Lancaster University
Management School

Economics Working Paper Series

2020/016

The Roles of Faith and Faith Schooling in Educational, Economic, and Faith Outcomes

Andrew McKendrick and Ian Walker

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

LUMS home page: http://www.lancaster.ac.uk/lums/

The Roles of Faith and Faith Schooling in Educational, Economic, and Faith Outcomes

Andrew McKendrick[a] and Ian Walker[b]

August 3, 2020

Keywords: Religiosity, faith schools, educational and long-term outcomes, faith persistence
JEL codes: I21, Z12

Abstract:

We examine the effects of intrinsic religiosity and faith-based schooling on short and longer-term outcomes among young people in England. Without an obvious quasi-experimental identification strategy we rely on a detailed dataset, a cohort study from England with an extensive range of household and school-level characteristics, to use Ordinary Least Squares (OLS), augmented by the Oster (2019) test. Inverse Probability Weighting and mediation analysis are also employed. We show that an individual's intrinsic religiosity is an important driver of short-term educational outcomes (age 16 test scores) and some longer-term outcomes (later Christian belief), while faith-based schooling plays a lesser role.

Acknowledgements: McKendrick was supported by a PhD studentship from the Economic and Social Research Council, while Walker was supported by a grants from the ESRC and the Nuffield Foundation that were concerned with wider school choice issues. We are grateful to participants in seminars at Lancaster University, and to Derek Neal, Olmo Silva, Jack Britton, Damon Clark, Emma Gorman, Giuseppe Migali, Maria Navarro Paniagua, Vincent O'Sullivan, and Anwen Zhang who gave helpful comments at various events**.** We are grateful to the Centre for Longitudinal Studies (CLS) at UCL's Institute of Education, and to the UK Data Service for making available the Next Steps and NPD data. However, neither CLS nor the UK Data Service bear any responsibility for the analysis or interpretation of these data. Access to the data is controlled by the UKDS after registration, application and training. Contact us (a.mckendrick@lancaster.ac.uk) if you require help and guidance with this process, or to access our code.

a. Lancaster University, contactable at a.mckendrick@lancaster.ac.uk
b. Lancaster University and IZA, contactable at ian.walker@lancaster.ac.uk

1 Introduction

There is a widely held view, among policy makers and parents alike, that faith-based schooling generates superior academic outcomes relative to the alternatives. In many countries faith schools are part of the private or charitable sectors and charge fees; in contrast, in England, faith schools are almost entirely state-funded and cannot charge fees. Faith schools must follow non-faith state schools in delivering the National Curriculum; their funding arrangements are closely comparable, with money following pupils; the requirements for teachers are the same (though faith schools are allowed to use faith as a criterion in hiring decisions); and both are regulated by the Office for Standards in Education (OFSTED). There is much greater comparability between faith and non-faith schools in England than there is elsewhere[1].

Of the state secondary schools in the England, 18.7% are faith schools a substantial portion of the education system for students aged 11 to 16 (Andrews and Johnes, 2016). Faith schools are popular, generally oversubscribed, and not just for those with religious belief; although faith schools can discriminate by faith in admissions to some extent. Around 20% of those pupils who attend faith schools have no religion or say their religion is of no importance to them. In contrast, approximately 60% of pupils in secular state schools say the same.

A simple comparison of the academic attainment of those in faith schools, compared to their secular equivalents, generally supports the view that they secure better pupil outcomes. Progression to an academic track post-age 16 in England is driven by performance in national exit examinations known as General Certificate of Secondary Education (GCSE). The usual benchmark by which both schools and students were judged was the attainment of at least five "good" grades at GCSE; with "good" being defined as achieving a grade between A* and C, where C is traditionally thought of as a pass.[2] In England, recent evidence (Andrews and Johnes, 2016) has shown faith schools average over 60 percent of pupils meeting that benchmark, compared to 57.4 percent in other state schools.

The impact of faith schools has been addressed extensively in the economics of education literature. What has *not* been examined (to the best of our knowledge) is the extent to which an individual's intensity of belief (a measure of intrinsic religiosity) could account for the impact

[1] See Long, Danechi, and Loft (2019) for an up-to-date summary of English faith school distinctiveness and how this might become more distinctive in the future. In the US, faith schools are increasingly covered by school choice programs that include vouchers, tax credit scholarships, and education savings accounts that can be used to offset private school fees.

[2] This A*-C grade benchmark applies to the cohort for whom we have data. Since 2017 a new grading system that uses numbers instead of letters has been in use.

that faith schools have, if indeed they have one. When it comes to the identification of causal effects it is obvious that issues will arise because pupils do not randomly select into the type of school they attend. Bias resulting from selection on unobserved characteristics[3] is easy to imagine in the context of both religious belief and faith school attendance. In the English case, no quasi-experimental method suggests itself, which is not uncommon in the literature. Altonji, Edler and Taber (2005b) argue that no satisfactory instrument exists for estimating the impact of faith schools and instead they develop a method to establish the sensitivity of estimates to selection on unobservables. This method (also referred to as AET hereafter) was later expanded upon and formalised by Oster (2019) and it is this version of the test that we apply here. The approach is well-suited to the particularly rich dataset that we have. Alongside this we use Inverse Probability Weighting Regression Adjustment (IPWRA), which models treatment and outcome separately (see Imbens and Wooldridge (2009)). In addition, we attempt to identify the pathway by which faithfulness and faith schooling are impacting outcomes using non-cognitive skills as mediators along the lines of Acharya, Blackwell, and Sen (2016).

The contribution of this paper is to explore the extent to which pupil outcomes are driven by the type of school they attend or by their own religiosity, defined as the intensity of their belief (rather than their affiliation or participation). We also contribute by examining longer-term outcomes including university attendance and religious belief at age 25. Finally, we continue the tradition in this literature of applying methods that are **not** robust to selection on unobservables but, instead, rely on testing for the potential bias from such selection. We provide lower bounds to effect sizes, conditional on particular assumptions, using the test developed from AET by Oster (2019).

This paper uses a powerful English dataset to attempt to answer these questions. The dataset combines administrative data from the National Pupil Database (NPD) for all children in all schools in England, with detailed survey data from Next Steps – a cohort study that began in 2004 comprising 15,770 individuals in its first wave. The survey randomly drew 21000 individuals age 13 or 14 (year group 9 of 1 at age 4/5 to11 at age 15/16 in the English system) from around 650 randomly selected English secondary schools and interviewed them each year, over seven waves, until they were age 20. A further, eighth, wave was conducted in 2015 when the respondents were age 25 and a further wave is planned at age 31. Next Steps is similar in

[3] Walker and Weldon (2020) conduct an extensive examination of the operation of the school choice mechanism. The mechanism is blind to prior ability (national test scores are conducted at age 11) with the exception of a very small number or remaining "grammar" schools that do select by prior ability. Although Faith schools are allowed to ask for further information from parents, that might facilitate cherry picking, the authors find no evidence that they do.

character to the well-known US National Longitudinal Survey of Youth (NSLY) – although it only contains a single cohort.

The findings are clear. Being faithful (having higher intrinsic religiosity) at age 14, compared to being not being faithful at the same age, is associated with higher attainment at GCSE, and with a greater likelihood of having a religious belief at age 25 – results that we show to be robust to unobserved confounders. Individual attainment in the national Advanced Level (A-level) examinations at the end of high school, around age 18, and university attendance are also significantly affected by religious belief, though these are less robust. Other outcomes such as: attending one of the more prestigious, so-called Russell Group, universities; university degree class; and the wage rate at age 25 do not appear to be significantly affected by faith. In contrast, the impact of faith schools seems to be much more equivocal. There is a suggestion that faith schools are effective at helping their pupils attain the five GCSE benchmark, but no other outcomes appear to be significantly impacted - except later religious belief. This raises questions as to why parents choose faith schools, which are examined in section 5.6.

The effectiveness of private, faith-based, education has become a more important issue in several countries. Our context is well suited to considering the impact of faith schooling when its provided free. In the US, a recent Supreme Court judgment suggests that tax dollars can now subsidize private schooling (which is often faith-based). In England, the government has committed to expanding existing faith schools and creating new ones, although a large minority of all publicly funded schools are already faith schools. Funding for such schools in England is equivalent to a US voucher that covers the entire costs.

The paper proceeds as follows. The relevant literature is outlined in section 2; the institutional setting and data description in section 3; the empirical strategy in section 4; and the results in section 5. Section 6 concludes.

2 Literature

This paper contributes to two distinct areas of literature – the faith schooling literature in educational outcomes, and the smaller literature on the impacts of religious belief. The faith school literature is extensive but is focussed on Catholic schools in the United States[4]. Much less is known beyond this. A number of papers have brought a range of (arguable) instruments to bear on the identification question. Hoxby (1994) uses an area's Catholic population as an

[4] English faith schools can be thought of as equivalent to US private schools with a voucher covering the full fees.

instrument for the presence of a Catholic school and finds a positive impact on area-wide achievement. Noting that exam-based attainment may mean little to some students, Evans and Schwab (1995) examine the impact of Catholic school attendance (at high school level) on both the probability of completing high-school and the probability of going to college. They use Catholic religious affiliation as an instrument for attending a catholic school. Though they go to some lengths to outline the validity of this approach, it is one that the literature as a whole no longer considers to be valid (see Altonji, Elder, and Taber (2005a) discussed below).

Correspondingly, Neal (1997) instruments Catholic school attendance with Catholic population density and the density of Catholic schools in a particular area to find a positive impact on wages of attendance at a Catholic school for urban minorities, a small effect for urban whites, and no discernible effect for suburban whites. Kim (2011) conducts very similar work and finds similar effects. The interesting element of the latter paper is that the data contains measures of school quality and that these explain large parts of the Catholic school effects (teacher quality being particularly important). These approaches have been criticised for their potential lack of validity. Neal's paper specifically is critiqued by Cohen-Zada and Sander (2008) who when replicating his work using different data find Catholic school effects are attenuated substantially by the inclusion of controls for religious affiliation.

Perhaps more convincingly, West and Woessmann (2010) and Allen and Vignoles (2016) each employ the historical religious population of an area as an instrument for the presence of a faith school in that area. The former finds a positive effect and the latter finds little evidence of an impact. However, if culture and values are persistent then the historical population in an area may still affect outcomes through wider cultural mechanisms rather than through religiosity. This could make the instrument invalid. Controversially, Carattini et al. (2012) use Catholic sex abuse scandals in the US as an instrument for the likelihood that an individual is enrolled in a Catholic school. The effect of Catholic school enrolment on public school test scores is then examined to judge if competition from Catholic schools implies better test scores; it does, suggesting that those schools themselves are better performing.

More broadly, Altonji, Elder, and Taber (2005a) argue that there is no convincing exogenous variation that would facilitate analysis of the impacts of faith schools. Though validity of instruments cannot be tested per se, the authors explore a number of routes to cast doubt on the instrumental variable strategies used in the literature. Instead they use data on public school eighth graders, few of whom attend Catholic school, to find a strong link between Catholic religion and educational attainment. In addition, they employ their own innovative method used

in their earlier work, that was ultimately published as Altonji, Elder, and Taber (2005b). The method uses the degree of selection on observables to infer the potential impact of unobserved selection (this was later formalised by Oster (2019)). Through it they find a positive impact of Catholic schools on the likelihood of attending university. A range of papers have employed this method, for example Cardak and Vecci (2013) also find small benefits of Catholic schools in terms of the likelihood of attending university.

A further contribution comes from Gihleb and Giuntella (2017) who exploit the rapid decline in the supply of teaching nuns lead to widespread closures of Catholic private schools following Vatican II (1962-1965) - a process of spiritual introspection and renewal for the Roman Catholic church. This is more convincing that early IV attempts because it exploits spatial and temporal variation and they find no effect on grade repetition, contrary to their OLS results. Interestingly, they use the AET method to examine the robustness of the OLS results and find that even a small degree of selection on unobservables is sufficient to drive the OLS results to zero, confirming their IV results.

Chingos and Peterson (2015) find that winning the New York City scholarship lottery had *no* overall impact on the college enrolment of students as a whole, but it did have positive impacts for students of color. Much of the effect was associated with Community College enrolment. The effects were larger for Catholic schools that other private schools.

Importantly for the English context, Gibbons and Silva (2011) argue that no credible instrument exists for attendance at a particular type of school. Given this, they use a number of techniques to analyse the impacts of faith schools. They combine their detailed dataset with prior subject-by-achievement-level fixed effects and home-postcode fixed effects; they then exploit the fact that selection occurs twice in choosing faith schools – at both primary and secondary level – to use secondary-type-by-postcode fixed effects to account for family and individual characteristics assuming selection at both secondary and primary level are comparable. If selection into secondary school type is driven by performance in primary then this method will be flawed – so the authors compare those who stay in a faith school between primary and secondary to those who stay in a non-faith school (the faith and non-faith stayers), and compare those who switch between the two types (the switchers). Assuming positive selection into faith schools the stayers will provide an upper bound of the faith school effect and the switchers a lower bound. Switchers are found to have virtually no effect and stayers a small positive effect. Finally, the authors implement the Altonji et al method to find that for stayers the moderate effect vanishes even if it is assumed that there are only modest degrees of unobserved selection.

Non-cognitive outcomes have also been examined. In the US, Elder and Jepsen (2014) find, using OLS, propensity score matching, and the Altonji et al method, little evidence of an impact on non-cognitive skills (along with a negative effect of Catholic schools on achievement in mathematics tests). Following the same approach Nghiem et al. (2015) find no effect of faith schooling on cognitive or non-cognitive skills in Australia. Their range of non-cognitive skills and controls is extensive which adds weight to their research even in the absence of a more conventional identification strategy. A number of papers also observe a positive relationship between Catholic school attendance in the US and subsequent religiosity (e.g. Sander (2001), and Wadsworth and Walker (2017)) and the notion that parents may explicitly send their children to faith schools in order to improve the likelihood that their child remains religious.

A recent report by Andrews and Johnes (2016) makes clear that the backgrounds of those attending faith schools do differ to those of students attending non-faith schools. Faith schools take fewer students from disadvantaged backgrounds (as measured by the proportion of pupils in receipt of free school meals), fewer students who have Special Educational Needs (SEN), and students who are already academically more able. Besides this, other unobserved characteristics exist that make pupils at faith schools different from those at non-faith schools.

Turning to religiosity, Hungerman (2014a) discusses religion in the context of club goods. Individuals can have the option of religious consumption and secular consumption. The presence of potential free-riders who want salvation without necessarily conforming to certain practises and rules leads religious groups to emphasise certain behaviours to screen out the unfaithful. These behaviours may include hard work which has implications for educational attainment and labour outcomes. McCullough and Willoughby (2009) similarly suggest that religion modifies an individual's priorities so that they want to accord with the prescribed practices. The promotion of honest toil and good behaviours would fit the educational context.

Endogeneity pervades the empirical analysis of the economic impacts of religion. Self-selection means that a particular kind of person could choose to be religious but would, in the absence of their belief, still perform better in the education system. Reverse causality too has been evidenced in compulsory schooling research in Canada and Turkey (Hungerman (2014b) and Cesur and Mocan (2018) respectively). Finally, the effect of education on religion may not be present for all faiths equally (McFarland, Wright, and Weakliem, 2011) – hence we decompose results between Protestants and Catholics below.

Some work claims to identify exogenous variation. Gruber (2005) innovated the religious market density instrument that employs the share of people of the same religious background in a particular area as an instrument for an individual's religiosity. It is not difficult to imagine spillovers that would make this instrument invalid.

Along more historical lines, Becker and Woessmann (2009) investigate whether a Protestant work ethic resulted in greater levels of economic prosperity in the 1500s. Using distance to Wittenberg (the epicentre of Lutheran Protestantism) as an instrument for Protestant belief there is found to be a positive and significant impact on literacy. In order to read the Bible, one has to be able to read which leads to other economic developments. Similarly, Spenkuch (2017) uses a 1555 treaty to engineer a fuzzy regression discontinuity design (RDD). Serfs followed the faith of their territorial lord (either Catholicism or Protestantism) creating a patchwork of religious populations that correlates strongly with the situation today. Protestants are found to be more likely to work longer hours, and though they do not earn higher wages, they earn more as a result of being paid for more hours of work. Evidently these instruments, though convincing, are not available for use in the setting of English schools from 2004 onwards.

Evidence is not limited to the Protestant case. Oosterbeek and van der Klaauw (2013) and Campante and Yanagizawa-Drott (2015) each use the timing of Ramadan to find a negative impact of religious practise on individuals' test scores in the case of the former and a nation's economic growth in the case of the latter. However, happiness is found to improve in the latter in accordance with the Hungerman (2014a) club good definition. The implications of this are unclear. The Becker and Woessman result does not have its origin in belief, but in an almost incidental need for literacy. The Spenkuch result points more clearly to religion, whilst the Ramadan-based work ultimately suggests an effect resulting from (a temporary decline in) nutrition, albeit among those with strong enough religiosity to adhere to the practice.

Other work also suggests a role for belief. Focusing on work ethic in an ordered probit analysis, Schaltegger and Torgler (2010) find Protestant faith is still statistically significant when interacted separately with both education and with intensity of religious belief within the same specification. Lehrer (2004), in the context of a model of supply and demand for education finance, finds conservative Protestant women who attend church regularly complete almost one additional year of schooling compared to the less observant.

At the intersection of these two literatures are those papers that examine religiosity and faith schooling effects together. As Cohen-Zada and Sander (2008) point out, there are studies that

control for religious affiliation in the estimation of Catholic school effects, though many do not control for religious groups other than Catholics. A UK-focused paper that does explicitly control for religiosity as part of its research quest ion, and is thus close to ours conceptually, is Sullivan et al. (2018), who examine the long-term impact of faith schooling whilst controlling explicitly for the individual's faith of upbringing.

Using the British Cohort Study (BCS) they are able to look at the long term achievement effects of faith schools and (affiliation-based) religiosity. They find effects of faith schooling and of religiosity on a range of outcomes. As this is close to what we propose it is worth articulating our contribution relative to Sullivan et al. Our measure of religiosity is different – intensity of belief instead of affiliation based. Especially in the context of faith of upbringing, wider cultural factors could be involved beyond religiosity. Our data provides a rich range of other school characteristics from administrative data to control for aspects of faith schooling that the Sullivan et al analysis does not. And, as the authors point out, their setting is schooling in the 1970s and the effects on the individual's examined at age 42. Hence our setting contributes to the literature by providing a more recent analysis (from 2004) and shows the short to medium-term impact of faith schooling and religiosity instead of the long-term effect.

A second paper that is close to ours is Adamczyk (2009) who uses the same religiosity measure as we have (intensity of belief, along with frequency of prayer and religious practise), and an indicator of Catholic school attendance, to estimate the impact of religiosity on the likelihood that a woman has had a premarital abortion in the United States. Neither religiosity nor religious practise have a significant impact, although being a more conservative Protestant does. Having more conservative Protestant peers has an impact but attending a religious school does not. The paper uses a hierarchical logistic regression that will be vulnerable to omitted variables. The range of controls are not as rich as contained in our paper, and the methodology does not address the robustness of the estimates reported.

Besides the papers above, there is evidence of religiosity impacting health and risky behaviours (Mendolia, Paloyo, and Walker, 2019); voting behaviour (Spenkuch and Tillmann, 2018); and the likelihood an individual pays their taxes (Torgler, 2006). We have not located any previous papers in the faith schooling literature that control for the intensity of belief. Neither is there existing work that matches our broad set of outcomes or provides such a rich range of controls.

3 Data

3.1 Institutional Background

3.1.1 Faith Schools

Religious institutions have been involved in English education for centuries - since the earliest schools were established. Historically, these schools were organised and run by religious institutions such as monasteries, and they were private in the sense that they were not maintained by the state even if they did not charge for the provision of their services. The 1902 Education Act brought free, compulsory, and Christian education for all to England, and most schools became part of the state-maintained system. This continued under the 1944 Education Act where faith and non-faith schools became distinct tracks of school (Department for Children, Schools and Famililies, 2007).

Faith schools at the time of the Next Steps cohort are generally voluntary controlled (VC) or voluntary aided (VA). In a VC school a religious body has influence in how the school is run but the school is mainly managed by the local authority (LA). In the VA case a religious institution may hold a stake in the buildings the school inhabits (or even own them completely) and have some small financial involvement in the school's operation. The religious body will also have a majority on the school's governing body (New Schools Network, 2015). More Church of England schools are VA than VC, while Catholic Schools are exclusively VA.[5]

A faith school in England is any that has an explicitly stated religious character. Whilst every LA in England has at least one faith school, there is a high degree of heterogeneity between LAs with nine LA areas in England having around 40% of their pupils in faith schools (Andrews and Johnes, 2016). Faith schools are allowed to use religious belief as a criterion for admitting pupils, for up to half of their pupils, if they are oversubscribed. Schools can be of a number of different denominations but the overwhelming majority are Christian. Of these, the lions' share are Roman Catholic (9.4% of all schools), with a smaller number being Church of England (6.1%) or of other Christian affiliation (2.3%). Jewish schools have existed since 1732, Muslim schools since the 1950s, and Sikh and Hindu schools since 1999 and 2008 respectively.

Crucially, as regards the ethos of a particular institution, or of the people who staff it, faith schools are allowed to apply religious criteria in their hiring processes. In practise, this means

[5] After the growth of Academy Schools and the inception of Free Schools, which are both state funded but independent of local authorities, the picture has become more fragmented. At the time of the Next Steps cohort, though, there were relatively few secondary academies and there were no Free Schools.

being able to choose one person over another if that person's beliefs align with that of the school. They also have freedom over what they choose to teach in Religious Studies classes, a GCSE level subject taught widely in schools but is outside of the National Curriculum.

It is clear that what is meant by a faith school in the context of England is distinct from what would be meant in the US context. In the US, Catholic schools are usually not public funded and operate very differently to an English faith school.

3.1.2 Key Stages

Children in the UK attend primary school from the ages of 4 or 5 up until age 11. Secondary schooling follows from ages 11 to 16. This applies to all students, and stratification occurs post-16 with the option to go into vocational training, as part of an apprenticeship scheme or full-time, or further academic studies.[6] Within primary and secondary schools, students are organised into 'Key Stages' (KSs). These are referred to as: KS1, which covers ages 5 to 7 (years 1, 2, and 3); KS2, which covers ages 8 to 11 (years 4, 5, and 6); upon moving to secondary school, ages 12 to 14 (years 7, 8, and 9) fall into KS3 with ages 15 to 16 (years 10 and 11) and 17 to 18 (years 12 and 13) belonging to KS4 and KS5 respectively. The Next Steps data covers KS3 onwards with some KS2 characteristics available.

At the end of each of these stages there are tests or national exams; GCSEs are completed at the end of KS4 and constitute the exit examinations from secondary schooling, whilst KS5 ends with the A-Level national examinations. KS5 is far narrower than the earlier KS levels with the typical student taking just three or four subjects. In contrast, KS4 normally includes 7-10 subjects with limited electives; 5 passing grades, usually including maths and English, are required to pursue an academic track, post-age 16, at KS5. One's A-level subjects, and the grades attained, determine access to undergraduate courses in universities.

3.2 Next Steps Data

This paper uses the Next Steps dataset (also known as the Longitudinal Study of Young People in England (LSYPE)).[7] The dataset is a cohort study beginning in 2004 with the sampling of approximately 21,000 Year 9 (KS3) pupils from 647 English State and Independent (fee-

[6] Since 2015 young people in England born after 1997 have had to stay in some form of education up until the age of 18. This does not apply to the Next Steps cohort, the overwhelming majority (99.8 percent among first wave participants) of whom were born in 1989 and 1990. As such, the Next Steps cohort face a leaving age of 16.

[7] The data and documentation can be found at University College London (2020). A full description of the dataset and its history can be found at Centre for Longitudinal Studies (2018).

paying) schools.[8] Questions were then asked of both the cohort member and their parents. In the first wave 74 percent of those contacted responded, yielding a sample of 15,770. Waves two to seven have response rates in excess of 85 percent (with the exception of the ethnic boost sample in wave 4 that has a 59 percent response rate). This represents a low level of attrition.

The study followed the cohort member (aged 13 or 14 in wave one) through their remaining years of education and up until the age of 20. The study then stopped until it was resurrected by the Centre for Longitudinal Studies for an eighth wave in 2015. This enables the same individual to be observed from age 13 to 25 and allows analysis to cover the impact of religiosity at an early age on later life outcomes. Whilst waves two to seven only include people who responded to the previous wave, wave eight saw all initial members (i.e. those who responded to wave 1) contacted and an ultimate sample size of 7,707 individuals (51 percent response) was achieved. We dropped from the wave one sample those who declared they were not Protestants, Catholics or of no religion. The numbers who declared themselves to be Muslims or to belong to other religions were too small to allow meaningful analysis, and only a small percentage of such people attended faith schools in our data.

This paper primarily makes use of Next Step's first wave (where respondents are 13 or 14 years old) with outcomes appearing from the eighth wave and some potential mediators coming from the second wave. Variables relating to religiosity, gender, ethnicity, and parental characteristics come exclusively from the first wave. Wave eight is used for information regarding university attendance, income, and other labour market outcomes. The sixth and seventh waves are used only to provide information on university attendance for those who were not present in wave eight. The outcome variables of interest are test scores at various different key stages, university attendance and performance, and the wage rate earned at age 25, eleven years after the faithfulness questions are asked. Christian affiliation at age 25 is also recorded.

In terms of test score outcomes: attaining the five A*-C grades benchmark is a binary outcome; number of GCSE passes is ordinal; and the GCSE point score is more granular and is calculated by attaching a score to each grade attained.[9] A-level attainment is measured by the University and College Admissions Service (UCAS) equivalent point score for the top three grades attained (even if more than three subjects were taken).[10] The UCAS score is used by universities

[8] In our analysis we examine state-maintained schools only.

[9] At GCSE, A* is worth 58 points, A is worth 52, B is worth 46, decreasing by 6 points until a grade G which is worth 16.

[10] For A-levels, when converted to UCAS points, A* is worth 140, A is worth 120, B is worth 100, C is worth 80, D is worth 60, and an E is worth 40.

when they review applications; focussing on the top three subject studied means that somebody who attained an A and three Bs does not appear to be performing better than somebody with three As. But the person with three As would, in practise, be much more likely to be accepted by a "better" university than the mostly-B-grade achieving student. We standardise both the GCSE point score and UCAS A-level point score.

The later, post high school, outcomes are defined as follows: university attendance is a binary variable taking the value one if the individual attended university by the time they were aged 25; Russell Group university attendance is a binary outcome indicating whether an individual attended one of the more prestigious research-intensive universities in the UK; degree classification is a binary variable that takes the value one if the individual was awarded a first class or upper second class degree at university and 0 otherwise (around ⅔rd of this cohort will have attained such a "good" degree). The wage rate is the log of labour income (in the individual's main job) divided by the number of hours worked per week in that job. Christian belief at age 25 is a binary variable that takes value one if the individual identifies as Christian in wave eight.

The Next Steps dataset is linked to the National Pupil Database (NPD). This is the government's administrative dataset for education in England.[11] This gives access to the academic achievement of the individuals in the Next Steps cohort and allows the denomination of the school the individual attended to be identified (i.e. faith versus non-faith). Additionally, a broad collection of school level characteristics are available, such as the ethnic mix of the school and the percentage of children eligible for free school meals (FSM), a proxy for lower socio-economic status.

Also available is the index of multiple deprivation (IMD) which measures the socio-economic status of the neighbourhood in which the cohort member lives, down to the postcode level (an average of 15 households). The multiple deprivation measure incorporates local levels of variables such as income; employment; health and disability; education, skills and training; barriers to housing services; and living environment and crime levels. Altogether, the available set of controls is extensive and suggests that a higher proportion of selection might be observed and explained than would be the case for other datasets.

[11] Education is a "devolved" policy area in the UK meaning that each nation within the United Kingdom has control over its education policy.

3.3 Measuring Religiosity

Religiosity is measured in Next Steps in accordance with recommendations made in McAndrew and Voas (2011). Three different facets of belief are covered: affiliation (or *extrinsic* religiosity); belief (or *intrinsic* religiosity, which we refer to as faithfulness in this paper) which we derive from the religion's importance in one's life question; and practise, which we derive from in the questions relating to participation in religious classes. The precise questions are given in Appendix Table A1.

The primary religiosity measure in this paper is 'faithfulness' or 'intrinsic religiosity', as it is often referred to. We use faithfulness throughout to more concretely refer to the treatment of interest, since intrinsic religiosity could easily, and confusingly, refer to other measures of practise (e.g. religious class attendance) as well as belief. Our faithfulness variable is the individual cohort member's response to the question: 'How important would you say your religion is to the way you live your life?' with four potential answers: "Not at all", "Not very", "Fairly" and "Very" faithful.

In the survey, those without a religion are not asked the faithfulness question, and are therefore coded here as not at all faithful. This variable is arguably superior to other measures of religiosity that are often used – variables based on affiliation and practise. For example, in a notionally Christian country, an individual may self-report being a Christian in survey data when, in practice they are not. Similarly, a young person may be taken to church by their parents but neither wish to be there nor care about what they hear whilst there. Thus, our definition, although subjective, does reduce a more explicit form of measurement error that could otherwise occur.

3.4 Descriptive Statistics

There are two treatments of interest – faithfulness and faith schooling. Additionally, interaction effects are examined later. If the faithfulness variable is kept as is, with four categories and with the separate faith school treatment then there will be four coefficients of interest (seven in the interaction case). This is arguably too many, since cell-sizes begin to get too small, to support precise estimates of the interaction treatments. Moreover, the Oster (2019) test, outlined below, is designed for binary treatments. We, therefore, collapse the three faithfulness treatment levels above the lowest into one, leaving a binary treatment that is 0 if the individual is not faithful, and 1 if the individual is faithful. Pairwise tests of coefficients each of the four

faithfulness levels (not shown) suggest that this is reasonable, with few statistically significant differences.

Figure 1 shows the percentage of each faithfulness level in the data, before the it is collapsed to binary. Figure 2 shows the way that faithfulness is broken down by Christian denomination and by Faith and Non-Faith School attendance. Those of no religion are, by construction, classified as not at all faithful. There are clearly more Protestants who are of lower faithfulness than there are Catholics of lower faithfulness. The same is true of non-faith schools relative to faith schools, as would be expected, with over half of non-faith school attendees being not at all faithful.

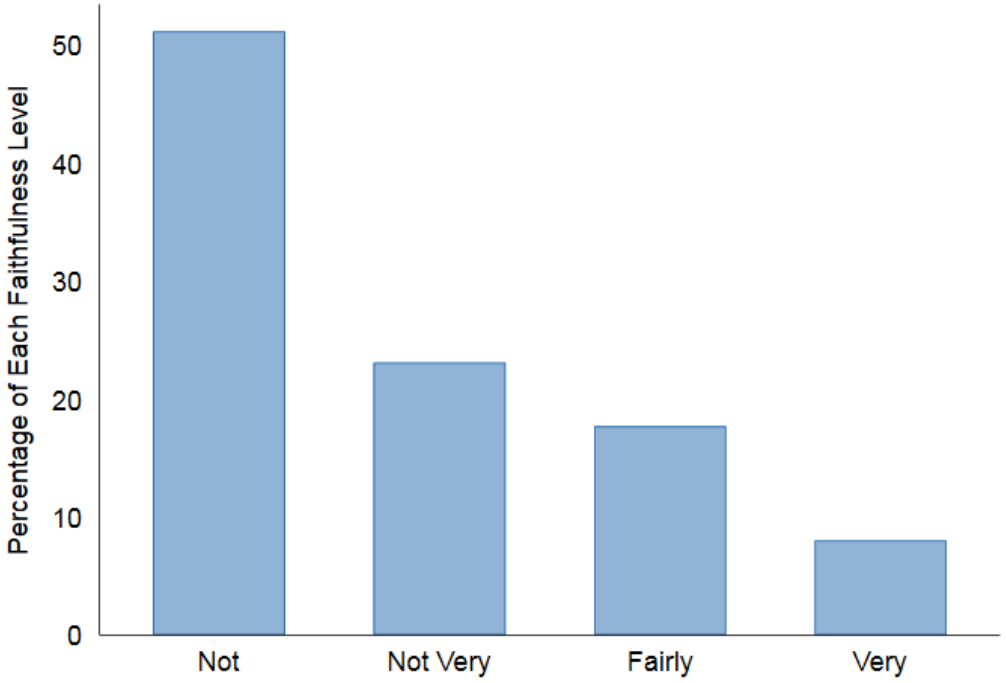
Table 1 shows summary statistics of cohort member's characteristics and Table 2 shows the outcome variables, broken down by faithfulness and school type.[12] The percentage of pupils on free school meals is very similar across all categories as are non-cognitive skills. Those attending faith secondary schools performed better at KS2 (i.e. at primary school), as did the faithful.

In Table 2 the faithful perform better in all of the schooling outcomes, but there are very small differences in the later outcomes, except for being a Christian at age 25 where there is a very large difference. There are also difference by school type in schooling outcomes and while there is a marked positive effect of faith schooling on attending university, there is no difference in university outcomes conditional on attending university. In the raw data, there appears to be a negative effect of faith schooling on income, but not of faithfulness. The faithful also appear to work longer hours per week – something that may reflect an independent effect of religion that has been noted in some previous work and motivates us to focus on a wage rate defined by the hourly rate of pay in analysis later in this paper.

Figures 3 and 4 selected outcome variables broken down by faith school and faithfulness respectively. These support the popular notion that faith schools have better educational outcomes. However, faithfulness shows patterns of effects on outcomes that are very similar to the effects of faith schooling. Figure 4 shows the potential for an interaction effect between faithfulness and faith schooling – there are more faithful individuals in faith schools.

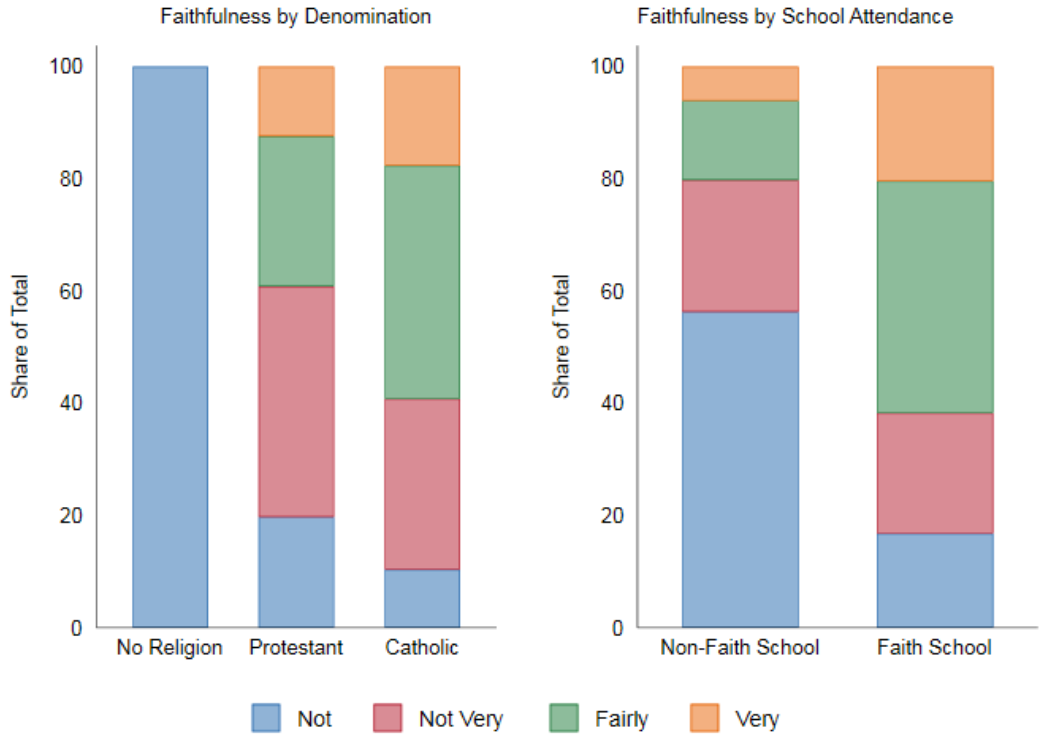
[12] Tables A2 and A3 in the Appendix give summary statistics for parental/household characteristics and school level characteristics respectively.

Figure 1 - Distribution of Faithfulness



Note: Displays levels of faithfulness where the survey responses were as appear with the word faithful after: e.g. "Not very" faithful. This is with the exception of "Not" which represents the survey response "not at all faithful".

Figure 2 - Faithfulness by Religion and School Attendance



Note: Displays levels of faithfulness where the survey responses were as appear with the word faithful after: e.g. "Not very" faithful. This is with the exception of "Not" which represents the survey response "not at all faithful".

Table 1 - Summary Statistics - Individual Characteristics

	Faithfulness			School Type		
	Unfaithful	Faithful	Total	Non-Faith	Faith	Total
			Mean/SD/N			
Faithful	0.00	1.00	0.49	0.44	0.83	0.49
	(0.00)	(0.00)	(0.50)	(0.50)	(0.37)	(0.50)
	5837	5568	11405	9732	1517	11249
Faith School	0.04	0.23	0.13	0.00	1.00	0.14
	(0.21)	(0.42)	(0.34)	(0.00)	(0.00)	(0.34)
	5733	5516	11249	9792	1531	11323
Religion	0.23	1.23	0.72	0.61	1.45	0.72
	(0.48)	(0.42)	(0.67)	(0.59)	(0.68)	(0.67)
	5837	5568	11405	9792	1531	11323
Gender	0.48	0.52	0.49	0.48	0.52	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
	5837	5568	11405	9792	1531	11323
Ethnicity	0.09	0.25	0.17	0.14	0.30	0.17
	(0.29)	(0.43)	(0.37)	(0.35)	(0.46)	(0.37)
	5837	5568	11405	9792	1531	11323
KS2 Average Score	553.20	579.08	565.73	562.43	591.30	566.37
	(212.25)	(206.58)	(209.92)	(212.28)	(190.42)	(209.66)
	5555	5210	10765	9241	1457	10698
KS3 Average Score	911.64	965.71	938.29	932.90	985.02	940.03
	(380.85)	(368.90)	(375.96)	(381.17)	(336.94)	(375.84)
	5575	5418	10993	9426	1492	10918
FSM Eligible	0.12	0.10	0.11	0.11	0.11	0.11
	(0.33)	(0.31)	(0.32)	(0.32)	(0.31)	(0.32)
	4548	4451	8999	7745	1194	8939
Locus of Control	2.77	2.80	2.78	2.78	2.84	2.79
(Wave 2)	(0.76)	(0.78)	(0.77)	(0.77)	(0.76)	(0.77)
	4240	4126	8366	7176	1140	8316
Self-Esteem	6.77	6.81	6.79	6.79	6.79	6.79
(Wave 2)	(1.62)	(1.56)	(1.59)	(1.59)	(1.60)	(1.59)
	4481	4397	8878	7609	1209	8818
Sociability	6.37	6.23	6.30	6.33	6.07	6.29
(Wave 2)	(1.93)	(1.94)	(1.94)	(1.94)	(1.91)	(1.94)
	4216	4161	8377	7218	1118	8336
Max N	5837	5568	11405	9792	1531	11323

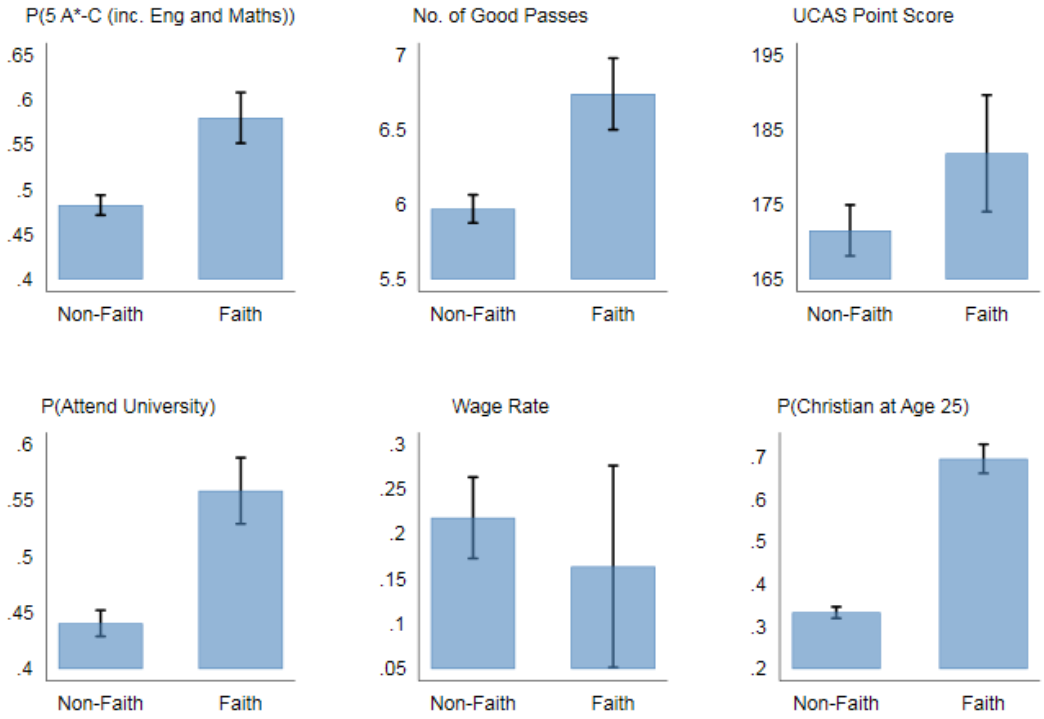
Note: Faithful is a binary indicator, 0 for unfaithful, 1 for faithful; faith school also takes the value 0 for a non-faith schools and 1 for a faith school. Religion is coded from 0 to 2, 0 is no religion, 1 is Protestant, and 2 is Catholic; Gender takes value 1 if the individual is female and 0 if male; Ethnicity is 1 for non-white individuals and 0 otherwise; KS2 and KS3 point scores are continuous; FSM eligible takes value 1 if the individual is on free school meals; internal Locus of control goes from 1 to 4 with 4 being the highest feeling of control over one's life and 1 the lowest; Self-Esteem goes from 1 to 8 and with 8 being the highest self-esteem and 1 the lowest; Sociability also goes from 1 to 8 with 8 being the highest and 1 the lowest.

Table 2 - Summary Statistics - Individual Outcomes

	Faithfulness			School Type		
	Unfaithful	Faithful	Total	Non-Faith	Faith	Total
			Mean/SD/N			
Achieved Five A*-C Grades	0.56	0.66	0.61	0.60	0.70	0.61
	(0.50)	(0.47)	(0.49)	(0.49)	(0.46)	(0.49)
	4587	4477	9064	7803	1202	9005
Achieved Five A*-C Grades (incl. English and Maths)	0.45	0.54	0.49	0.48	0.58	0.50
	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)
	4587	4477	9064	7803	1202	9005
Number of Good Passes	5.53	6.57	6.04	5.97	6.74	6.07
	(4.28)	(4.13)	(4.24)	(4.26)	(3.98)	(4.23)
	4587	4477	9064	7803	1202	9005
GCSE Point Score	360.81	401.63	380.97	378.35	406.70	382.14
	(160.92)	(145.44)	(154.82)	(155.66)	(142.00)	(154.20)
	4587	4477	9064	7803	1202	9005
A-Level Point Score (Top 3 Grades)	167.98	176.87	172.99	171.56	181.83	173.20
	(122.85)	(120.77)	(121.75)	(122.06)	(119.73)	(121.73)
	2567	3310	5877	4926	939	5865
Attended University	0.38	0.53	0.46	0.44	0.56	0.46
	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
	4050	4082	8132	7006	1092	8098
Degree Class	0.73	0.70	0.71	0.71	0.70	0.71
	(0.44)	(0.46)	(0.45)	(0.45)	(0.46)	(0.45)
	504	700	1204	1012	190	1202
Attended Russell Group University	0.23	0.21	0.22	0.22	0.22	0.22
	(0.42)	(0.41)	(0.42)	(0.42)	(0.41)	(0.42)
	1413	1979	3392	2822	562	3384
Weekly Income	319.06	321.60	320.35	321.97	310.56	320.42
	(67.08)	(68.47)	(67.80)	(68.07)	(65.14)	(67.79)
	2707	2818	5525	4770	747	5517
Hours Worked	38.85	39.30	39.09	39.03	39.37	39.08
	(11.45)	(11.39)	(11.42)	(11.48)	(10.99)	(11.41)
	2145	2366	4511	3880	629	4509
Christian Age 25	0.18	0.57	0.38	0.33	0.70	0.38
	(0.39)	(0.50)	(0.49)	(0.47)	(0.46)	(0.49)
	2611	2729	5340	4612	721	5333
Max N	4587	4477	9064	7803	1202	9005

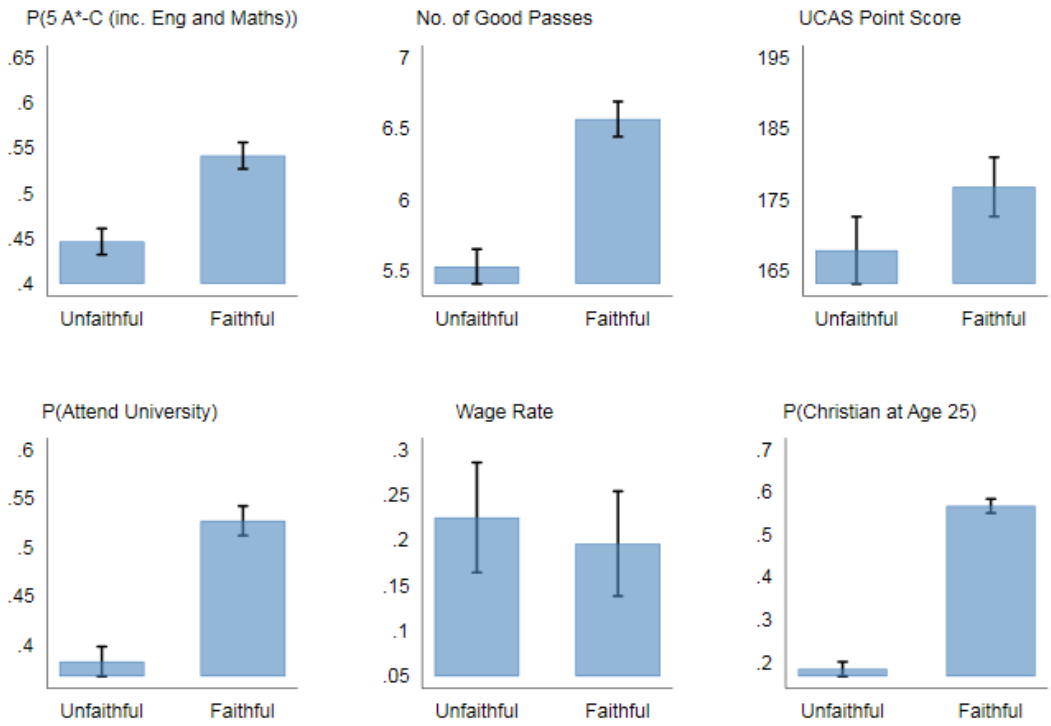
Note: Outcomes are described in detail on Page 12. Five A*-C, Five A*-C (incl. English and Maths), Attended University, Degree Class, Attended Russell Group University, and Christian at Age 25 are all binary. Each takes value one if the condition is true. Degree class takes the value one if the individual got a first or an upper second class degree, and 0 otherwise. Wage, GCSE point score and A-Level point score are continuous variables. Number of Good Passes is discrete.

Figure 3 – Mean Outcomes by School Attendance (Faith vs Non-Faith School) (95% CI)



Note: Chart shows selected outcomes for those who are "Unfaithful" (or not faithful) and "Faithful" individuals. Probability of attaining 5 A*-C grades, probability of attending university, and probability of being Christian at age 25 are all binary; number of good passes, UCAS (A-level) point score, and wage rate are not.

Figure 4 – Mean Outcomes by Faithfulness (Faithful vs Unfaithful) (95% CI)



Note: Chart shows selected outcomes for those who are "Unfaithful" (or not faithful) and "Faithful" individuals. Probability of attaining 5 A*-C grades, probability of attending university, and probability of being Christian at age 25 are all binary; number of good passes, UCAS (A-level) point score, and wage rate are not.

4 Empirical Strategy

4.1 Specification

Our analysis begins with an Ordinary Least Squares (OLS) estimation of a linear specification:

$$Y_{is} = \beta_0 + \beta_1 F_{is} + \beta_2 FS_{is} + \beta_3 \boldsymbol{X}_{is} + \beta_4 \boldsymbol{S}_i + \epsilon_{is} \quad (1)$$

where Y_{is} is some outcome for individual i in school s: GCSE attainment, A-level attainment, whether or not the individual attends university, attends a Russell group university, attains a "good" degree class, the wage rate at age 25, and whether or not they are a Christian at age 25. F_{is} is a binary variable taking the value zero if the individual says their faith is not at all important in their everyday life, and one if the individual says their faith is more important than that (i.e. not very, fairly, or very faithful). FS_{is} takes the value zero if the individual did not attend a faith school, and one if they did. $\boldsymbol{X}_{is}$ is a vector of controls including gender, ethnicity, religion, parental religion, parental education, parental employment status, and number of dependent children in the household. $\boldsymbol{S}_i$ is a vector of school level characteristics of the cohort member. These include the ethnic mix, the share of pupils on free school meals (FSM), whether the school has a single sex intake, is academically selective, and has a "sixth form" (senior high school) attached to it. Standard errors (ϵ_i) are clustered at the school level as this is the primary randomisation unit for the data sampling. Evidently the set of controls is both rich and varied.

Religiosity is measured when the cohort members are aged 13 or 14 – in Next Steps' first wave. Though the faithfulness question is asked in subsequent waves the analysis is based on the first wave information only. This is to ensure that our measure of religiosity is recorded pre-treatment – if we were to use wave 3 faithfulness, after GCSE high stakes exams have been taken, there may be an issue of reverse causality between attainment and religiosity. As there is no quasi-experimental variation here it makes sense to minimise issues such as this.

Sensitivity analysis in empirical research is traditionally conducted by observing how treatment effect estimates change as additional control variables are included; if there is little movement in the estimated treatment effects then the threat of unobserved selection is said to be low. However, as pointed out in Oster (2019) this may not be enough.[13] Hence, we augment the

[13] The example, in the introduction of Oster (2019), is the effect of education on wages. There are two orthogonal components of ability, one that has high variance and the other low variance – if both were included all variation would be explained. Controlling for the low variance ability component would not change coefficient sizes all that much – leading to the conclusion that selection on unobservables was not an issue. But the bias would still exist by omitting the second, high variance, ability control.

OLS estimates with the test suggested in Oster (2019). The test extends prior work by Altonji, Elder, and Taber (2005b). Their paper suggested that it might be reasonable to assume that the amount of selection on unobservables could be bounded from above by the amount of selection on observables. If covariates to be included in estimations were picked at random from the full set of possible covariates, selection on unobservables would be less than or equal to selection on observables. As researchers do not pick covariates at random but based on other empirical studies and theoretical justification for their inclusion, in reality selection on unobservables in a rich data set is likely to be less than that which is observed and controlled for. Bounds on OLS estimates can also be produced using their method.

Oster (2019) points out that observed selection is only informative about unobserved selection if the two are distributed in the same way. Assuming that it is in a rich dataset, it will be the case that explaining all variation, i.e. attaining an R^2 value of one, is impossible. This is due to measurement error in research data. As a result, the Oster test provides a procedure to use the observed R^2 value from estimated regression specifications multiplied by something larger than one. Oster suggests, on the basis of comparison of randomised controlled trial estimates with non-experimental estimates from a range of previous studies, that 1.3 would be appropriate. More conservatively, estimates are also provided in the tables below that use double the R^2.

The test can be used in two ways – firstly to infer the degree of unobserved selection the that would need to exist to reduce the magnitude of the treatment coefficient to zero. This is the δ value. The threshold for robustness in this case is one – equal observed and unobserved selection. The second way is to bound estimates assuming a particular degree of unobserved selection – the β value. The test is not a silver bullet that enables causal inference, but it substantially augments the usefulness of OLS estimates in that it may allow researchers to argue selection bias is unlikely to bias estimates substantially.

It is useful to articulate the nature of the expected omitted variable bias. The most obvious is likely innate ability. For example, in Table 1 the faithful and those attending faith schools have higher KS2 scores than the unfaithful and those in secular schools. As test scores are likely to only imperfectly capture an individual's true innate ability, there will be omitted sources of ability that could have an impact on estimates. Family background would behave similarly. As such, the expected sign of omitted variable bias on our outcomes of interest is positive, and so if estimates are indeed vulnerable to it, they will be biased upwards.

We also employ Inverse Probability Weighted Regression Adjustment (IPWRA) as an alternative way of better ensuring robustness (see Imbens and Wooldridge (2009) for a more in-depth description of the method). IPWRA models both the treatment (faithfulness or faith school) and the outcome in two separate equations. Taking the treatment equation first, a propensity score is estimated that suggests the probability of treatment based on included observables. This propensity score is then used to weight the second stage in an attempt to generate better counterfactuals and so strip out the possibility of selection into treatment from the outcome equation.

Based on selection on observables, IPWRA can get closer to causal estimates than OLS by accounting for two levels of selection – in treatment and outcome. It also possesses the so-called "double robustness" property that means it produces consistent estimates if only one of the two equations is incorrectly specified. In the analysis below, IPWRA is conducted on one treatment at a time controlling for the other treatment, as in the OLS specifications. The estimate of the propensity score in the first stage requires there to be sufficient "overlap" – that both treatment and control groups have a similar distribution of propensity scores.

As a degree of experimentation occurs in the selection of covariates in order to produce sufficient overlap, there might be concerns about cherry-picking the specification that yields the results that look most desirable. We avoid this by randomly generating a variable that is used as the outcome variable until the specification that will be used for subsequent analysis has been chosen on the basis of balance and overlap. In our case, the same treatment equation (the first stage) ultimately produced good overlap for both the faithful and faith school treatments.[14] The coefficient balances and overlap figures are given in Tables A15 and A16, and Figures B2 and B3 respectively.

4.2 Mediation

Once the effects of faithfulness or faith schooling are identified it is useful to try to explain the mechanism(s) through which those estimates operate. One set of potential mediators are non-cognitive skills or personality traits, that are recorded in Next Step's second wave. We refer to them as non-cognitive skills from here onwards, though there is a suggestion (see e.g. Borghans et al. (2008)) that to refer to them as such suggests these traits are "devoid of cognitive processing". The literature often refers to the Big Five factors, known as such since Goldberg

[14] These variables were gender, FSM status, KS2 achievement, IMD, mother's education, mother's age, number of dependent children in the household, region of residence, whether the individual has a single mother, and whether either of the parents was aged less than 20 when the individual was born.

(1971); these are – Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticisim (also referred to as OCEAN). Though there have been other suggestions, such as that from the University of Chicago Consortium on Chicago School Research, that conclude that those non-cognitive skills most related to academic performance (the context in which we are working) are academic behaviours (such as participation in class), academic perseverance (hard work), academic mindsets (e.g. belief that ability can grow with effort), and social skills (Farrington et al. 2012; Bjorklund-Young 2016).

The non-cognitive skills that we have in Next Steps that relate to these are work ethic, internal locus of control (the degree to which one believes they can shape their own outcomes), self-esteem, and sociability. The survey questions relating to these variables are in Table A2. As well as being related to academic outcomes, these traits may be linked to religious belief. Work ethic has an association with religion, stretching back decades in sociology to the work of Max Weber, through the idea that Protestants are called upon to work hard for its own sake (Weber, 2001). Locus of control may be lower among those who think that an external force has determined what will happen in their lives. Self-esteem could be higher as depression has been shown to be higher among those who can use their faith as a form of support mechanism (Fruehwirth, Iyer, and Zhang, 2019). Equally, sociability could make an individual better at team-working or studying with others, and this could be improved by faithfulness if that makes one attend church social events. Taken together this suggests their suitability as mediators.

The mediation analysis is based on Acharya, Blackwell, and Sen (2016). Their analysis stems from the observation that including potential mediators that are simultaneously determined with the treatment could risk biasing the treatment effect of interest through "intermediate variable bias", where some unobserved factor is correlated with the potential mediator, the treatment, and the outcome. They generate examples that suggest this could be a genuine issue and they apply their method to previous political science empirical examples to show how results change with their method, relative to the usual approach of simply including the potential mediator as a control in a single stage estimation.

Their method estimates what they term the "average controlled direct effect" (ACDE) - the effect of the treatment on the outcome when the effect of the mediator is fixed at some value for all units. The method is implemented in two steps via "sequential g-estimation". First, we estimate the effect of the potential (post-treatment) mediator (M_{is}) on the outcome, controlling for treatments (F_{is} and FS_{is}), pre-treatment controls ($\mathbf{X}_{is}$ and $\mathbf{S}_{is}$, notation as above), and

"intermediate" (or post-treatment) control variables (I_{is}) that are contemporaneous to the potential mediator.

$$Y_{is} = \beta_0 + \beta_m M_{is} + \beta_1 F_{is} + \beta_2 FS_{is} + \beta_3 \boldsymbol{X}_{is} + \beta_4 \boldsymbol{S}_i + \beta_I I_{is} + \epsilon_{is} \qquad (2)$$

In the education timeline our first outcomes (GCSE attainment) come from wave three (academic-year 11). All controls in OLS and IPWRA specifications prior to mediation analysis come from wave one. This enables mediation analysis to take place using wave two variables.

The pre-treatment controls in our setting are school characteristics as these are from wave one (year 9) and individual and household characteristics from wave one. Most of the controls available are "fixed" in the data, such as gender, or school characteristics which only exist for wave one. As such, the intermediate controls (I_{is}) are mother's education and mother's employment status interacted with dummies denoting missingness for father characteristics as well as parental marital status and faithfulness. The mediators are non-cognitive skills listed in Table A4 of the Appendix.

Next, we transform the outcome variable by the estimated coefficient ($\tilde{Y} = Y - \beta_m M_{is}$) to *demediate* the dependent variable and run the second stage with just the treatment and the pre-treatment controls to estimate the effect of the treatment on the demediated outcome. The resulting impact should show how the treatment acts on the outcome independently of any post-treatment factors. The difference between the initial treatment effect and the treatment effect after the mediation analysis is the impact of the mediator.[15]

In practise, we perform principle component analysis (PCA) on the four mediators and include the first principle component as a mediator – this is because we are interested in the underlying variation these survey responses capture rather than the specific coefficients attached to each mediator. The first component has an eigenvalue greater than one, meaning that it passes the Kaiser-Guttman criterion and are said to summarise more variation than any single variable (Guttman 1954; Jackson 1993). The second component's eigenvalue is just below one in the case shown below (the Protestant sample) and so is not included initially (though we conducted robustness tests with it included though these are not shown in this paper). The Screeplot is also provided in the Figure B1.

[15] See Huber (2019) for an overview of mediation analysis.

5 Results

5.1 OLS Specifications

Regression results are presented below (with additional tables in the Appendix) and the pattern of controls is the same for each OLS table. The whole sample includes those of no religion, Protestants, and Catholics. Controls for the two Christian denominations are added in column (2). Individual characteristics are added in column (3). These are gender, ethnicity, month of birth, month of interview, and the individual's academic performance at KS2. Parental/household characteristics (added in (4)) are the index of multiple depravation (IMD), whether the child is on free school meals (FSM), mother's education, mother's ethnicity, mother's employment status, whether the child has a single mother, the number of dependent children in the household, and the region of residence. The mother's employment and education are interacted with whether the father's characteristics of the same variable are missing. Parental belief (in their religion and how important it is to them) is added in column (5). School characteristics (added in (6)) are whether the school has a particular specialism (for which they had been awarded additional funding), the percentage of students on FSM, whether the school is academically selective, whether the school has a "sixth form" (a senior high school for post-compulsory education), the size of the school and the size of the previous school attended, the percentages of students who have special needs, who are white, speak a first language that is not English, and whether the school has single sex intake.[16] The column (6) controls are used in the mediation analysis and in the outcome equation of the IPWRA later.

Tables 3 and 4 report regression results for two GCSE outcomes; whether or not the individual attained 5 A*-C grades, including English and Maths, and the number of "good" passes that the individual achieved – i.e. how many grades were they awarded at C or above. Columns (1) to (6) show results for the whole sample whilst columns (7) and (8) show only the Protestant (P) and Catholic (C) subsamples. Each table, as will be the case in each of the OLS tables, presents four panels. The first (Panel A) shows the regression results, across numerous specifications, where faithfulness is not included - the only "treatment" is attendance at a faith school. The second (Panel B) is the opposite, faith school is not included - the only treatment is whether the individual is faithful. Panel C includes both treatments together. The reasoning behind presenting the results in such depth is to show the stability of coefficients upon the inclusion of both treatments of interest together. An obvious concern if only panel C was shown

[16] Without a sixth form, students wishing to pursue an academic track post-age 16 would go to a separate institution.

would be that one treatment was sapping the significance associated with the other due to the obvious correlation between being more devout and wanting to attend a faith school. This concern is all the more valid considering the papers cited above that suggest a positive effect of faith school attendance on religiosity (e.g. Wadsworth and Walker 2017)). Panel D, to be explained later, presents the Oster (2019) test results.

Looking at Table 3 it appears that both faith school and faithfulness have impacts on the likelihood that an individual attains the benchmark of five A*-C grades (with the added condition that those grades include English and Maths) in a linear probability model. These are each significant; at the five and ten percent levels respectively. The magnitudes are not insubstantial, attending a faith school appears to increase the likelihood of attaining the benchmark by around four percentage points, and being faithful compared to unfaithful by three percentage points. Comparing the coefficients in Panel C to the corresponding coefficients in Panels A and B, it is easy to see that the inclusion of both treatments simultaneously does not seem to alter the coefficient magnitudes by any meaningful amount – indeed the difference is never different in a statistical sense.

Turning to Table 4, where the outcome is number of good passes at GCSE, a number of points stand out. The first is that faith schooling does not appear to have an impact once exogenous individual characteristics (gender, ethnicity, month of birth, and prior attainment at primary school) are accounted for. The second is that faithfulness does – and it has a large impact at that. It is also always significant at the one percent level, except in the Catholic case where the significance is at the five percent level in Panel C. Taking column 6, which includes the whole sample and the full range of covariates, it appears that around 0.4 of an additional pass could be gained by being faithful compared to unfaithful. These numbers, as in Table 3, are remarkably stable when comparing the panels that include the two treatments separately with their simultaneous inclusion.[17] This same pattern is repeated for the GCSE point score outcome – essentially the same outcome but more granular. This is given in the Appendix Table A6.[18]

[17] As we are interested in faith school effects, we cannot include school fixed effects as the faith school coefficient is omitted from regressions. When faithfulness alone is the treatment, however, we can and do include fixed effects (not shown) – the results are virtually unchanged from what is shown in this paper when the full range of controls is included.

[18] Interestingly, the same pattern is also replicated by subject. Regression results (not shown) for highest English grade attained, maths grade attained, and highest science grade attained show the same pattern as the number of good passes outcome. This suggests that faith schools, as well not being stronger overall, are not any stronger in particular subjects.

Table 3 OLS Results for Attained Five A*-C Grades (incl. English and Maths)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**Attained Five A*-C Grades (Including English and Maths) at GCSE**							
	Whole Sample						**P only**	**C only**
Panel A								
Faith School	0.102***	0.075**	0.035	0.043**	0.032*	0.042**	0.053*	0.031
	(0.027)	(0.030)	(0.021)	(0.019)	(0.019)	(0.019)	(0.028)	(0.033)
N	9,005	9,005	8,660	7,931	7,905	7,849	3,761	957
R^2	0.004	0.011	0.427	0.455	0.456	0.462	0.457	0.533
Panel B								
Faithful	0.104***	0.060***	0.054***	0.042***	0.029*	0.031*	0.016	0.102**
	(0.013)	(0.021)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)	(0.045)
N	9,064	9,064	8,713	7,980	7,956	7,803	3,727	945
R^2	0.011	0.012	0.427	0.455	0.456	0.461	0.455	0.534
Panel C								
Faith School	0.066**	0.068**	0.028	0.037*	0.029	0.039**	0.052*	0.009
	(0.027)	(0.030)	(0.021)	(0.019)	(0.019)	(0.019)	(0.028)	(0.034)
Faithful	0.091***	0.053**	0.052***	0.039**	0.028*	0.029*	0.015	0.099**
	(0.013)	(0.021)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)	(0.046)
N	8,945	8,945	8,604	7,883	7,859	7,803	3,727	945
R^2	0.012	0.013	0.427	0.454	0.455	0.462	0.455	0.534
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		0.010	0.174	0.259	0.179	0.252	0.702	0.093
$R^2_{max} = 2\widehat{R^2}$		0.708	0.233	0.310	0.214	0.294	0.840	0.093
$R^2_{max} = 1.3\widehat{R^2}$		1.902+	0.776	1.032+	0.713	0.978	2.788+	0.269
Faithful								
$R^2_{max} = 1$		0.003	0.154	0.121	0.078	0.079	0.220	0.863
$R^2_{max} = 2\widehat{R^2}$		0.181	0.206	0.145	0.093	0.092	0.263	0.863
$R^2_{max} = 1.3\widehat{R^2}$		0.522	0.682	0.483	0.309	0.308	0.877	2.477+
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. + indicates passing of robustness threshold in the Oster (2017) test.

Table 4 OLS Regression Results for Number of Good Passes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**Number of Good Passes (A*-C) at GCSE**							
	Whole Sample						**P only**	**C only**
Panel A								
Faith School	0.843***	0.632**	0.225	0.243	0.163	0.215	0.363	-0.071
	(0.245)	(0.270)	(0.185)	(0.170)	(0.171)	(0.178)	(0.239)	(0.280)
N	9,005	9,005	8,660	7,931	7,905	7,849	3,761	957
R^2	0.004	0.016	0.530	0.561	0.561	0.563	0.557	0.594
Panel B								
Faithful	1.123***	0.711***	0.613***	0.518***	0.449***	0.430***	0.368***	0.658*
	(0.113)	(0.178)	(0.120)	(0.120)	(0.125)	(0.122)	(0.133)	(0.325)
N	9,064	9,064	8,713	7,980	7,956	7,803	3,727	945
R^2	0.017	0.018	0.531	0.562	0.563	0.562	0.557	0.590
Panel C								
Faith school	0.430*	0.551**	0.150	0.176	0.124	0.181	0.344	-0.205
	(0.241)	(0.267)	(0.184)	(0.170)	(0.171)	(0.178)	(0.238)	(0.291)
Faithful	1.015***	0.637***	0.589***	0.493***	0.432***	0.419***	0.362***	0.713**
	(0.112)	(0.178)	(0.121)	(0.121)	(0.126)	(0.122)	(0.133)	(0.361)
N	8,945	8,945	8,604	7,883	7,859	7,803	3,727	945
R^2	0.017	0.019	0.531	0.561	0.561	0.563	0.557	0.590
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		0.015	0.161	0.202	0.133	0.202	0.667	-0.396
$R^2_{max} = 2\widehat{R^2}$		0.751	0.161	0.202	0.133	0.202	0.667	-0.396
$R^2_{max} = 1.3\widehat{R^2}$		2.158+	0.475	0.526	0.347	0.523	1.762+	-0.917
Faithful								
$R^2_{max} = 1$		0.004	0.252	0.227	0.181	0.169	0.839	1.247+
$R^2_{max} = 2\widehat{R^2}$		0.202	0.252	0.227	0.181	0.169	0.839	1.247+
$R^2_{max} = 1.3\widehat{R^2}$		0.555	0.739	0.590	0.469	0.438	2.211+	2.865+
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Some other outcomes display significant results. A-level (UCAS) points (Table A7) show the same pattern as number of good passes – faith schooling does not have any impact that is significantly different from zero; faithfulness does for the whole sample and the Protestant-only sample. Coefficient movements between Panels A and B, and Panel C are a little larger in this case but still do not represent marked changes. University Attendance (Table A8) shows the same result. The remaining outcome which displays significant coefficients is Christian belief at age 25 (Table A12). This differs from the previous outcomes discussed in that significance at the one percent level is near universal. The magnitudes are large – a twelve percentage point increase in the likelihood that an individual is a Christian 11 years later for those who are faithful at age 14 compared to those who are not; and a 10 percentage point increase for those who attend a faith school compared to those who do not.

A number of other outcomes, shown in the Appendix, do not have robust results. Indeed, in some cases no statistically significant estimates are to be found. These outcomes are attendance at a Russell Group university (Table A9); degree classification (Table A10), where Catholics, whilst included in the whole sample, are not shown separately due to too small a sample size existing for that survey response; and the wage rate earned at age 25 (Table A11).

The results here present the pattern that will be repeated throughout – GCSE, A-level, university attendance, and Christian belief at age 25 are the outcomes where significant results are identified. It is these that will be the subject of robustness and heterogeneity analysis.

In Sullivan et al (2018), an affiliation-based measure of religiosity is used. We control for affiliation but, due to the way results are presented, have not yet shown the coefficients on Protestant and Catholic affiliation. Table A13 does this. Panel A shows the coefficients on religion in the equivalent to the column 6 specifications in Tables 3 and 4 but without including faithfulness. Panel B includes faithfulness. As can be seen, effects are generally attributed to affiliation-based religiosity until intensity of belief is included. This, we believe, adds weight to our analysis that faithfulness is a more relevant measure of religiosity to include.[19]

[19] Regressions are also run that replace the faithful treatment with attendance at religious classes whilst including faith school and all other controls from column 6 of Tables 3 and 4. The faith school coefficients tell broadly the same story as when faith school is included alongside faithfulness, whilst the religious classes coefficients are only significant in the case of Christian religion at age 25. This adds to the case that intensity of belief, as measured by the faithful variable, is indeed a better measure than other measures of religiosity such as practise. These are shown in Table A14. It is worth noting that we have some reservations about this variable; as the question given in Table A1 shows the question asks about religious classes – it is difficult to know how this would be interpreted by respondents – is it the same as attending church? Or is it some specific additional "class" such as a Bible study?

5.2 Oster Tests

5.2.1 Degree of Unobserved Selection Needed to Nullify Observed Results (Oster δs)

In each OLS table there is a fourth panel (D) not yet discussed. These present δ values resulting from Oster tests conducted on each specification other than the first, column (1), specification of Panel C[20]. The δ value represents the amount of unobserved selection that would be needed to drive the results, that have been identified, to zero - assuming that the unobserved selection takes a similar distribution to that which is observed.

Three rows appear in each Panel D. The first row in each assumes that the maximum possible proportion of variation that could be explained is one. In reality, this is unrealistically high due to measurement error in survey responses. The noise this creates means that explaining all variation becomes impossible. Additionally, if R^2 values get beyond 0.8, then suspicions of severe multi-collinearity may arise.

To be concise, only the final whole sample column, column (6), and the Protestant- and Catholic-only columns, columns (7) and (8), will be discussed. But δ values are given for each column giving rise to the possibility that the δ value is high enough to meet the standard in one column but not another. This is possible as the test assumes that the unobserved selection takes the same form as the observed selection, meaning that the test becomes more reliable the richer the range of included covariates. Where columns disagree, later columns take precedence.

The Oster test can be conducted with amendments made to the assumption that the maximum attainable R^2 is one. Instead the maximum R^2 can be set to be some multiple of the R^2 value that is observed in the regression specifications in each table. The standards adopted are an R^2_{max} of twice the observed R^2 from the estimated regression, and an R^2_{max} of 1.3 times the observed R^2.[21] If results withstand unobserved selection to the same degree as that which is observed (i.e. δ =1) then the result is thought to be robust. This sets a high bar. As is evident, the R^2 values for the later columns of Tables 3 and 4 range from 0.45 to 0.55 suggesting around half of all possible variation is explained. If true, it seems unrealistic to suppose that as much unobserved selection could exist. But given the non-experimental nature of the methodology employed, it makes sense to use such a stringent standard.

[20] The same tests were conducted for Panels A and B and showed the same patterns as those from Panel D.

[21] A good example paper is Clemens, Montenegro, and Pritchett (2019).

In Table 3, for the five A*-C grades benchmark, the faith school coefficient satisfies the Oster test standard for the Protestant-only sample, and virtually so for the whole sample, when using $R^2_{max} = 1.3\widehat{R^2}$. The faithful coefficient satisfies it for the Catholic-only sample. In Table 4, the faithful coefficient for the Protestant-only sample clears the threshold, as is the case for the Catholic-only sample. The whole sample does not satisfy the standard which could be due to greater heterogeneity among those in the full sample.

Taking the other significant results from the OLS tables in the Appendix, A-level point score (Table A7) has a faithful coefficient that is robust for both the whole sample and the Protestant-only sample. The whole sample faithful coefficient for the university attendance outcome (Table A8) was the only one that displayed significant results, but it does not meet the Oster standard. Christian belief at age 25 (Table A12) contains robust results for the Protestant-only sample for the faithful coefficient, and for both the Protestant- and Catholic-only samples for the faith school coefficient. Again, in the whole sample the results are not robust, though in the faith school coefficient case they get close with a little over 66% as much unobserved selection needed as observed selection needed to nullify the result – high considering the rich range of covariates contained in the analysis.

5.2.2 Bounded Estimates (Oster βs)

The above leads to a suggestion that the most robust results are to be found among the GCSE outcomes, A-level results, and Christian belief at age 25 – a list that is virtually unchanged to the summary given at the end of Section 5.1. The omission is university attendance which did not meet the Oster threshold. Moreover, the most common source of robustness is the Protestant-only sample.

The Protestant-only sample will therefore provide the basis for employing the second capability of the Oster test – to generate bounding estimates. Without clearing the Oster threshold value of 1, the bounds begin to dip below zero – so the Protestant-only sample provides the best case in which to show bounded estimates across a range of outcomes. These are given in Table 5. The table uses the maximum R^2 as 1.3 times the observed R^2 in each case and assumes a δ value of 1 – equal unobserved and observed selection.

Panel A, column (1) shows the outcomes for which there were significant and robust faith school coefficients in analysis above; Panel B gives the same for the faithful coefficients. Column (2) gives the lower bound on those estimates. Standard errors for the lower bounds are

calculated using bootstrapping. Only GCSE point score and Christian at age 25 have faithful coefficients that remain significant in this case – though this is assuming a high degree of unobserved selection that may be unrealistic. Given this, these two outcomes can be said to display the most robust results. No faith school coefficients meet this standard. Column (3) tests whether the two estimates, the original and the lower bound, can be said to be statistically different. In no case do the resulting test statistics get close to suggesting that the lower bound is significantly different from the original OLS estimate.

Table 5 Bounded Estimates (Protestant-Only Sample)

Outcomes	(1) Original Coefficient	(2) Lower Bounds (Oster βs)	(3) Test (1) – (2)	N
Panel A – Faith School (Protestants Only)				
Five A*-C	0.052*	0.035	0.445	3,727
	(0.027)	(0.027)		
Five A*-C (incl.	0.052*	0.036	0.435	3.727
English and Maths)	(0.028)	(0.026)		
Christian at Age 25	0.075*	0.037	0.610	2,088
	(0.044)	(0.044)		
Panel B – Faithful (Protestants Only)				
No. Good Passes	0.362***	0.209	0.766	3,727
	(0.133)	(0.149)		
GCSE Points	0.121***	0.090***	0.634	3,727
	(0.032)	(0.035)		
A-level Points	0.092*	0.082	0.170	2,274
	(0.051)	(0.029)		
Christian Age 25	0.111***	0.087***	0.560	2,088
	(0.028)	(0.030)		

Notes: Each specification is estimated using the column (7) controls from Table 3. Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The maximum R^2 in the Oster specification is $R^2_{max} = 1.3\widehat{R^2}$. The assumed δ value is 1 – equal unobserved selection to observed selection. Column (3) tests the different between the original coefficient and the lower bound using: $\frac{\beta_{ols}-\beta_{lowerbound}}{\sqrt{SE(\beta_{ols})^2+SE(\beta_{lowerbound})^2}}$, which is based on Clogg, Petkova, and Haritou (1995), and gives the resulting test statistic.

5.3 IPWRA

A complementary approach to the analysis conducted above is to use IPWRA to model both the treatment (i.e. faith school attendance or faithfulness) and the outcome. As outlined in the empirical strategy, overlap and coefficient balance are important issues. Overlap charts and coefficient balances are given in the Appendix (Figures B3 and B3 and Tables A15 and A16).

The outcome variables analysed through IPWRA are those summarised at the end of section 5.1 – GCSE, A-level, university attendance, and future Christian belief. Only the significant results from the OLS are analysed for each of the two treatments, though for each outcome that is analysed coefficients are given for the whole, Protestant-only, and Catholic-only samples, even if they were not significant in all samples initially. Table 6 shows the results for the faith school treatment on two different versions of the five A*-C grades benchmark and for future Christian belief. The coefficient magnitudes are generally smaller than the OLS estimates, and only two of the treatments remain robust – Christian belief at age 25 in the whole sample and five A*-C grades including English and maths for Protestants. Taken together, the IPWRA results do not provide a strong endorsement for there being faith schooling impacts.

In contrast, Table 7 shows broad agreement between what has been presented in previous sections and the IPWRA analysis of the faithful treatment. Large effects are found for GCSE point score and number of good passes, but not for the five A*-C benchmark (with the exception of Catholics). A-level points show no impacts of faithfulness, and the impacts for university attendance are also weak. Christian belief at age 25 does show significant result.

Table 6 IPWRA Average Treatment Effect (ATE) Estimates for Faith School

	Whole	Protestant	Catholic
5 A*-C Grades			
Faith School	0.025	0.045	0.031
	(0.021)	(0.028)	(0.029)
Non-Faith School Mean Outcome	0.621***	0.653***	0.659***
	(0.006)	(0.008)	(0.027)
N	7,802	3,727	945
5 A*-C Grades (inc. English and Maths)			
Faith School	0.021	0.058**	0.037
	(0.021)	(0.029)	(0.029)
Non-Faith School Mean Outcome	0.501***	0.530***	0.544***
	(0.006)	(0.008)	(0.026)
N	7, 802	3,727	945
Christian at Age 25			
Faith School	0.168***	0.072	0.065
	(0.038)	(0.049)	(0.047)
Non-Faith School Mean Outcome	0.368***	0.476***	0.681***
	(0.008)	(0.012)	(0.040)
N	4,195	2,088	502

Note: Each specification is estimated using the column (6) controls from Table 3, excluding religious affiliation for the Protestant and Catholic columns. Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 7 IPWRA Average Treatment Effect (ATE) Estimates for Faithful

	Whole	Protestant	Catholic
Five A*-C Grades (inc. English and Maths)			
Faithful	0.024	0.013	0.121**
	(0.015)	(0.022)	(0.055)
Unfaithful Mean Outcome	0.477***	0.523***	0.471***
	(0.011)	(0.021)	(0.055)
N	7,802	3,727	945
Number of Good Passes			
Faithful	0.412***	0.301**	1.040**
	(0.116)	(0.151)	(0.410)
Unfaithful Mean Outcome	5.917***	6.316***	5.782***
	(0.083)	(0.15)	(0.413)
N	7,802	3,727	945
GCSE Point Score			
Faithful	0.125***	0.122***	0.121
	(0.027)	(0.037)	(0.09)
Unfaithful Mean Outcome	-0.021	0.041	0.082
	(0.019)	(0.037)	(0.09)
N	7,802	3,727	945
A-Level Point Score (Top 3)			
Faithful	0.05	0.066	.
	(0.047)	(0.068)	.
Unfaithful Mean Outcome	-0.057	-0.039	.
	(0.035)	(0.066)	.
N	4,462	2,274	.
Attend University			
Faithful	0.035*	0.024	0.027
	(0.018)	(0.025)	(0.07)
Unfaithful Mean Outcome	0.445***	0.488***	0.531***
	(0.013)	(0.024)	(0.069)
N	6,547	3,177	800
Christian at Age 25			
Faithful	0.151***	0.205***	.
	(0.026)	(0.031)	.
Unfaithful Mean Outcome	0.296***	0.300***	.
	(0.017)	(0.029)	.
N	4,195	2,088	.

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Each specification is estimated using the column (6) controls from Table 3, excluding religious affiliation in the case of the Protestant and Catholic columns. Blank spaces show where IPWRA specifications would not converge.

5.4 Mediation Analysis

The set of results presented above is suggestive of a relationship between faithfulness (or intrinsic religiosity) and short-term educational outcomes as well as with longer term belief. Weaker evidence exists of a relationship between faithfulness and medium- to long-term educational outcomes. The fact that the effect seems weaker for outcomes further into the future is not surprising. Indeed, the fact that there are any significant coefficients for university attendance for example, suggests that further analysis of religiosity and university attendance would be fruitful.

For those significant results that have been found, it is important to establish whether they act through a channel of their own (in essence a direct faithfulness effect) or if they act through some other variable. Here we conduct mediation analysis along the lines of Acharya et al (2016) (outlined in the empirical strategy). The most obvious potential mediators are non-cognitive skills: work ethic, internal locus of control, self-esteem, and sociability. We use the first principal component of these four variables as a potential mediator.

Table 8 shows the result for the Protestant sample for each variable where results were found previously.[22] The first two columns show mediation of the faith school treatment, and the remaining columns show mediation of the faithful treatment. Sample sizes differ from previous tables as Table 8 only includes those who are present in wave two (whilst treatment has occurs in wave one), when the non-cognitive skill questions are asked. As a result, effect sizes can also be different to previous tables, which is evident in the faith school effect for Christian religion at age 25 is not significant (though only marginally so), and the A-Level point score and University attendance variables are not either. The analysis points squarely to the non-cognitive skills doing little to dampen the effect of faithfulness. Each of the coefficients provided hardly moves. This is also true when the second principal component is included (not shown).

As a further piece of analysis (not shown), the number of good GCSE passes at age 16 was used as a mediator for attainment at A-level at 18, and the A-level point score was used as a mediator for university attendance. In the case of the former little mediation effect was found, whilst in the latter the A-level point score was enough to reduce the faithfulness effect on university attendance to near zero. This indicates that this is the channel through which any possible effects on university attendance are working.

[22] Though not presented, similar results exist for the whole sample and Catholic sample.

Table 8 Mediation Analysis

Outcome	5 A*-C	Christian at age 25	5 A*-C	Good passes	GCSE points	A-level points	Attends university	Christian at age 25
Treatment to be mediated:	Faith School		Faithful					
Panel A – Initial Coefficient								
Initial	0.052*	0.081	0.018	0.373**	0.118***	0.081	0.036	0.141***
	(0.031)	(0.053)	(0.020)	(0.162)	(0.040)	(0.056)	(0.024)	(0.033)
N	2717	1535	2717	2717	2717	1745	2337	1535
R^2	0.455	0.087	0.455	0.539	0.520	0.321	0.281	0.087
Panel B – Mediation Analysis								
Mediated	0.050*	0.080	0.013	0.318**	0.105***	0.074	0.030	0.140***
	(0.028)	(0.050)	(0.020)	(0.157)	(0.038)	(0.058)	(0.025)	(0.034)
Principal component	0.044***	0.005	0.04[4]***	0.566***	0.136***	0.070***	0.047**	0.007**
	(0.009)	(0.016)	(0.009)	(0.077)	(0.017)	(0.026)	(0.011)	(0.016)
N	2717	1535	2717	2717	2717	1745	2337	1535
R^2	0.456	0.087	0.456	0.543	0.524	0.323	0.282	0.087
Oster Tests of Panel B								
$R^2_{max} = 1$	0.752	0.056	0.190	0.769	1.271	0.991	0.119	0.144
$R^2_{max} = 2\widehat{R^2}$	0.897	0.579	0.227	0.769	1.271	2.075	0.304	1.381
R^2_{max}=$1.3\widehat{R^2}$	2.978	1.842	0.755	2.147	3.816	6.884	1.010	3.718

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Each column represents a different outcome from those where significant results were previously identified. Panel A provides the initial coefficients from earlier analysis whilst Panel B gives the mediated coefficient and the coefficient on the principle component. For the initial coefficient cluster robust standard errors are reported, for the mediation analysis standard errors result from bootstrapping. More details of the mediation procedure we follow can be found in Acharya, Blackwell, and Sen (2016)

5.5 Heterogeneity Analysis

Throughout, the coefficients for Protestant and Catholics were reported separately in all specifications. A reasonable question would be whether any interaction effects between faith and faith schooling exist. Perhaps the faithful benefit from faith schools but nobody else does? Such an explanation might be used to argue that the evidence so far presented has missed some effects of faith schooling. Table 9 reports the previous column (6) specification for every possible outcome, even those not fully discussed before if they did not previously present significant results. A one sentence summary of the table is that the interaction effect does not seem to be important. The only meaningful exception being A-level points where the interaction effect is large. A question might be also raised as to whether significant heterogeneity exists by gender. This is shown in Table A17 in the Appendix. Again, a single sentence summary would say that there were few meaningful differences by gender. Panel C

of the table shows statistical tests of the coefficients – in only three cases are they different. This includes the faithfulness effect on A-level points – significant for females but not males. This justifies not stratifying previous analyses by gender, but instead by religious affiliation.

Finally, we have explored heterogeneity by family income (indicated by receipt of Free School Meals – which is confined to around 15% of pupils) and by prior ability (measured by quartiles of the KS2 test score distribution from the final year of elementary schooling (year 6)). Our prior is that the ethos of faith schools would be consistent with a redistribution of school effort so as to compensate for sources of prior disadvantage. That is, we feel that we should expect bigger effects for such groups. We do find some heterogeneity by ability and income group for the all-important five A*-C benchmark outcome, but not in the expected direction.[23]

5.6 Why Choose a Faith School?

Of the evidence presented so far it would appear that there is some reason to believe that one's own innate religiosity is important in shaping some short-term educational outcomes - as well as future belief. It would also appear that faith schooling plays a lesser role, with the evidence of any effects for educational outcomes existing only for the five A*-C Benchmark (Table 3). This benchmark, while an important educational indicator, the lack of robustness of this effect casts doubt on the finding. Moreover, the fact that such an effect does not exist for other outcomes raises questions as to why people choose faith schools and whether those choices constitute mistakes. An obvious reason to choose a faith school is one's faith. Faith schools provide parents with an opportunity to have their children schooled in an environment that has an explicit religious association – one that they themselves may value and hence promote. In the results presented above, this appears to leverage a greater chance that the child will be faithful in later life. But other reasons may be important, too. Tables 10 and 11 give a range of outcomes that suggest the ethos and environment of faith schools are preferable to their secular counterparts. The outcomes in Table 10 are: How satisfied parents are with discipline in their child's school; How satisfied they are with the progress their child is making; and, How well they feel their child fits in and gets on with others at the school.[24] In Table 11 they are: Whether

[23] The faith school coefficient was significant only for the highest quartile of prior (KS2) ability. Moreover, the minority in receipt of free school meals do not benefit from faith schools whilst the majority who are not do. Faithfulness effects are present across quartiles of prior ability for the sample as a whole, and for those **not** on free school meals. That is, the effects appear to be among the more able and the less disadvantaged. Results are available on request.

[24] Related work by Green et al. (2014) shows that there are negative effects for child well-being associated with private Catholic schools in Spain. English faith schools appear to do better.

the child has ever been bullied at the school; Ever been sent abusive texts; and, Experienced physical violence. The definition of ever bullied includes abusive texts and physical violence, but also covers name calling, social exclusion, and extortion (made to hand over money).[25]

Table 10 shows: positive and significant results associated with satisfaction with discipline for all samples; positive and significant coefficients on satisfaction with progress; and positive and significant effects for getting on with others for the whole sample and for the Protestant-only sample. Oster tests are broadly supportive with robustness to the proportional selection standard seen in several columns for the second and third rows of Panel B. Similarly, the coefficients in Table 11 are all negative, although not all statistically significant. For the whole sample and for Protestants the chance of being sent abusing texts or experiencing physical violence is lower in faith schools. For Catholics only the latter is significant. These tables lead to some positive conclusions for faith schools.

An interesting feature of the Next Steps dataset is the questions that ask parents why they actually chose their child's school. Figure 5 shows the main reasons for those who attend secular (or non-faith) schools and those who attend faith schools, conditional on the school they are currently at being their first-choice school. Parents can give multiple reasons for choosing their first-choice school, so the charts below are generated by regressing each answer on a dummy for faith school attendance and the number of reasons the parent gave for their choice. Parents of faith school pupils tend to give around 0.2 choices more than those in secular schools, a difference which, whilst small, is statistically significant.

As expected, those in faith schools are much more likely to choose their school because of its religion than those attending secular schools – in line with the finding that faith school attendance "buys" higher religiosity in later life in OLS Table A12. But the biggest reason for choosing a faith school, other than that, is examination results. This is significantly higher than the number of non-faith attendees who cited the same reason, followed closely by friends or siblings attending, which is lower than for those attending secular schools. Bullying features, but to check the scale on the axis of the bullying panel and compare it to the exam results panel is to notice a substantial difference. Reputation and the school being local are the two remaining reasons, the former being significantly less common among faith attendees than secular school attendees.

[25] Gorman et al. (2020) study the impacts of bullying on educational, labour market, and health outcomes using Next Steps.

Table 9 **Including Interaction Effect in Column (6) OLS Specifications from Previous Tables (Whole Sample)**

Outcomes:	(1) Five A*-C	(2) Five A*-C including Eng+Math	(3) Number of Good Passes	(4) GCSE Points	(5) A-level Points	(6) Attends University	(7) Attends Russell Group	(8) Degree Class	(9) Wage Rate	(10) Christian at Age 25
Faith School	0.020	-0.014	0.045	0.071	-0.098	-0.011	-0.017	0.035	-0.036	0.177***
	(0.040)	(0.039)	(0.322)	(0.102)	(0.118)	(0.037)	(0.063)	(0.142)	(0.028)	(0.054)
Faithful	0.027**	0.022**	0.380***	0.123***	0.031	0.041***	0.005	-0.041	0.002	0.242***
	(0.012)	(0.011)	(0.091)	(0.021)	(0.034)	(0.014)	(0.018)	(0.038)	(0.015)	(0.019)
Faith School * Faithful	0.011	0.063	0.043	-0.062	0.254**	0.040	0.047	-0.021	0.020	-0.017
	(0.038)	(0.040)	(0.321)	(0.096)	(0.122)	(0.045)	(0.068)	(0.153)	(0.053)	(0.057)
N	7,803	7,803	7,803	7,803	4,463	6,548	2,872	961	3,609	4,196
R^2	0.449	0.462	0.562	0.543	0.307	0.303	0.174	0.117	0.011	0.204
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Belief	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Each specification is estimated using the column (6) controls from Table 3. The sample includes those of no religion, Protestants and Catholics. EM in column (2) means including English and Maths. Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 10 Parental Satisfaction with the Ethos of the Child's School

Outcomes:	(1) Discipline	(2) Progress	(3) Gets-on	(4) Discipline	(5) Progress	(6) Gets-on	(7) Discipline	(8) Progress	(9) Gets-on
Panel A Faith School	**Whole sample**			**Protestants**			**Catholics**		
Faith School	0.254***	0.057*	0.044*	0.261***	0.091**	0.077**	0.274***	-0.038	0.015
	(0.048)	(0.031)	(0.026)	(0.061)	(0.038)	(0.037)	(0.085)	(0.060)	(0.052)
Protestant	-0.048	-0.041	-0.000						
	(0.039)	(0.032)	(0.026)						
Catholic	-0.063	-0.073*	0.017						
	(0.052)	(0.041)	(0.034)						
N	7,687	7,784	7,803	3,682	3,717	3,727	937	944	945
R^2	0.076	0.092	0.045	0.079	0.089	0.060	0.115	0.151	0.114
Panel B Oster Tests of Faith School Coefficient									
$R^2_{max} = 1$	0.067	0.061	0.094	0.148	0.168	-0.054	0.174	1.300$^+$	-0.032
$R^2_{max} = 2\widehat{R^2}$	0.765	0.600	1.970$^+$	1.608$^+$	1.691$^+$	-0.301	1.207$^+$	19.873$^+$	-0.245
$R^2_{max} = 1.3\widehat{R^2}$	2.176$^+$	1.975$^+$	6.442	4.554$^+$	5.458$^+$	-0.996	3.155$^+$	63.477$^+$	-0.817
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Belief	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Each specification is estimated using the column (6), (7) and (8) controls from Table 3 for the whole, protestant-only, and catholic-only samples respectively. Discipline takes value 1 if the cohort member's parents are satisfied with the disciplinary policies of the schools, 0 otherwise; Progress takes value one if the cohort member's parents are satisfied with how much progress their child is making at school, 0 otherwise; and Gets-on takes value 1 if the cohort member's parents are satisfied with how the child gets on with others at the school, 0 otherwise. Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. $^+$ indicates passing of robustness threshold in the Oster (2017) test.

Table 11 Experiences of Bullying at School

Outcomes:	(1) Any Bullying	(2) Texts	(3) Violence	(4) Any Bullying	(5) Texts	(6) Violence	(7) Any Bullying	(8) Texts	(9) Violence
Panel A Faith School	**Whole Sample**			**Protestants**			**Catholics**		
Faith School	-0.001	-0.047**	-0.048***	0.008	-0.054*	-0.041*	-0.008	-0.046	-0.062*
	(0.013)	(0.021)	(0.017)	(0.015)	(0.031)	(0.023)	(0.025)	(0.040)	(0.035)
Protestant	0.026***	0.037*	-0.003						
	(0.010)	(0.021)	(0.016)						
Catholic	0.014	0.010	0.003						
	(0.015)	(0.026)	(0.022)						
N	7,803	7,803	7,803	3,727	3,727	3,727	945	945	945
R^2	0.016	0.059	0.049	0.025	0.066	0.060	0.065	0.136	0.118
Panel B Oster Tests of Faith School Coefficient									
$R^2_{max} = 1$	0.001	0.051	0.080	-0.022	0.262	0.335	-0.039	2.220	0.743
$R^2_{max} = 2\widehat{R^2}$	0.076	0.799	1.519$^+$	-0.870	3.679$^+$	5.172$^+$	-0.570	14.008$^+$	5.411$^+$
$R^2_{max} = 1.3\widehat{R^2}$	0.254	2.594$^+$	4.816$^+$	-2.880	11.890$^+$	16.719$^+$	-1.894	45.650$^+$	16.821$^+$
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Belief	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Each specification is estimated using the column (6), (7) and (8) controls from Table 3 for the whole, protestant-only, and catholic-only samples respectively. Any Bullying takes value 1 if the individual reports facing any kind of bullying at schools, and takes value 0 otherwise; Texts takes value 1 if the individual has been sent abusive text messages, 0 otherwise; and Violence takes value one if the individual has been threatened with or actually experience physical violence, 0 otherwise. Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. $^+$ indicates passing of robustness threshold in the Oster (2017) test.

These results on parental motivations suggest a possible contradiction. Although parents report being satisfied with the ethos of their faith school (based on Tables 10 and 11 above), after having made their choice of school, this does not seem to be the reason that most of them give for having chosen that school in the first place. Bullying and reputation, those responses that are closest to those from Tables 10 and 11, turn out to be much less frequently mentioned by parents than religion and examination results. This suggests that there are benefits associated with the ethos of faith schools that parents do value - but that they are not necessarily driving school choice. It could be that although schools may be able to credibly communicate their academic quality, they cannot easily do the same for their ethos and that parents need to experience this.

Figure 5 Parent's Reasons for Choosing their Child's School (95% CI)

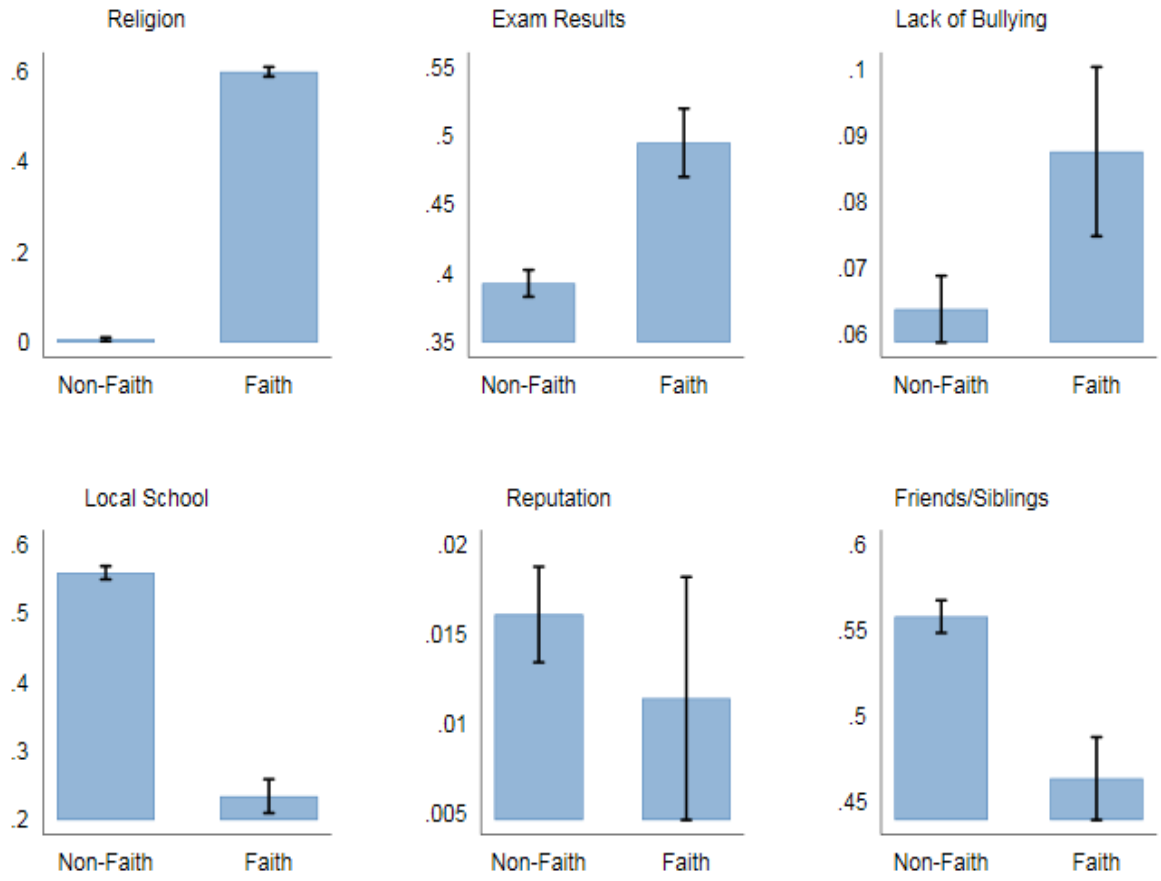


Figure Note: Faith means a faith school and Non-Faith a secular school. Parents could make more than one choice. The six most popular reasons are given. Each is a binary outcome that takes value 1 if it was mentioned and 0 otherwise. Due to the difference in popularity of each option each panel has a different axis range.

6 Discussion and Conclusion

This paper presents robust evidence of a relationship between one's innate religiosity (faithfulness) and a range of outcomes. The strongest results are found for GCSE point scores and Christian belief at age 25, but less robust associations exist for A-level point score and university attendance too. This university association, if true, is interesting, as the lack of an effect on one's degree class suggests that higher religiosity may help an individual get to a university, but not increase how they perform once there. Across the results presented, the most robust associations are found among Protestants. In addition, its important to note that our results **only** apply to Christian schools - sample sizes prohibit us from analysing non-Christian schools and it would be entirely inappropriate to assume that our results are general ones.

The combination of a rich range of covariates from a unique English dataset supported by the Oster (2019) tests as well as results from inverse probability weighting regression adjustment, and analysis of non-cognitive skills as potential mediators, all point to relationships beyond simple correlations. Our measure of religiosity also captures intensity of belief better than measures of practise. The degree to which a causal relationship can be pinned down in the absence of quasi-experimental methods is, of course, difficult to argue, but given the focus of the literature around faith schools and the emphasis on the fact that no convincing instrument has yet been found, it is arguable that this evidence represents a strong attempt to unearth effects that could be relied on for policy purposes.

Possible explanations for the mechanism by which the faithfulness effect is operating are in line with Hungerman (2014a) - that religion prescribes certain behaviours that the faithful enact, and with McCullough and Willoughby (2009), that faith provides a coping mechanism for stress. The former explanation should be accounted for in the mediation analysis where work ethic is included, though it is conceivable that an individual following a prescribed behaviour to work hard does not see themselves as hard-working as opposed to doing their duty. The latter should be accounted for in the mediation analysis too, in the variables relating to locus of control and self-esteem, but for similar reasons the Next Steps variables may not capture them perfectly. As it stands, the evidence presented above is consistent with a "purely" religiosity effect for which mediators have not yet been identified.

Despite the attention it has received in the literature, we find that faith schooling, in contrast to faithfulness, does not seem to improve many outcomes; although there is evidence of an association between faith school attendance and the five A*-C benchmark (including English

and Maths) and with future religious belief. The benchmark is an important one in terms of English educational attainment since it is used to judge the success of both students and schools alike. Moreover, it is an important milestone that facilitates a transition to the academic post-compulsory track. Faith schooling also seems to impact future religious belief, a finding that is in tune with Wadsworth and Walker (2017).

Fewer outcomes are positively impacted by faith schooling in our initial specifications than are impacted by faithfulness. This result reflects the literature, where a number of papers have found mixed effects of faith schools. In particular, research that examines faith school effects in England generally do not observe significant impacts. For example, Gibbons and Silva (2011) find very small effects of faith primary schools that are generally not robust to exposure to the AET method – a finding that is reflected in our work for faith secondary schools. We expand on their analysis by explicitly examining the role of intrinsic religiosity.

Where faith schools do appear to perform better in the context of how satisfied parents are with their ethos and environment. Parents of those in faith schools appear to be more satisfied with the progress their child is making, how their child gets on with others, and the discipline of the school. Fewer faith-school-attending children also report having been bullied than their secular equivalents. This seems to be something parents are pleased with in retrospect, as bullying and reputation do not seem to be driving parents to choose the schools in the first place.

The policy environment around faith schools has recently taken on renewed importance. In the US, a recent Supreme Court ruling states that private school choice voucher programmes cannot exclude religious schools (SCOTUSblog, 2020). In such a setting, evidence relating to the impacts of such schools is crucial – and our English evidence on the limited impacts of faith schools seems applicable, given that the funding policy in England is equivalent to a voucher that covers the whole cost of faith schooling.[26] Moreover, there may be a case for thinking that, if the only long-run effect of faith schooling is to perpetuate faithfulness, then the Supreme Court judgment may run foul of the US First Amendment Establishment Clause that, broadly speaking, makes it illegal for the US government to promote religion, for example with taxes. In the UK, our findings cast doubt on those of Sullivan et al. (2018) and on the soundness of the current policy in England to facilitate the expansion of faith schools, alongside expanding academically selective grammar schools, whereas other state schools have been

[26] Evidence relating to voucher programmes more broadly can be found on the Urban Institute Website – here: https://www.urban.org/research/publication/do-voucher-students-attain-higher-levels-education

constrained from doing so (*BBC,* 2018). There is a presumption that schools rated as outstanding can expand without compromising performance. Moreover, there is also an assumption that there is something about how faith schools operate that makes them good candidates for expansion. Our work suggests that it is the pupils in those schools and that simply expanding faith schools may not have the intended effect, if the result were that a higher proportion of the non-faithful then have access.

There are several of avenues for further work suggested by our analysis that our data has the potential to illuminate. Firstly, the association between faithfulness and a range of outcomes suggests that it could be the type of person who attends faith schools that make such schools perform better. The peer effects literature has been focussed almost entirely on the ability of peers rather than any other characteristics. Our data clearly shows that there are many nonfaithful pupils in faith schools and vice versa – sufficient to consider the possibility that the proportion of faithful peers might affect the outcomes, both academic and other, of the non-faithful pupils as well as the faithful. Allen and Vignoles (2016) find little effect of the proximity of a faith school on the performance of the non-faith schools, but this is consistent with, at best, very indirect evidence of (an absence of) faithful-peer effects. In addition, the analysis here uses faith at age 14 as the treatment and does not consider the evolution of religiosity with age – in schools where faithfulness declines slowly we would expect to find more faithful peers remaining at the age when high stakes tests occur and it may be at this age, in particular, that positive externalities on the never-faithful peers might be most apparent. To the extent that faith is correlated with parental socio-economic background this might have important effects on social mobility – although it's not at all clear, right now, whether the existence of state-funded faith schooling promotes social mobility or not.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining Causal Findings without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110 (3): 512–29.

Adamczyk, Amy. 2009. "Understanding the Effects of Personal and School Religiosity on the Decision to Abort a Premarital Pregnancy." *Journal of Health and Social Behavior* 50 (2): 180–95.

Allen, Rebecca, and Anna Vignoles. 2016. "Can School Competition Improve Standards? The Case of Faith Schools in England." *Empirical Economics* 50 (3): 959–73.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005a. "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling." *Journal of Human Resources* 40 (4): 791–821.

———. 2005b. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–84.

Andrews, Jon, and Rebecca Johnes. 2016. "Faith Schools, Pupil Performance and Social Selection." *Education Policy Institute*, 1–46.

BBC. 2018. "Grammar Schools and Faith Schools Get Green Light to Expand," May 2018.

Becker, Sascha O, and Ludger Woessmann. 2009. "Was Weber Wrong? A Human Capital Theory of Protestant Economic History." *Quarterly Journal of Economics* 124 (2): 531–96.

Bjorklund-Young, Alanna. 2016. "What Are Non-Cognitive Skills? Why Are They Important? How Can Schools Develop Students' Non-Cognitive Skills ?", Institute for Education Policy Working Paper, Johns Hopkins University.

Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Capital* 43 (4): 972–1059.

Campante, Filipe, and David Yanagizawa-Drott. 2015. "Does Religion Affect Economic Growth and Happiness? Evidence from Ramadan." *Quarterly Journal of Economics* 130 (2): 615–58.

Carattini, Juliana F., Angela K. Dills, Sean E. Mulholland, and Rachel B. Sederberg. 2012. "Catholic Schools, Competition, and Public School Quality." *Economics Letters* 117 (1): 334–36.

Cardak, Buly A., and Joe Vecci. 2013. "Catholic School Effectiveness in Australia: A Reassessment Using Selection on Observed and Unobserved Variables." *Economics of Education Review* 37: 34–45.

Centre for Longitudinal Studies. 2018. "Next Steps Sweep 8 - Age 25 Survey: User Guide." Edited by Lisa Calderwood, Institute of Education, UCL.

Cesur, Resul, and Naci Mocan. 2018. "Education, Religion, and Voter Preference in a Muslim Country." *Journal of Population Economics* 31 (1): 1–44.

Chingos, Matthew M., and Paul E. Peterson. 2015. “Experimentally Estimated Impacts of School Vouchers on College Enrollment and Degree Attainment.” *Journal of Public Economics* 122: 1–12.

Clemens, Michael A, Claudio E. Montenegro, and Lant Pritchett. 2019. “The Place Premium: Bounding the Price Equivalent of Migration Barriers.” *The Review of Economics and Statistics* 101 (2): 201–13.

Cohen-Zada, Danny, and William Sander. 2008. “Religion, Religiosity and Private School Choice: Implications for Estimating the Effectiveness of Private Schools.” *Journal of Urban Economics* 64 (1): 85–100.

Department for Children, Schools and Families. 2007. “Faith in the System.” Whitehall, London.

Elder, Todd, and Christopher Jepsen. 2014. “Are Catholic Primary Schools More Effective than Public Primary Schools?” *Journal of Urban Economics* 80: 28–38.

Evans, William N., and Robert M. Schwab. 1995. “Finishing High School and Starting College: Do Catholic Schools Make a Difference?” *The Quarterly Journal of Economics* 110 (4): 941–74.

Farrington, Camille A, Melissa Roderick, Elaine M Allensworth, Jenny Nagaoka, Tasha Seneca Keyes, David W Johnson, and Nicole O Beechum. 2012. “Teaching Adolescents To Become Learners: The Role Of Noncognitive Factors in Shaping School Performance: A Critical Literature Review.” *University of Chicago Consortium on Chicago School Research*.

Fruehwirth, Jane Cooley, Sriya Iyer, and Anwen Zhang. 2019. “Religion and Depression in Adolescence Sriya Iyer Anwen Zhang.” *Journal of Political Economy* 127 (3): 1178–1209.

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 and Organization* 137: 191–213.

Goldberg, Lewis R. 1971. “A Historical Survey of Personality Scales and Inventories.” In *Advances in Psycholigical Assessment*, edited by Paul McReynolds, 293–336. Palo Alto, CA: Science and Behaviour Books.

Gorman, Emma, Colm Harmon, Silvia Mendolia, Anita Staneva, and Ian Walker. 2020. “Adolescent School Bullying Victimisation and Later Life Outcomes.” 20–05.

Green, Colin P., María Navarro-Paniagua, Domingo P. Ximénez-de-Embún, and María Jesús Mancebón. 2014. “School Choice and Student Wellbeing.” *Economics of Education Review* 38: 139–50.

Gruber, Jonathan H. 2005. “Religious Market Structure, Religious Participation, and Outcomes: Is Religion Good for You?” *The BE Journal of Economic Analysis & Policy* 5 (1): Article 5.

Guttman, Louis. 1954. “Some Necessary Conditions for Common-Factor Analysis.”

Psychometrika 19 (2): 149–61.

Hoxby, Caroline M. 1994. "Do Private Schools Provide Competition for Public Schools?" Working Paper No.4978. National Bureau of Economic Research.

Huber, Martin. 2019. "A Review of Causal Mediation Analysis for Assessing Direct and Indirect Treatment Effects." 500.

Hungerman, Daniel M. 2014a. "Do Religious Proscriptions Matter? Evidence from a Theory-Based Test." *Journal of Human Resources* 49 (4): 1053–93.

———. 2014b. "The Effect of Education on Religion: Evidence from Compulsory Schooling Laws." *Journal of Economic Behavior & Organization* 104: 52–63.

Imbens, Guido W, and Jeffrey M Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.

Jackson, Donald A. 1993. "Stopping Rules in Principal Components Analysis : A Comparison of Heuristical and Statistical Approaches." *Ecology* 74 (8): 2204–14.

Kim, Young Joo. 2011. "Catholic Schools or School Quality? The Effects of Catholic Schools on Labor Market Outcomes." *Economics of Education Review* 30 (3): 546–58.

Lehrer, Evelyn. 2004. "Religiosity as a Determinant of Educational Attainment: The Case of Conservative Protestant Women in the United States." *Review of Economics of the Household* 2 (2): 203–19.

Long, Robert, Shadi Danechi, and Philip Loft. 2019. "Faith Schools in England: FAQs (Briefing Paper Number 06972)."

McAndrew, Siobhan, and David Voas. 2011. "Measuring Religiosity Using Surveys." *SURVEY QUESTION BANK: Topic Overview*. Vol. 4.

McCullough, Michael E, and Brian L B Willoughby. 2009. "Religion, Self-Regulation, and Self-Control: Associations, Explanations, and Implications." *Psychological Bulletin* 135 (1): 69.

McFarland, Michael J., Bradley R.E. Wright, and David L. Weakliem. 2011. "Educational Attainment and Religiosity: Exploring Variations by Religious Tradition." *Sociology of Religion: A Quarterly Review* 72 (2): 166–88.

Mendolia, Silvia, Alfredo Paloyo, and Ian Walker. 2019. "Intrinsic Religiosity, Personality Traits, and Adolescent Risky Behaviors." *B.E. Journal of Economic Analysis and Policy* 19 (3): 1–16.

Neal, Derek. 1997. "The Effects of Catholic Secondary Schooling on Educational Achievement." *Journal of Labor Economics* 15 (1, Part 1): 98–123.

New Schools Network. 2015. "Comparison of Different Types of School A Guide to Schools in England."

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.

Oosterbeek, Hessel, and Bas van der Klaauw. 2013. “Ramadan, Fasting and Educational Outcomes.” *Economics of Education Review* 34: 219–26.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business and Economic Statistics* 37 (2): 187–204.

Sander, William. 2001. “The Effects of Catholic Schools on Religiosity, Education, and Competition.” 32.

Schaltegger, Christoph A, and Benno Torgler. 2010. “Work Ethic, Protestantism, and Human Capital.” *Economics Letters* 107 (2): 99–101.

SCOTUSblog. 2020. “Espinoza v. Montana Department of Revenue.” 2020.

Spenkuch, Jörg L. 2017. “Religion and Work: Micro Evidence from Contemporary Germany.” *Journal of Economic Behavior & Organization* 135: 193–214.

Spenkuch, Jörg L., and Philipp Tillmann. 2018. “Elite Influence? Religion and the Electoral Success of the Nazis.” *American Journal of Political Science* 62 (1): 19–36.

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): 806–22.

Torgler, Benno. 2006. “The Importance of Faith: Tax Morale and Religiosity.” *Journal of Economic Behavior and Organization* 61 (1): 81–109.

University College London, UCL Institute of Education, and Centre for Longitudinal Studies. 2020. “Next Steps: Sweeps 1-8, 2004-2016: Secure Access [Data Collection].” UK Data Service SN:7104.

Wadsworth, Allyssa A., and Jay K. Walker. 2017. “Religiosity and the Impact of Religious Secondary Schooling.” *Journal of School Choice* 11 (1): 131–47.

Walker, Ian, and Matthew Weldon. 2020. “School Choice, Admission, and Comparing the Relative Access to Good Schools in England.” Lancaster University Management School, Working Paper 2020/011.

Weber, Max. 2001. “The Protestant Ethic and the Spirit of Capitalism.” Translated by Talcott Parsons [1930], London, UK: Routledge Classics.

West, Martin, and Ludger Woessmann. 2010. “‘Every Catholic Child in a Catholic School’: Historical Resistance to State Schooling, Contemporary Private Competition and Student Achievement across Countries.” *Economic Journal* 120 (546): F229–55.

Appendix A - Tables

Table A1 - Measures of Religiosity in Next Steps

Variable Name	**Survey Question** (Response)	**Type of Religiosity**
Extensive Margin of Belief		
Religion	**What if any is your religion?** None, Christian, Buddhist, Hindu, Jewish, Muslim, Sikh, Another Religion	Affiliation
Christian Denomination	**Which of these would you say it is?** Roman Catholic, Church of England, Methodist, URC/Presbyterian/Congregationalist, Baptist, Other	Affiliation
Intensive Margin of Belief		
Faithfulness*	**How important would you say your religion is to the way you live your life – is it...** Very important, Fairly important, Not very important, Not at all important?	Belief
Religious Classes	**In the last 12 months have you ever gone to classes or courses connected with any religious establishment you might go to?** Yes, No	Practise
Frequency of Attendance at Religious Classes	**How often would you say you have gone to classes like this in the last 12 months – would you say that, on average, you have gone...** More than once a week, Once a week, Two or three times a month, About once a month, Less than once a month?	Practise

* faithfulness is the primary measure of religiosity used in this paper.

Table A2 – Summary Statistics - Parental and Household

	Faithfulness			School Type		
	Unfaithful	Faithful	Total	Non-Faith	Faith	Total
			Mean/SD/N			
IMD Score	23.59	23.63	23.61	22.96	27.18	23.53
	(17.06)	(16.87)	(16.97)	(16.84)	(17.46)	(16.98)
	5829	5564	11393	9781	1531	11312
Mother Has a Degree	0.09	0.11	0.10	0.10	0.14	0.10
	(0.29)	(0.32)	(0.30)	(0.30)	(0.34)	(0.30)
	5283	5094	10377	8885	1421	10306
Mother in Employment	0.26	0.24	0.25	0.25	0.24	0.25
	(0.44)	(0.43)	(0.43)	(0.43)	(0.43)	(0.43)
	5837	5568	11405	9792	1531	11323
Mother's Ethnicity	0.10	0.24	0.17	0.15	0.29	0.17
	(0.30)	(0.43)	(0.38)	(0.36)	(0.45)	(0.38)
	5837	5568	11405	9792	1531	11323
Young Parent	0.14	0.09	0.12	0.12	0.08	0.12
	(0.35)	(0.29)	(0.32)	(0.33)	(0.27)	(0.32)
	5837	5568	11405	9792	1531	11323
Mother is Faithful	0.44	0.89	0.66	0.63	0.87	0.66
	(0.50)	(0.32)	(0.47)	(0.48)	(0.33)	(0.47)
	5837	5568	11405	9792	1531	11323
Single Mother	0.28	0.25	0.27	0.26	0.28	0.27
	(0.45)	(0.43)	(0.44)	(0.44)	(0.45)	(0.44)
	5537	5313	10850	9310	1469	10779
Number of Dependent Children	2.22	2.21	2.22	2.22	2.23	2.22
	(1.05)	(1.02)	(1.04)	(1.04)	(1.00)	(1.04)
	5790	5522	11312	9713	1517	11230
Experienced Any Bullying	0.91	0.91	0.91	0.91	0.91	0.91
	(0.28)	(0.29)	(0.29)	(0.29)	(0.29)	(0.29)
	5837	5568	11405	9792	1531	11323
Been Called Names or Sent Abusive Messages	0.34	0.30	0.32	0.33	0.26	0.32
	(0.47)	(0.46)	(0.47)	(0.47)	(0.44)	(0.47)
	5837	5568	11405	9792	1531	11323
Threatened with or Experienced Violence	0.20	0.16	0.18	0.19	0.15	0.18
	(0.40)	(0.37)	(0.39)	(0.39)	(0.35)	(0.39)
	5837	5568	11405	9792	1531	11323
Parents Satisfied with Discipline	3.06	3.22	3.13	3.10	3.39	3.14
	(0.91)	(0.85)	(0.88)	(0.89)	(0.77)	(0.88)
	5627	5425	11052	9486	1495	10981
Parent's Satisfied with Discipline	3.25	3.39	3.32	3.31	3.39	3.32
	(0.76)	(0.68)	(0.72)	(0.73)	(0.69)	(0.72)
	5741	5496	11237	9648	1509	11157
Parents Satisfied with how Child Gets on With Others	3.58	3.61	3.60	3.59	3.63	3.60
	(0.65)	(0.63)	(0.64)	(0.64)	(0.62)	(0.64)
	5768	5514	11282	9688	1514	11202
Max N	5837	5568	11405	9792	1531	11323

Table A3 – Summary Statistics - School

	Faithfulness			School Type		
	Unfaithful	Faithful	Total	Non-Faith	Faith	Total
			Mean/SD/N			
Protestant School	0.02	0.06	0.04	0.00	0.28	0.04
	(0.13)	(0.23)	(0.19)	(0.00)	(0.45)	(0.19)
	5715	5440	11155	9792	1437	11229
Catholic School	0.02	0.16	0.09	0.00	0.72	0.09
	(0.15)	(0.37)	(0.29)	(0.00)	(0.45)	(0.29)
	5715	5440	11155	9792	1437	11229
Percent on FSM	15.07	16.01	15.53	15.58	15.44	15.56
	(12.47)	(14.56)	(13.54)	(13.53)	(13.63)	(13.55)
	5733	5516	11249	9792	1531	11323
Grammar School	0.03	0.05	0.04	0.04	0.01	0.04
	(0.18)	(0.21)	(0.19)	(0.20)	(0.07)	(0.19)
	5766	5524	11290	9792	1531	11323
Has Sixth Form	0.57	0.60	0.58	0.58	0.63	0.58
	(0.50)	(0.49)	(0.49)	(0.49)	(0.48)	(0.49)
	5733	5516	11249	9792	1531	11323
KS3 School Size	1125.13	1103.67	1114.63	1130.66	1007.64	1114.03
	(352.26)	(349.81)	(351.21)	(352.24)	(329.24)	(351.73)
	5766	5524	11290	9792	1531	11323
KS2 School Size	327.07	322.54	324.89	327.59	305.75	324.64
	(136.45)	(140.90)	(138.62)	(139.30)	(133.63)	(138.74)
	5583	5182	10765	9252	1444	10696
Percentage with Special Educational Needs	2.48	2.29	2.39	2.41	2.21	2.39
	(1.62)	(1.48)	(1.56)	(1.59)	(1.34)	(1.56)
	5704	5480	11184	9680	1531	11211
Percent White	86.28	79.68	83.04	83.89	77.56	83.04
	(17.28)	(23.75)	(20.97)	(20.34)	(24.05)	(20.99)
	5733	5516	11249	9792	1531	11323
Percent who do not have English as First Language	6.47	10.94	8.66	8.41	10.35	8.67
	(11.55)	(16.96)	(14.63)	(14.61)	(14.79)	(14.65)
	5733	5516	11249	9792	1531	11323
School's Average KS3 Score	33.94	34.49	34.21	34.07	35.18	34.22
	(3.40)	(3.53)	(3.47)	(3.57)	(2.59)	(3.47)
	5745	5496	11241	9743	1531	11274
% Achieving Five A*-C Grades (inc. Eng and Maths)	39.21	43.61	41.37	40.13	48.81	41.32
	(19.41)	(20.87)	(20.26)	(20.33)	(18.03)	(20.25)
	5666	5456	11122	9665	1531	11196
Single Sex Intake	0.08	0.12	0.10	0.09	0.14	0.10
	(0.27)	(0.33)	(0.30)	(0.29)	(0.34)	(0.30)
	5766	5524	11290	9792	1531	11323
Max N	5766	5524	11290	9792	1531	11323

Table A4 – Next Steps' Survey Questions for Bullying and Non-Cognitive Skills

Variable	Question (Response) [P denotes question asked to parents]
Bullying [P]	**As far as you know, have any of these things happened to (insert name) at his/her school since we last spoke to you last year?** Called names by other pupils, sent offensive or hurtful text messages or emails, shut out from groups of other pupils or from joining in things, made to give other pupils his or her money or belongings, threatened by their pupils with being hit or kicked or with other violence, actually being hit or kicked or attacked in any other way by other pupils, experience any type of racist behaviour by other pupils, any other sort of bullying, no none of these things have happened in the last 12 months.
Work Ethic	**How strongly do you agree with the following statement: "I work as hard as I can."** Strongly disagree, Disagree, Agree, Strongly Agree.
Locus of Control	**How strongly do you agree with the following statement: "I decide what happens in my life."** Strongly Disagree, Disagree, Agree, Strongly Agree.
Self-Esteem	**How much have you been losing confidence in yourself?** Much more than usual, About the same as usual, Less than usual, Much less than usual. **How much have you been thinking of yourself as worthless?** Much more than usual, About the same as usual, Less than usual, Much less than usual.
Sociability	**How many times have you had friends round in the last week?** None, Once or twice, 3-5 times, 6 times or more. **How often have you gone out with friends in the last week?** None, Once or twice, 3-5 times, 6 times or more. **How many times per week on average do you attend clubs or societies?** Never, Less than once a week, Once or twice a week, Less than once a week, 5 times a week or more.

Table A5 OLS Regression Results for Five A*-C Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**Attained Five A*-C Grades at GCSE**							
	Whole Sample						**P only**	**C only**
Panel A								
Faith School	0.102***	0.075**	0.035	0.043**	0.032*	0.042**	0.053*	0.031
	(0.027)	(0.030)	(0.021)	(0.019)	(0.019)	(0.019)	(0.028)	(0.033)
N	9,005	9,005	8,660	7,931	7,905	7,849	3,761	957
R^2	0.004	0.011	0.427	0.455	0.456	0.462	0.457	0.533
Panel B								
Faithful	0.106***	0.057***	0.049***	0.037**	0.029*	0.030*	0.024	0.045
	(0.013)	(0.021)	(0.016)	(0.016)	(0.017)	(0.016)	(0.018)	(0.047)
N	9,064	9,064	8,713	7,980	7,956	7,803	3,727	945
R^2	0.012	0.013	0.424	0.450	0.451	0.448	0.430	0.493
Panel C								
Faith school	0.064**	0.075***	0.034	0.038*	0.036*	0.036*	0.052*	-0.009
	(0.026)	(0.029)	(0.021)	(0.019)	(0.020)	(0.021)	(0.027)	(0.035)
Faithful	0.091***	0.049**	0.046***	0.035**	0.029*	0.028*	0.024	0.048
	(0.013)	(0.021)	(0.016)	(0.016)	(0.017)	(0.016)	(0.018)	(0.048)
N	8,945	8,945	8,604	7,883	7,859	7,803	3,727	945
R^2	0.013	0.014	0.424	0.450	0.450	0.449	0.431	0.493
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$			0.230	0.262	0.235	0.231	0.606	-0.080
$R^2_{max} = 2\widehat{R^2}$		0.854	0.313	0.321	0.288	0.283	0.801	-0.082
$R^2_{max} = 1.3\widehat{R^2}$		2.220+	1.041+	1.067+	0.957	0.942	2.657+	-0.273
Faithful								
$R^2_{max} = 1$			0.131	0.106	0.079	0.076	0.322	0.644
$R^2_{max} = 2\widehat{R^2}$		0.169	0.178	0.129	0.097	0.093	0.426	0.663
$R^2_{max} = 1.3\widehat{R^2}$		0.498	0.591	0.431	0.323	0.309	1.418+	2.200+
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A6 OLS Regression Results for GCSE Point Score (Standardised)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**GCSE Point Score (Standardised)**							
				Whole Sample			**P only**	**C only**
Panel A								
Faith School	0.203***	0.147**	0.049	0.052	0.038	0.054	0.075	0.056
	(0.059)	(0.067)	(0.052)	(0.052)	(0.053)	(0.055)	(0.063)	(0.079)
N	9,005	9,005	8,660	7,931	7,905	7,849	3,761	957
R^2	0.004	0.017	0.519	0.552	0.552	0.541	0.537	0.528
Panel B								
Faithful	0.286***	0.196***	0.170***	0.148***	0.138***	0.131***	0.122***	0.194**
	(0.026)	(0.041)	(0.029)	(0.029)	(0.030)	(0.029)	(0.032)	(0.082)
N	9,064	9,064	8,713	7,980	7,956	7,803	3,727	945
R^2	0.020	0.021	0.521	0.555	0.555	0.543	0.539	0.525
Panel C								
Faith school	0.095	0.124*	0.028	0.033	0.025	0.043	0.068	0.019
	(0.059)	(0.066)	(0.052)	(0.052)	(0.053)	(0.054)	(0.062)	(0.081)
Faithful	0.261***	0.176***	0.162***	0.141***	0.132***	0.129***	0.121***	0.189**
	(0.026)	(0.041)	(0.029)	(0.029)	(0.030)	(0.029)	(0.032)	(0.084)
N	8,945	8,945	8,604	7,883	7,859	7,803	3,727	945
R^2	0.020	0.021	0.521	0.554	0.554	0.543	0.540	0.525
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		0.016	0.121	0.153	0.110	0.199	0.613	0.429
$R^2_{max} = 2\widehat{R^2}$		0.696	0.121	0.153	0.110	0.199	0.613	0.429
$R^2_{max} = 1.3\widehat{R^2}$		2.047+	0.371	0.411	0.296	0.559	1.740+	1.296+
Faithful								
$R^2_{max} = 1$		0.005	0.267	0.257	0.220	0.197	1.208+	1.111+
$R^2_{max} = 2\widehat{R^2}$		0.218	0.267	0.257	0.220	0.197	1.208+	1.111+
$R^2_{max} = 1.3\widehat{R^2}$		0.578	0.811	0.684	0.587	0.549	3.405+	3.309+
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A7 OLS Regression Results for A-level (UCAS) Points Score (Top 3, Standardised)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**A-level (UCAS) Point Score (Standardised)**							
	Whole Sample						**P only**	**C only**
Panel A								
Faith School	0.104*	0.030	0.063	0.038	0.016	0.071	0.111	0.013
	(0.055)	(0.060)	(0.047)	(0.044)	(0.045)	(0.049)	(0.067)	(0.094)
N	5,865	5,865	5,597	4,500	4,485	4,485	2,289	618
R^2	0.001	0.003	0.241	0.289	0.291	0.307	0.305	0.417
Panel B								
Faithful	0.100***	0.118**	0.144***	0.120***	0.100**	0.112**	0.092*	0.143
	(0.030)	(0.046)	(0.044)	(0.046)	(0.048)	(0.048)	(0.050)	(0.156)
N	5,877	5,877	5,604	4,504	4,491	4,463	2,274	611
R^2	0.002	0.004	0.241	0.289	0.291	0.307	0.305	0.417
Panel C								
Faith school	0.071	0.018	0.048	0.025	0.011	0.067	0.106	0.008
	(0.056)	(0.059)	(0.047)	(0.045)	(0.046)	(0.049)	(0.068)	(0.096)
Faithful	0.086***	0.116**	0.139***	0.117**	0.099**	0.109**	0.092*	0.141
	(0.030)	(0.047)	(0.045)	(0.047)	(0.048)	(0.048)	(0.051)	(0.157)
N	5,838	5,838	5,570	4,476	4,463	4,463	2,274	611
R^2	0.003	0.004	0.242	0.289	0.291	0.307	0.306	0.417
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		0.001	0.110	0.078	0.029	0.258	1.553	-0.032
$R^2_{max} = 2\widehat{R^2}$		0.132	0.345	0.192	0.071	0.582	3.516	-0.045
$R^2_{max} = 1.3\widehat{R^2}$		0.432	1.147+	0.638	0.238	1.936+	11.647+	-0.149
Faithful								
$R^2_{max} = 1$		0.002	0.245	0.223	0.163	0.200	0.859	2.235+
$R^2_{max} = 2\widehat{R^2}$		0.343	0.766	0.547	0.397	0.451	1.946+	3.123+
$R^2_{max} = 1.3\widehat{R^2}$		0.774	2.517+	1.810+	1.315+	1.495+	6.437+	10.343+
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A8 OLS Regression Results for University Attendance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**University Attendance**							
	Whole Sample						**P only**	**C only**
Panel A								
Faith School	0.110***	0.078***	0.044**	0.038*	0.016	0.033	0.041	0.062
	(0.027)	(0.028)	(0.021)	(0.021)	(0.020)	(0.020)	(0.032)	(0.040)
N	8,098	8,098	7,722	6,656	6,633	6,589	3,208	810
R^2	0.005	0.017	0.242	0.288	0.293	0.302	0.289	0.367
Panel B								
Faithful	0.140***	0.111***	0.086***	0.069***	0.045**	0.041**	0.030	0.067
	(0.013)	(0.019)	(0.017)	(0.018)	(0.019)	(0.020)	(0.020)	(0.061)
N	8,132	8,132	7,751	6,680	6,659	6,548	3,177	800
R^2	0.020	0.020	0.244	0.291	0.295	0.303	0.289	0.371
Panel C								
Faith school	0.055**	0.064**	0.032	0.028	0.011	0.029	0.040	0.053
	(0.026)	(0.028)	(0.021)	(0.020)	(0.020)	(0.020)	(0.032)	(0.040)
Faithful	0.128***	0.101***	0.080***	0.064***	0.042**	0.040**	0.029	0.053
	(0.012)	(0.019)	(0.017)	(0.018)	(0.019)	(0.020)	(0.020)	(0.061)
N	8,039	8,039	7,667	6,613	6,592	6,548	3,177	800
R^2	0.021	0.021	0.244	0.290	0.294	0.303	0.290	0.372
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		0.015	0.086	0.089	0.031	0.090	0.182	0.507
$R^2_{max} = 2\widehat{R^2}$		0.637	0.267	0.217	0.074	0.207	0.445	0.854
$R^2_{max} = 1.3\widehat{R^2}$		1.862+	0.888	0.722	0.245	0.688	1.477+	2.821+
Faithful								
$R^2_{max} = 1$		0.006	0.080	0.082	0.047	0.045	0.124	0.253
$R^2_{max} = 2\widehat{R^2}$		0.233	0.248	0.200	0.113	0.104	0.304	0.427
$R^2_{max} = 1.3\widehat{R^2}$		0.585	0.809	0.658	0.376	0.345	1.009+	1.415+
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A9 OLS Regression Results for Russell Group University Attendance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**Russell Group University Attendance**							
				Whole Sample			**P only**	**C only**
Panel A								
Faith School	-0.004	-0.030	-0.027	-0.036	-0.032	0.011	-0.009	0.030
	(0.028)	(0.028)	(0.027)	(0.026)	(0.027)	(0.027)	(0.034)	(0.057)
N	3,384	3,384	3,239	2,893	2,886	2,883	1,512	422
R^2	0.000	0.002	0.109	0.152	0.152	0.174	0.173	0.291
Panel B								
Faithful	-0.013	-0.019	-0.015	-0.021	-0.015	-0.010	-0.008	-0.062
	(0.016)	(0.026)	(0.025)	(0.026)	(0.026)	(0.026)	(0.028)	(0.081)
N	3,392	3,392	3,246	2,900	2,893	2,872	1,505	418
R^2	0.000	0.002	0.109	0.153	0.153	0.175	0.174	0.297
Panel C								
Faith school	0.001	-0.029	-0.028	-0.038	-0.036	0.008	-0.009	0.027
	(0.028)	(0.028)	(0.028)	(0.026)	(0.027)	(0.027)	(0.034)	(0.058)
Faithful	-0.013	-0.015	-0.011	-0.016	-0.012	-0.011	-0.008	-0.067
	(0.016)	(0.026)	(0.025)	(0.026)	(0.026)	(0.026)	(0.028)	(0.083)
N	3,372	3,372	3,227	2,882	2,875	2,872	1,505	418
R^2	0.000	0.002	0.110	0.153	0.153	0.175	0.174	0.298
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		-0.002	-0.155	-0.317	-0.319	-0.065	0.038	-0.102
$R^2_{max} = 2\widehat{R^2}$		-0.787	-1.253	-1.751	-1.757	-0.309	0.181	-0.240
$R^2_{max} = 1.3\widehat{R^2}$		-1.803	-4.140	-5.774	-5.802	-1.030	0.605	-0.798
Faithful								
$R^2_{max} = 1$		0.001	0.049	0.153	0.080	0.088	0.134	-0.533
$R^2_{max} = 2\widehat{R^2}$		0.542	0.399	0.847	0.444	0.416	0.636	-1.254
$R^2_{max} = 1.3\widehat{R^2}$		1.603$^+$	1.327$^+$	2.820$^+$	1.477$^+$	1.385$^+$	2.120$^+$	-4.135
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A10 OLS Regression Results for Degree Classification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**Degree Classification**							
				Whole Sample			**P only**	**C only**
Panel A								
Faith	-0.026	0.009	0.017	0.029	0.028	0.033	0.068	.
School	(0.040)	(0.044)	(0.047)	(0.048)	(0.049)	(0.053)	(0.074)	.
N	1,202	1,202	1,156	963	962	962	519	.
R^2	0.000	0.002	0.046	0.089	0.095	0.117	0.142	.
Panel B								
Faithful	-0.014	-0.032	-0.020	-0.049	-0.052	-0.054	-0.100*	.
	(0.026)	(0.042)	(0.041)	(0.045)	(0.049)	(0.050)	(0.056)	.
N	1,204	1,204	1,158	965	964	961	519	.
R^2	0.000	0.003	0.046	0.088	0.094	0.117	0.146	.
Panel C								
Faith	-0.023	0.014	0.021	0.038	0.033	0.038	0.068	.
school	(0.041)	(0.044)	(0.047)	(0.048)	(0.049)	(0.053)	(0.073)	.
Faithful	-0.009	-0.034	-0.023	-0.054	-0.056	-0.056	-0.100*	.
	(0.026)	(0.042)	(0.041)	(0.045)	(0.049)	(0.050)	(0.056)	.
N	1,200	1,200	1,154	962	961	961	519	.
R^2	0.001	0.003	0.046	0.090	0.097	0.118	0.148	.
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		-0.001	-0.016	-0.041	-0.041	-0.051	-0.178	.
$R^2_{max} = 2\widehat{R^2}$		-0.251	-0.334	-0.412	-0.381	-0.382	-1.020	.
$R^2_{max} = 1.3\widehat{R^2}$		-0.783	-1.103	-1.354	-1.256	-1.262	-3.322	.
Faithful								
$R^2_{max} = 1$		0.022	-0.277	-0.196	-0.250	-0.312	-1.105	.
$R^2_{max} = 2\widehat{R^2}$		6.839	-5.668	-1.968	-2.310	-2.317	-6.190	.
$R^2_{max} = 1.3\widehat{R^2}$		16.959	-18.698	-6.371	-7.501	-7.554	-19.020	.
Religion	No	Yes	Yes	Yes	Yes	Yes	No	.
Individual	No	No	Yes	Yes	Yes	Yes	Yes	.
Parental	No	No	No	Yes	Yes	Yes	Yes	.
Parental Belief	No	No	No	No	Yes	Yes	Yes	.
School	No	No	No	No	No	Yes	Yes	.

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A11 – OLS Regression Results for Wage Rate at Age 25

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**Wage Rate at Age 25**							
	Whole Sample						**P only**	**C only**
Panel A								
Faith School	-0.034*	-0.018*	-0.025*	-0.017	-0.022	-0.008	-0.038	-0.016
	(0.018)	(0.010)	(0.013)	(0.016)	(0.021)	(0.027)	(0.045)	(0.019)
N	4,509	4,509	4,329	3,649	3,639	3,632	1,827	438
R^2	0.000	0.000	0.004	0.006	0.007	0.011	0.038	0.115
Panel B								
Faithful	-0.024	0.025**	0.019*	0.014	0.011	0.014	0.014	-0.034*
	(0.031)	(0.012)	(0.010)	(0.011)	(0.010)	(0.011)	(0.012)	(0.018)
N	4,511	4,511	4,331	3,649	3,639	3,609	1,810	432
R^2	0.000	0.000	0.003	0.006	0.007	0.011	0.038	0.121
Panel C								
Faith school	-0.026**	-0.020*	-0.028*	-0.020	-0.024	-0.010	-0.039	-0.016
	(0.013)	(0.011)	(0.014)	(0.017)	(0.021)	(0.028)	(0.045)	(0.020)
Faithful	-0.020	0.027**	0.022**	0.017	0.014	0.015	0.014	-0.030
	(0.031)	(0.012)	(0.011)	(0.012)	(0.010)	(0.011)	(0.012)	(0.021)
N	4,477	4,477	4,299	3,626	3,616	3,609	1,810	432
R^2	0.000	0.000	0.004	0.006	0.007	0.011	0.038	0.123
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		0.000	0.004	0.004	0.006	0.002	-0.083	-1.893
$R^2_{max} = 2\widehat{R^2}$		0.587	1.222+	0.594	0.810	0.211	-2.079	-13.189
$R^2_{max} = 1.3\widehat{R^2}$		1.793+	3.998+	1.966+	2.675+	0.703	-6.899	-41.715
Faithful								
$R^2_{max} = 1$		0.000	-0.001	-0.001	-0.001	-0.002	0.024	-0.513
$R^2_{max} = 2\widehat{R^2}$		-0.200	-0.225	-0.212	-0.168	-0.180	0.603	-3.544
$R^2_{max} = 1.3\widehat{R^2}$		-0.575	-0.741	-0.703	-0.557	-0.598	2.009+	-10.955
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A12 – OLS Regression Results for Christian Belief at Age 25

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome:	**Christian Belief at Age 25**							
	Whole Sample						**P only**	**C only**
Panel A								
Faith School	0.359***	0.174***	0.173***	0.161***	0.122***	0.110***	0.086**	0.120*
	(0.024)	(0.028)	(0.029)	(0.031)	(0.030)	(0.031)	(0.043)	(0.061)
N	5,333	5,333	5,090	4,262	4,249	4,223	2,108	509
R^2	0.060	0.175	0.184	0.191	0.205	0.213	0.078	0.213
Panel B								
Faithful	0.361***	0.171***	0.160***	0.167***	0.126***	0.123***	0.112***	0.156*
	(0.015)	(0.023)	(0.024)	(0.026)	(0.026)	(0.026)	(0.028)	(0.087)
N	5,340	5,340	5,098	4,264	4,252	4,196	2,088	502
R^2	0.141	0.176	0.184	0.192	0.205	0.215	0.082	0.205
Panel C								
Faith school	0.228***	0.156***	0.156***	0.142***	0.113***	0.101***	0.075*	0.108*
	(0.024)	(0.028)	(0.029)	(0.031)	(0.031)	(0.031)	(0.044)	(0.063)
Faithful	0.320***	0.154***	0.142***	0.150***	0.116***	0.118***	0.111***	0.132
	(0.014)	(0.023)	(0.024)	(0.026)	(0.026)	(0.026)	(0.028)	(0.087)
N	5,294	5,294	5,054	4,234	4,222	4,196	2,088	502
R^2	0.163	0.184	0.192	0.199	0.209	0.218	0.083	0.211
Panel D: Oster Tests of Panel C								
Faith School								
$R^2_{max} = 1$		0.094	0.097	0.093	0.072	0.057	0.052	0.198
$R^2_{max} = 2\widehat{R^2}$		0.402	0.397	0.364	0.268	0.205	0.564	0.721
$R^2_{max} = 1.3\widehat{R^2}$		1.216+	1.206+	1.130+	0.854	0.662	1.806+	2.240+
Faithful								
$R^2_{max} = 1$		0.035	0.034	0.038	0.028	0.030	0.110	0.441
$R^2_{max} = 2\widehat{R^2}$		0.149	0.139	0.148	0.106	0.106	1.131+	1.617+
$R^2_{max} = 1.3\widehat{R^2}$		0.452	0.428	0.452	0.337	0.338	3.235+	5.096+
Religion	No	Yes	Yes	Yes	Yes	Yes	No	No
Individual	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Parental	No	No	No	Yes	Yes	Yes	Yes	Yes
Parental Belief	No	No	No	No	Yes	Yes	Yes	Yes
School	No	No	No	No	No	Yes	Yes	Yes

Notes: Controls are listed at the beginning of Section 5.1 (page 22). Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.10. + indicates passing of the robustness threshold in the Oster (2017) test.

Table A13 - Coefficients on Religion

Variables	(1) Five A*-C	(2) Five A*-C (EM)	(3) Good Passes	(4) GCSE Points	(5) A-level Points	(6) Attends University	(7) Russell	(8) Degree Class	(9) Wage Rate	(10) Christian (Age 25)
Panel A - Without Faithful										
Protestant	0.021*	0.019*	0.278***	0.086***	-0.016	0.037***	0.015	-0.016	-0.003	0.250***
	(0.011)	(0.011)	(0.089)	(0.022)	(0.036)	(0.014)	(0.021)	(0.039)	(0.015)	(0.018)
Catholic	0.007	0.013	0.071	0.046	0.096*	0.020	0.046	-0.060	-0.023	0.387***
	(0.020)	(0.019)	(0.155)	(0.041)	(0.058)	(0.023)	(0.031)	(0.067)	(0.021)	(0.033)
N	7,849	7,849	7,849	7,849	4,485	6,589	2,883	962	3,632	4,223
R-squared	0.449	0.462	0.563	0.541	0.307	0.302	0.174	0.117	0.011	0.213
Panel B - Including Faithful										
Protestant	0.001	-0.002	-0.023	-0.008	-0.100*	0.009	0.024	0.026	-0.014	0.163***
	(0.016)	(0.016)	(0.121)	(0.031)	(0.052)	(0.020)	(0.030)	(0.052)	(0.014)	(0.026)
Catholic	-0.015	-0.007	-0.242	-0.053	0.008	-0.009	0.056	-0.018	-0.036	0.294***
	(0.024)	(0.023)	(0.179)	(0.046)	(0.071)	(0.028)	(0.039)	(0.076)	(0.024)	(0.040)
Faithful	0.028*	0.029*	0.419***	0.129***	0.109**	0.040**	-0.011	-0.056	0.015	0.118***
	(0.016)	(0.016)	(0.122)	(0.029)	(0.049)	(0.020)	(0.026)	(0.050)	(0.011)	(0.026)
N	7,803	7,803	7,803	7,803	4,463	6,548	2,872	961	3,609	4,196
R-squared	0.449	0.462	0.563	0.543	0.307	0.303	0.175	0.118	0.011	0.218
Panel C - Differences between Panels A and B – P-values of tests of coefficient differences										
Protestant	0.084	0.095	0.001	0.000	0.033	0.065	0.503	0.125	0.305	0.000
Catholic	0.054	0.295	0.023	0.002	0.826	0.230	0.393	0.105	0.439	0.013
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Belief	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; the Panel C tests were conducted using the suest command in Stata. Each column uses the full range of controls from column (6) of Table 3.

Table A14 - OLS Specifications with Religious Classes as Independent Variable

Variables	(1) Five A*-C	(2) Five A*-C (EM)	(3) Good Passes	(4) GCSE Point	(5) A-level Point	(6) Attends University	(7) Russell	(8) Degree Class	(9) Wage Rate	(10) Christian (Age 25)
Classes	0.024	0.025	0.186	0.027	0.040	0.035*	-0.038	-0.013	-0.030	0.086***
	(0.015)	(0.017)	(0.125)	(0.031)	(0.046)	(0.020)	(0.026)	(0.053)	(0.023)	(0.030)
Faith School	0.026	0.040**	0.064	0.021	0.143***	0.032	0.024	0.007	-0.048	0.163***
	(0.019)	(0.018)	(0.172)	(0.048)	(0.053)	(0.021)	(0.026)	(0.056)	(0.043)	(0.033)
Observations	4,713	4,713	4,713	4,713	2,906	4,015	1,934	641	2,263	2,614
R-squared	0.436	0.465	0.557	0.527	0.308	0.293	0.176	0.136	0.029	0.098
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Belief	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Each column uses the full range of controls from column (6) of Table 3.

Table A15 – Coefficient Balances for Faith School (Whole Sample)

		Raw	Weighted
	Total	7,802	7,802.00
	Treated	1,059	3,892.30
	Control	6,743	3,909.70
	Standardized Differences		**Variance Ratio**
Variables	**Raw**	**Weighted**	**Raw**	**Weighted**
Gender	0.076	-0.007	0.999	1
FSM	-0.005	-0.002	0.989	0.996
KS2 Achievement	0.112	-0.003	0.848	0.928
IMD	0.265	0.032	1.095	0.841
Mother's Education	0.127	-0.005	1.161	1.048
Mother's Age	0.137	-0.018	0.978	1.039
Number of Dependent Children in HH	-0.021	-0.009	0.958	0.931
Region	-0.186	-0.031	0.985	0.963
Single Mother	0.058	-0.004	1.07	0.995
Young Parent	-0.127	0.026	0.696	1.066
Month of Birth	0.035	-0.001	1.046	1.037

Table A16 – Coefficient Balance for Faithful (Whole Sample)

		Raw	Weighted
	Total	7,802	7,802.00
	Treated	1,059	3,892.30
	Control	6,743	3,909.70
	Standardized Differences		**Variance Ratio**
Variables	**Raw**	**Weighted**	**Raw**	**Weighted**
Gender	0.000	1.003	1.000	0.098
FSM	0.053	0.003	0.872	1.007
KS2 Achievement	0.109	0.001	0.921	0.968
IMD	0.034	0.001	0.931	0.943
Mother's Education	0.122	0.001	1.011	0.916
Mother's Age	0.234	0.004	0.923	0.935
Number of Dependent Children in HH	0.025	0.001	0.962	0.962
Region	-0.009	0.001	0.996	0.997
Single Mother	0.081	0.001	0.906	1.002
Young Parent	0.187	0.002	0.610	1.006
Month of Birth	0.003	0.000	0.975	0.982

Table A17 Analysis by Gender

Variables	(1) Five A*-C	(2) Five A*-C (EM)	(3) Good Passes	(4) GCSE Points	(5) A-level Points	(6) Attends University	(7) Russell	(8) Degree Class	(9) Wage Rate	(10) Christian (Age 25)
Panel A - Males										
Faith School	0.023	0.032	-0.019	0.045	0.081	-0.022	0.057	0.198**	-0.124	0.164***
	(0.020)	(0.020)	(0.156)	(0.037)	(0.063)	(0.025)	(0.038)	(0.081)	(0.166)	(0.035)
Faithful	0.024*	0.029**	0.405***	0.121***	0.021	0.028	0.066**	-0.009	-0.026	0.240***
	(0.014)	(0.014)	(0.109)	(0.026)	(0.046)	(0.018)	(0.028)	(0.057)	(0.117)	(0.024)
N	3,971	3,971	3,971	3,971	2,102	3,253	1,285	407	1,664	1,875
R^2	0.452	0.454	0.564	0.554	0.303	0.305	0.208	0.223	0.036	0.219
Panel B - Females										
Faith School	0.039**	0.050**	0.143	0.005	0.005	0.047*	0.028	-0.063	-0.024	0.165***
	(0.019)	(0.020)	(0.149)	(0.035)	(0.035)	(0.025)	(0.030)	(0.060)	(0.020)	(0.030)
Faithful	0.030**	0.019	0.354***	0.118***	0.118***	0.055***	-0.044*	-0.078	0.007	0.261***
	(0.014)	(0.015)	(0.112)	(0.026)	(0.026)	(0.019)	(0.025)	(0.051)	(0.015)	(0.023)
N	3,832	3,832	3,832	3,832	3,832	3,295	1,587	554	1,945	2,321
R^2	0.429	0.467	0.549	0.513	0.513	0.294	0.182	0.124	0.047	0.238
Panel C - Differences by Gender – P-values of tests of coefficient differences										
Faith School	0.571	0.517	0.435	0.418	0.281	0.047	0.530	0.004	0.514	0.993
Faithful	0.786	0.630	0.745	0.929	0.064	0.297	0.003	0.333	0.305	0.545
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parental Belief	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Note: Each column includes the column (6) controls from Table 3. The Panel C tests were conducted using the suest command in Stata.

Appendix B – Figures

Figure B1 - Screeplot for Non-cognitive Skills Mediators

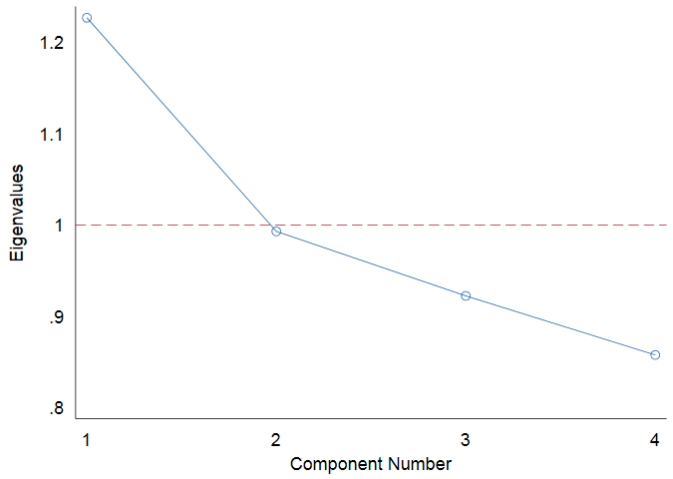


Figure B2 – Overlap for Faith School (Whole Sample)

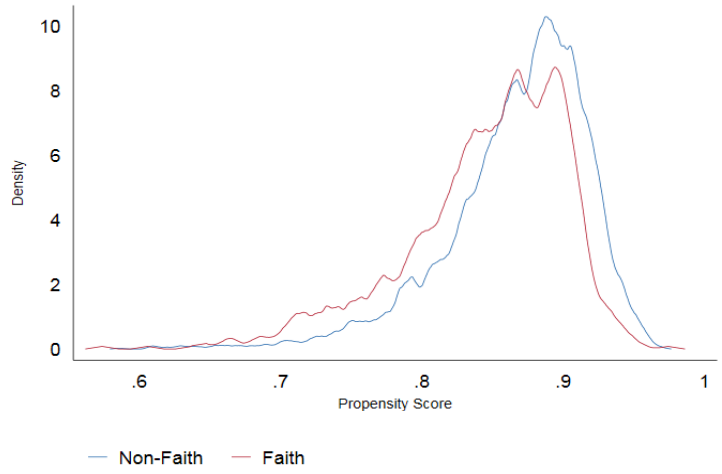


Figure B3 – Overlap for Faithful (Whole Sample)