by Sex and Family Structure

In **Table 3**, we perform the analyses by children's sex, origin, and family structure, for which heterogeneous associations have been found in prior work (Altonji, Elder and Taber 2005b, Elder and Jepsen 2014, Gibbons and Silva 2011). To operationalize differences by family structure, we have constructed an indicator for an intact family (living with both parents, irrespective of marital status) and not-intact family (not living with both parents, i.e. parent and stepparent or single parent). Columns 1 and 2 show the results for girls and boys separately. Overall, we find no heterogeneity by gender. Although coefficients are sometimes higher for boys than for girls, tests of equality of coefficients do not reject the null hypothesis of equal coefficients. We also find no heterogeneity by origin in columns 3 and 4.[8] On the other hand, we do find some differences by family structure in columns 5 and 6. Specifically, Catholic, mainstream Protestant, and Orthodox Protestant schools seem to be more beneficial for children in not-intact families (significant difference at the 5% level). By contrast, Islamic schools are equally beneficial for children in both intact and not intact families, whereas Hindu schools are particularly beneficial for children in intact families.

TABLE 3 – THE INFLUENCE OF RELIGIOUS SCHOOLS ON CHILDREN'S STANDARDIZED TEST SCORES AT THE END OF PRIMARY EDUCATION BY SEX, ORIGIN, AND FAMILY STRUCTURE

	Girls	Boys	Minority	Native	Not-intact family	Intact family
	(1)	(2)	(3)	(4)	(5)	(6)
School denom. (ref: public)						
Catholic	.030***	.025***	.030***	.028***	.030***	.023***
	(.003)	(.003)	(.005)	(.002)	(.002)	(.004)
Mainstream Protestant	.033***	.031***	.031***	.034***	.034***	.025***
	(.003)	(.003)	(.006)	(.002)	(.002)	(.005)
Orthodox Protestant	.142***	.151***	-.137***	.160***	.153***	.093***
	(.006)	(.006)	(.026)	(.005)	(.005)	(.017)
Jewish	-.167*	.075	.098	-.115	-.085	.098
	(.070)	(.069)	(.092)	(.059)	(.054)	(.114)
Islamic	.275***	.206***	.235***	.167**	.237***	.252***
	(.013)	(.014)	(.011)	(.053)	(.011)	(.024)
Hindu	.221***	.156***	.212***	.327*	.102***	.302***
	(.029)	(.031)	(.023)	(.139)	(.028)	(.033)
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of children	778,793	780,885	246,361	1,313,317	1,223,443	336,235
$\mathbf{R^2}$	.146	.137	.108	.123	.134	.111

Note: Standard errors clustered at the municipality (postcode) level are in parentheses.
* p < .05; ** p < .01; *** p < .001 (two-tailed t-tests).

[8] Note that a comparison of coefficients for children in Orthodox Protestant schools is not meaningful as these schools include only a few minority children.

5.4. Children in Religious Schools and Later School Outcomes

Table 4 compares children in religious schools versus public schools on diploma attainment in secondary education and enrolment into tertiary education. Children in our sample were born between 1998 and 2007, so we can only analyze children from earlier cohorts (1998 to 2000) as they alone are old enough to have graduated from secondary education. Overall, it appears that the influence of religious schooling is mainly concentrated in the short run. Once we control for children's background, columns 2 and 4 show that only Catholic and Protestant schools exert a significant positive influence on diploma attainment and tertiary education enrolment. However, these coefficients are very small, about 1% to 2% of a standard deviation. Thus, the influence of religious schooling seems to disappear in the longer run.

TABLE 4: THE INFLUENCE OF RELIGIOUS SCHOOLS ON CHILDREN'S DIPLOMA ATTAINMENT IN SECONDARY EDUCATION AND ENROLMENT INTO TERTIARY EDUCATION

	Diploma attainment		Tertiary education enrolment	
	(1)	(2)	(3)	(4)
School denom. (ref: public)				
Catholic	.018***	.009***	.028***	.010***
	(.003)	(.001)	(.003)	(.002)
Mainstream Protestant	.019***	.010***	.022***	.018***
	(.002)	(.002)	(.003)	(.002)
Orthodox Protestant	.047***	.014***	.023**	.014***
	(.004)	(.002)	(.008)	(.002)
Jewish	-.079**	-.059	.047	.019
	(.028)	(.031)	(.028)	(.037)
Islamic	-.059***	-.005	-.069***	.022
	(.007)	(.007)	(.012)	(.018)
Hindu	-.051**	-.001	-.018	.037
	(.018)	(.017)	(.014)	(.025)
Additional controls	No	Yes	No	Yes
Number of children	445,015	445,015	445,015	445,015
$\mathbf{R^2}$	.002	.021	.013	.055

Note: Standard errors (not shown for brevity) are clustered at the neighborhood level.
* p < .05; ** p < .01; *** p < .001 (two-tailed t-tests).

5.5. Bounding the Causal Religious School Effect

To better understand potential selection effects, we performed Oster's (2019) bounding approach. This approach exploits coefficient and R^2 movements before and after the inclusion of observed characteristics to bound the treatment effect. **Table 5** shows the effect of religious schools on different school outcomes when the selection on unobserved characteristics is just as high as the selection on observed characteristics. We find in column 1 that children in religious schools perform better than children in public schools on the standardized test in primary education even after the introduction of unobserved characteristics. We further calculated how large the selection on unobserved characteristics should be to neutralize the positive effect of religious schooling. Our results indicate that the selection on unobserved characteristics should be at least 2.5 times higher than the selection on observed characteristics to render the coefficients insignificant. Although it is possible that unobserved characteristics such as type of parenting and family values associated with religion could exercise a considerable influence on the results, we do include several of the main observable characteristics such as parental education, origin, age, and income, as well as a variable for a child's neighborhood. For diploma attainment in secondary education and enrolment into tertiary education in columns 2 and 3, coefficients turn insignificant once the selection on unobserved characteristics is as large as the selection on observed characteristics. Thus, it appears that the positive effect of religious schooling disappears by the end of secondary education.

Nevertheless, it is important to note that this bounding method assumes that selection on observed characteristics is informative about selection on unobserved characteristics. Moreover, the bounding analysis is dependent on the choice of R_{max} which is set at 1.3 times R^2 from the regression including all controls as recommended by Oster (2019). Although our results are robust to using a higher value of R_{max} set to twice the R^2 from the controlled regression (see **Table S9** in the online supplement), in the absence of studies that used this type of analysis in a similar context, it is unclear how large R_{max} should be.

TABLE 5: BOUNDING THE CAUSAL EFFECT OF RELIGIOUS SCHOOLS ON CHILDREN'S ACHIEVEMENT, DIPLOMA ATTAINMENT, AND ENROLMENT INTO TERTIARY EDUCATION

	Standardized test score in primary education	Diploma attainment in secondary education	Tertiary education enrolment
	(1)	(2)	(3)
School denom. (ref: public)			
Catholic	.024*	.009	.004
	(.011)	(.008)	(.004)
Mainstream Protestant	.047***	.012	.024
	(.010)	(.009)	(.021)
Orthodox Protestant	.177***	.033	.034
	(.048)	(.028)	(.029)
Jewish	-.101	-.049	.015
	(.098)	(.041)	(.013)
Islamic	.409***	.016	.058
	(.108)	(.010)	(.049)
Hindu	.275***	.019	.060
	(.088)	(.011)	(.054)
Number of children	1,559,678	445,015	445,015

Note: Bootstrapped standard errors are in parentheses (1,000 replications). R^2max is set at 1.3 times R^2 from the regression including all controls as recommended by Oster (2019). The selection ratio is set at 1, indicating that the selection on unobserved characteristics is as high as the selection on observed characteristics.
* $p < .05$; *** $p < .001$ (two-tailed t-tests).

6. Discussion

Using an administrative population dataset from the Netherlands and a range of rigorous methods, this article compared school outcomes of children in religious versus public schools. The results indicated that children in relgious schools outperformed children in public schools on standardized tests in primary education, with no heterogeneity by gender. We found no evidence of selection bias as the influence of religious schooling ran contrary to the expectations based on socioeconomic status. In the longer run, the influence of religious schooling faded out. Once we accounted for selection effects, children in religious schools were just as likely as children in public schools to graduate from high school and to enrol into tertiary education.

Overall, these results contrast with most recent studies that claimed that the seeming religious school advantage occurs through sorting of children into religious schools according to pre-existing characteristics and preferences (Altonji, Elder and Taber 2005b, Elder and Jepsen 2014, Gibbons and Silva 2011, Gihleb and Giuntella 2017, Nghiem, et al. 2015). It also contrasts with earlier studies that claimed that religious schools are unequivocally beneficial for children. Instead, we argue that children do seem to benefit from religious schooling in primary education but that these differences disappear over time. We attribute these differences in findings to the unique Dutch context that enabled us to study religious schools that are all publicly funded, that do not charge any fees, and that include a range of religious denominations. As a result, our results were less plagued by selection bias than in the previous literature.

It is possible to address some of the theoretical mechanisms that may be driving our results. Our heterogeneity analyses by family structure partially support the social capital theory. We found some evidence that children who lived with a parent and a stepparent or a single parent, rather than both parents, especially benefit from religious schools. As children from not-intact families are likely to have a lower social capital at home than children living both parents, the social capital resulting from religious schooling may compensate for the lack of social capital at home.

Religious schools may also produce better test scores if they spend more resources on children and have higher-quality teachers. Although our administrative data does not include information on teachers, we do observe class sizes. We find that Catholic, mainstream Protestant, and Orthodox Protestant schools have about a one child larger average class size (24 children) than public schools (23 children). On the other hand, Jewish schools have about 22 children in the class, Islamic schools 21 children and Hindu schools 23 children. Given the very small differences in class size that are not consistent between religious schools, it is unlikely that class size explains our results.

As a final potential explanation of our results, we test the discrimination theory based on the Islamic schools analyses. Specifically, the Islamic terrorist attacks may have changed the association between Islamic school enrolment and children's test scores. If discrimination theory drives religious school advantage, we would expect a more positive association after terrorist attacks than before as Islamic children may be more discriminated against in public schools after terrorist attacks. In **Table S10** in the online supplement, we estimate the influence of Islamic schooling on standardized test scores before and after the three coordinated suicide bombings by the Islamic State (IS) in Belgium in 2016. We do indeed find a slightly larger coefficient after the terrorist attacks than before, suggesting that religious schools may exert a protective effect on minority children in primary education. Thus, the social capital theory and the discrimination theory provide partial explanations for our findings.

Although we attempted to offer a comprehensive analysis of the religious schooling influence using population data over a substantial period and a range of models, this study is not without limitations. First, our administrative data do not include a measure of the family's religious affiliation. Thus, we may be confounding the effect of religious schooling with the effect of religious upbringing (Pearce, Uecker and Lundquist Denton 2019). Although our bounding analyses may have partially addressed this issue, it is likely that some of the selection issues remain. We also solely studies school outcomes and made no assessment of any behavioural outcomes nor long-run outcomes such as earnings or criminal outcomes. Similarly, our sample of children in Jewish schools may be too small to draw any strong conclusions and we could only study diploma attainment and tertiary education enrolment for the earlier cohorts. Finally, further research should continue refining the insights from these administrative findings with qualitative analyses that better assess the religious schooling mechanisms.

References

Agirdag, Orhan, Geert Driessen, and Michael S. Merry. 2017. "The Catholic school advantage and common school effect examined: a comparison between Muslim immigrant and native pupils in Flanders." *School Effectiveness and School Improvement* 28 (1): 123-135.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005a. "An evaluation of instrumental variable strategies for estimating the effects of Catholic schools." *Journal of Human Resources* 40 (4): 791-821.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005b. "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy* 113 (1): 151-184.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* New Jersey: Princeton University Press.

Bibler, Andrew, and Stephen B. Billings. 2020. "Win or Lose: Residential Sorting After a School Choice Lottery." *Review of Economics and Statistics* 102 (3): 457-472.

Billings, Stephen B., David J. Deming, and Jonah Rockoff. 2014. "School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg." *Quarterly Journal of Economics* 129 (1): 435-476.

Blackwell, Matthew, Stefano Iacus, Gary King, and Giuseppe Porro. 2009. "CEM: Coarsened Exact Matching in Stata." *Stata Journal* 9 (4): 524-546.

Bressler, Marvin, and Charles Westoff. 1963. "Catholic Education, Economic Values, and Achievement." *American Journal of Sociology* 69 (3): 225-233.

Brutsaert, Herman. 1998. "Home and School Influences on Academic Performance: State and Catholic elementary schools in Belgium compared." *Educational Review* 50 (1): 37-43.

Bulman, May. 2017. *Muslim children as young as nine branded terrorists in wake of recent terror attacks, reveals Childline.* 28 June. Accessed November 20, 2020. https://www.independent.co.uk/news/uk/home-news/nine-year-old-children-terrorists-race-faith-hate-bullying-childline-manchester-arena-terror-attack-westminster-london-a7810051.html.

Cameron, Colin A., and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50 (2): 317-372.

Candipan, Jennifer. 2020. "Choosing Schools in Changing Places: Examining School Enrollment in Gentrifying Neighborhoods." *Sociology of Education* 93 (3): 215-237.

Carbonaro, William, and Elizabeth Covay. 2010. “School Sector and Student Achievement in the Era of Standards Based Reforms.” *Sociology of Education* 83 (2): 160-182.

Central Institute of Test Development. 2010. *Reflection and Results 2010 [Terugblik en resultaten 2010].* Arnhem, the Netherlands: Central Institute of Test Development.

Coleman, James S. 1988. “Social Capital in the Creation of Human Capital.” *American Journal of Sociology* 94 (Supplement): S95-S120.

Coleman, James S., and Thomas Hoffer. 1987. *Public and Private High Schools: Impact of Communities.* New York City: Basic Books.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1982. *High school achievement: Public, Catholic, and private schools compared.* New York City: Basic Books.

Coleman, James S., Thomas Hoffer, and Sally Kilgore. 1981. *Public and Private Schools.* Chicago: National Center for Education Statistics.

Darnell, Alfred., and Darren, E. Sherkat. 1997. “The Impact of Protestant Fundamentalism on Educational Attainment.” *American Sociological Review* 62 (2): 306-315.

De Witte, Kristof, and Chris Van Klaveren. 2014. “The influence of closing poor performing primary schools on the educational attainment of students.” *Educational Research and Evaluation* 20 (4): 290-307.

Deen, Shulem. 2018. “Why Is New York Condoning Illiteracy?” The New York Times, 4 April.

Driessen, Geert, Orhan Agirdag, and Michael S. Merry. 2016. “The gross and net effects of primary school denomination on pupil performance.” *Educational Review* 68 (4): 466-480.

Elder, Todd, and Christopher Jepsen. 2014. “Are Catholic primary schools more effective than public primary schools?” *Journal of Urban Economics* 80: 28-38.

Elwert, Felix, and Christopher Winship. 2014. “Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable.” *Annual Review of Sociology* 40: 31-53.

Epple, Dennis, Richard E. Romano, and Miguel Urquiola. 2017. “School Vouchers: A Survey of the Economics Literature.” *Journal of Economic Literature* 55 (2): 441-492.

Espinoza et al. v. Montana Department of Revenue et al. 2020. 18-1195 (Supreme Court of the United States, 30 June).

Evans, William N., and Robert M. Schwab. 1995. “Finishing high school and starting college: Do Catholic schools make a difference?” *Quarterly Journal of Economics* 110 (4): 941-974.

Figlio, David, and Joe A. Stone. 1999. “School choice and student performance: Are private schools really better?” *Research on Labour Economics* 18: 115-140.

Geronimus, Arline T., Sanders Korenman, and Marianne M. Hillemeier. 1994. “Does Young Maternal Age Adversely Affect Child Development? Evidence from Cousin Comparisons in the United States.” *Population and Development Review* 20 (3): 585-609.

Gibbons, Stephen, and Olmo Silva. 2011. “Faith Primary Schools: Better Schools or Better Pupils?” *Journal of Labor Economics* 29 (3): 589-635.

Gihleb, Rania, and Osea Giuntella. 2017. “Nuns and the effects of Catholic schools. Evidence from Vatican II.” *Journal of Economic Behavior & Organization* 137: 191-213.

Goldhaber, Dan, Vanessa Quince, and Roddy Theobald. 2017. “Has It Always Been This Way? Tracing the Evolution of Teacher Quality Gaps in U.S. Public Schools.” *American Educational Research Journal* 55 (1): 171-201.

Grogger, Jeffrey, and Derek Neal. 2000. “Further evidence on the effects of Catholic secondary schooling.” *Brookings-Wharton Papers on Urban Affairs* 151-201.

Iacus, Stefano M., Garry King, and Giuseppe Porro. 2012. “Causal Inference without Balance Checking: Coarsened Exact Matching.” *Political Analysis* 20 (1): 1-24.

Jackson, Kirabo C., Rucker C. Johnson, and Claudia Persico. 2016. “The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms.” *Quarterly Journal of Economics* 131 (1): 157-218.

Jepsen, Christopher. 2003. “The Effectiveness of Catholic Primary Schooling.” *Journal of Human Resources* 38 (4): 928-941.

Jeynes, William H. 2012. “A Meta-Analysis on the Effects and Contributions of Public, Public Charter, and Religious Schools on Student Outcomes.” *Peabody Journal of Education* 87 (3): 305-335.

Lenski, Gerhard Emmanuel. 1961. *The Religious Factor: A Sociological Study of Relgion's Impact on Politics, Economics and Family Life.* New York City: Doubleday and Company.

Mazrekaj, Deni, Kristof De Witte, and Sofie Cabus. 2020. “School Outcomes of Children Raised by Same-Sex Parents: Evidence from Administrative Panel Data.” *American Sociological Review* 85 (5): 1-27.

McCrory Calarco, Jessica. 2020. “Avoiding Us versus Them: How Schools' Dependence on Privileged "Helicopter" Parents Influences Enforcement of Rules.” *American Sociological Review* 85 (2): 223-246.

Morgan, Stephen L., and Aage B. Sørensen. 1999. “Parental Networks, Social Closure, and Mathematics Learning: A Test of Coleman's Social Capital Explanation of School Effects.” *American Sociological Review* 64 (5): 661-681.

National Center for Education Statistics. 2019. *School Choice in the United States: 2019.* Accessed November 20, 2020. https://nces.ed.gov/programs/schoolchoice/ind_03.asp#f1.

Neal, Derek. 1997. “The effects of Catholic secondary schooling on educational achievement.” *Journal of Labor Economics* 15 (1): 98-123.

Nederlandse Omroep Stichting. 2016. *Rutte: no clear indications attack in the Netherlands [Rutte: geen concrete aanwijzingen aanslag in Nederland].* 22 March. Accessed November 20, 2020. https://nos.nl/artikel/2094403-rutte-geen-concrete-aanwijzingen-aanslag-in-nederland.html.

Nghiem, Hong Son, Ha Trong Nguyen, Rasheda Khanam, and Luke B. Connelly. 2015. “Does school type affect cognitive and non-cognitive development in children? Evidence from Australian primary schools.” *Labour Economics* 33: 55-65.

Noell, Jay. 1982. “Public and private schools: A reanalysis of "Public and private schools".” *Sociology of Education* 55 (2): 123-132.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 37 (2): 1-18.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 37 (2): 1-18.

Pearce, Lisa D., Jeremy E. Uecker, and Melinda Lundquist Denton. 2019. “Religion and Adolescent Outcomes: How and Under What Conditions Religion Matters.” *Annual Review of Sociology* 45: 201-222.

Pearman, Francis A., and Walker A. Swain. 2017. “School Choice, Gentrification, and the Variable Significance of Racial Stratification in Urban Neighborhoods.” *Sociology of Education* 90 (3): 213-235.

Portes, Alejandro. 1998. “Social Capital: Its Origins and Applications in Modern Sociology.” *Annual Review of Sociology* 24: 1-24.

Richardson, Hannah. 2014. “'Radicalisation risk' at six Muslim private schools, says Ofsted.” 21 November. Accessed January 31, 2021. https://www.bbc.co.uk/news/education-30129645.

Rosenbaum, Paul R. 1984. “The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment.” *Journal of the Royal Statistical Society (Series A)* 147 (5): 656-666.

RTL Nieuws. 2019. “Inspectie onderzoekt islamitische scholen vanwege homofoob lesmateriaal.” 12 September. Accessed January 31, 2021. https://www.rtlnieuws.nl/nieuws/politiek/artikel/4846411/inspectie-islamitische-scholen-homofobie-salafisme-homoseksualiteit.

Sander, William. 1996. “Catholic grade schools and academic achievement.” *Journal of Human Resources* 31 (3): 540-548.

Sander, William. 2010. “Religious background and educational attainment: The effects of Buddhism, Islam, and Judaism.” *Economics of Education Review* 29: 489-493.

Scheerens, Jaap, Melanie C. M. Ehren, Peter J. C. Sleegers, and Renske de Leeuw. 2012. *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands.* The Hague, The Netherlands: Dutch Ministry of Education.

Sullivan, Alice, Samantha Parsons, Francis Green, Richard D. Wiggins, George Ploubidis, and Timmy Huynh. 2018. “Educational attainment in the short and long term: was there an advantage to attending faith, private, and selective schools for pupils in the 1980s?” *Oxford Review of Education* 44 (6): 806-822.

Taylor, Kate. 2015. “New York City Questions English, Math and Science Taught at Yeshivas.” The New York Times, 31 July.

Times NL. 2016. *Four years in prison for terrorist attack on Enschede mosque.* 27 October. Accessed November 20, 2020. https://nltimes.nl/2016/10/27/four-years-prison-terrorist-attack-enschede-mosque.

Timmermans, Anneke C., Hester de Boer, Hilda T. A. Amsing, and Greetje van der Werf. 2018. “Track recommendation bias: Gender, migration background and SES bias over a 20-year period in the Dutch context.” *British Educational Research Journal* 44 (5): 847-874.

Vable, Anusha M., Mathew V. Kiang, Maria M. Glymour, Joseph Rigdon, Emmanuel F. Drabo, and Sanjay Basu. 2019. “Performance of Matching Methods as Compared With Unmatched Ordinary Least Squares Regression Under Constant Effects.” *American Journal of Epidemiology* 188 (7): 1345-1354.

Weir, Bruce S., Amy D. Anderson, and Amanda B. Hepler. 2006. “Genetic relatedness analysis: modern data and new challenges.” *Nature Reviews Genetics* 7: 771-780.

Witte, Griff, Souad Mekhennet, and Michael Birnbaum. 2016. *Islamic State claims responsibility for the Brussels attacks.* 22 March. Accessed November 20, 2020.

https://www.washingtonpost.com/world/brussels-on-high-alert-after-explosions-at-airport-and-metro-station/2016/03/22/b5e9f232-f018-11e5-a61f-e9c95c06edca_story.html.

Online Supplement

Do Children Perform Better in Religious Schools? Evidence From Population Data

By DENI MAZREKAJ AND CHRISTIAAN MONDEN

Tables in the Online Supplement are currently awaiting approval for publication by Statistics Netherlands. Nonetheless, the conclusions of the tables are summarized in the main text.