Essays in the Economics of Education and Identity

Andrew McKendrick

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in the Department of Economics

LANCASTER UNIVERSITY

2021

i

Own Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgement, the work presented is entirely my own.

Supervisors' Declaration

To the Faculty of Lancaster University:

The supervisors appointed to guide and examine the thesis of Andrew McKendrick find it satisfactory and recommend that it be accepted.

Dr Maria Navarro Paniagua and Professor Ian Walker

Acknowledgements

Essays in the Economics of Education and Identity

Abstract

This thesis covers three research topics, each separate chapter a distinct entity designed to be submitted to (and hopefully published in) a good quality journal. Both chapters 1 and 2 make use of the Next Steps cohort study with linked administrative data. The first chapter examines the impact of faithfulness, a form of intrinsic religiosity, on a range of educational and other outcomes. Once the Oster (2019) sensitivity test is applied, faithfulness is robustly and positively associated with educational attainment at GCSE and with Christian affiliation at age 25 but with no other outcomes. Faith schooling does not have any robust impacts other than a positive association with future Christian affiliation. The popular perception that faith schools are important for outcomes, therefore, appears to be misplaced – faith matters more. The second chapter examines the impact of the Education Maintenance Allowance (EMA) on a range of outcomes covering educational attainment, risky behaviours, and employment. Positive impacts are found on retention in further education and on university attendance; negative effects are found on the chance of being on an insecure employment contract at age 25. These estimates are causal under assumptions of unconfoundedness as they are estimated using Inverse Probability Weighting Regression Adjustment (IPWRA). Treatment heterogeneity is analysed using a machine learning approach called Causal Forests, with some interesting dimensions of heterogeneity identified. Chapter 3 uses the secure access version of the Labour Force Survey to analyse the effect of education on national identity. Education is instrumented using the Raising of the School Leaving Age (RoSLA) reform. Although OLS results suggest positive and significant effects on national identity, IV estimates are mixed, with signs turning negative and significance disappearing on one of the two outcomes, despite the instrument being very strong. Each of these topics has meaningful policy implications.

Table of Contents

Table of Tables

*

Table of Figures

Thesis Introduction

The formative years of our lives are spent in school. The myriad details of the world's education systems are poured over by economists and others, eagerly identifying the effects of government policy or quirks in the rules that lead to disparities between individuals. Two such changes appear in chapters 2 and 3 of this thesis – the Education Maintenance Allowance (EMA) and the Raising of the School Leaving Age (RoSLA). But schools do not exist in an isolated state, separate from society, certain types of pupils attend them. The way those students think, or, more relevantly for this thesis, what those students believe, could shape how they interact with their school and the wider education system. Ultimately, this may impact how highly they achieve. By the same token, in exposing their students to new ideas and bringing them together with people from different backgrounds, schools could shape an individual's identity. This thesis covers these three facets – the effect of individual identity on educational attainment, and other outcomes (chapter 1); the effect of an education policy on outcomes (chapter 2); and the impact of education on individual identity (chapter 3).

In the first chapter, co-authored with Ian Walker, we estimate the effects of "faithfulness", as measured by responses to the question of the importance of one's religion for how one lives one's life, on both short-run educational outcomes and later outcomes. We rely on a cohort study of 15 years olds, containing an extensive range of pupil, household, and school-level characteristics, linked to administrative data. Because of the well-known difficulty of sourcing exogenous variation in this context, we use the Oster (2019) test - to both gauge the vulnerability of our estimates to unobserved selection, and to provide a lower bound to our least squares estimates. We show that faithfulness is an important driver of short-term educational age 16 test scores – separately for faithful Protestants and for faithful Catholics relative to their less than faithful co-believers. And we find only *one* longer-term outcome (Christian affiliation at 25) that is significantly affected by faithfulness. These findings are

robust to the inclusion of numerous non-cognitive skills and further tests. We also examine numerous dimensions of heterogeneity. We hope that our findings, that it is faithfulness that drives outcomes, will contribute to the demise of the popular perception that it is faith schools that improve outcomes.

The second chapter examines the impact of the Education Maintenance Allowance, a conditional cash transfer in England that was available nationally from 2004 to 2011, on a range of short- and long-term outcomes. Average treatment effects are identified using OLS and Inverse Probability Weighting Regression Adjustment (IPWRA). Treatment effect heterogeneity is examined using Causal Forests, a new machine learning approach. I find beneficial impacts on retention, university attendance and, for the first time, insecure work, as measured by the probability of being on a "zero hours" contract. Other outcomes (educational attainment, risky behaviours, and labour market outcomes) are found not to be impacted.

In the third chapter I analyse the impact of education on national identity using the commonly employed Raising of the School Leaving Age (RoSLA) reform of 1972 as an instrumental variable (IV). According to Woodin, McCulloch and Cowan (2013) the reform "was conceived as an agency for the promotion of culture, discrimination, and civilisation." Using Labour Force Survey data (LFS), OLS estimates suggest statistically significant relationships between years of education and two measures of national identity, whilst IV estimates suggest statistically meaningful effects on only one measure of national identity. In a separate dataset, the British Social Attitudes Survey, containing comparable outcomes, results of similar magnitudes are identified but the none of the IV estimates are significant.

Each chapter has distinct policy implications in important areas. Faith schools are popular among parents and policy makers. Recent government policy was to allow faith schools, along with grammar schools, to expand as a way of increasing the number of "outstanding" (as rated

by the Office for Standards in Education) school places. The evidence presented in chapter one suggests that this might not be a good idea, unless the government's only aim is to achieve its promise on paper. If faith schools are not important as institutions and instead it is the faith of those who go that drives some important effects, then faith schools only appear better as a result of congregating believers together. This raises important questions for schools policy.

In the second chapter – EMA – the positive benefits of a conditional cash transfer are evident, in improved attendance and reduced likelihood of future insecure work, without the possible drawbacks of a corresponding increase in consumption of demerit goods. The discussion around EMA has receded since attendance in some form of education or training became compulsory in England, but in light of the COVID-19 pandemic and the disruption to learning for young people, the policy may deserve another look. Two of those cohorts who suffered disruption to their schooling have now left full-time education. Some have gone into higher education, but others may be in a sub-optimal position in the labour market due to having a lower human capital stock than in a world where the pandemic did not happen. A new EMA, targeted at these groups and others may provide a way to "catch up" those that current catch up efforts in the education system will miss as they have already left school for good. Finally – chapter 3 casts doubt on the potency of the impact more (British) schooling has on national identity. In other settings, such as Catalan language teaching (Clots-Figueras and Masella, 2013) or Chinese textbooks (Cantoni et al, 2017)), content-based elements of education have been shown to have effects. In this chapter, the focus has been on quantity of schooling, though one might expect similar effects; more time spent around other British people might foster a sort of national spirit. In the empirical results, this does not seem to be the case. Much discussion has taken place in recent years about the promotion of British culture in schools. Whilst this paper focuses on British natives and not newcomers, it does not provide much support for this sort of approach.

1 Chapter One – The Effect of Faith (and the Non-effect of Faith Schooling) on Educational Outcomes and Beyond

1.1 Introduction

A rapidly growing literature seeks to assess the impact of a person's religious beliefs on outcomes of interest. The field began in earnest with (Iannaccone, 1998), although, as has been pointed out by Iyer (2016) and other authors, it dates back as least as far as Adam Smith. Since then, a range of studies have examined the impact of individual affiliation and belief on various behaviours, including tax morale (Torgler, 2006), labour supply (Spenkuch, 2017), political beliefs (Spenkuch & Tillmann, 2018) and health and risky behaviours (Mendolia et al., 2019). This paper examines the effect of (what we refer to as) faithfulness on a wide range of educational outcomes, as well as wage rates, and religious affiliation 10 years later, in England.

Our religiosity measure comes from self-reported responses to the question "How important would you say your religion is *to the way you live* your life?" (Our emphasis), rather than questions of church attendance or affiliation. Attendance might be occasional and imply minimal sacrifice. Each may be coerced or be socially determined. Indeed, in a nation, like England, which has an "established" church (the Church of England), one might affiliate as a Christian simply because one identifies as English. We rely on the italicised words above to argue that our question encapsulates religiosity – belief in religious teachings that are sufficiently deep for them to affect your behaviour. Indeed, we find it difficult to see how one might answer this particular question in a way that would *not* signify faithfulness, whether habitually induced or otherwise acquired. We could see how one might say that one's religion was *important to you*, or even *important to your life*, even though you do not really believe in it. But the question specifically implies that you follow in the tenets of your religion enough for it to actually affect *the way you live* your life. We refer to our measure as *faithfulness* in what follows. We only have one measure, rather than the index of variables available in some surveys. On the other hand, our data has many outcomes and a rich range of controls that are

unavailable in other datasets. We believe we are better served by richness in the explanatory variables, and we are comfortable that our treatment variable captures the essence of what it means to be faithful.

Our research makes several contributions. First and foremost, we estimate the effect of our faithfulness measure at age 15 on a range of outcomes, which are primarily educational but extend to wages and religious affiliation at age 25. Secondly, we can separate out this faithfulness effect from the effect of attending a faith school, which we think is important given the attention that parents and policymakers give to the *apparent* effectiveness of faith schools. Third, we apply unusually rich administrative and survey data to the issue, within a context where faith schools and secular schools are virtually identical and where there is very limited scope to select by ability. Finally, employing Oster's (2019) test, we can assess the vulnerability of our conclusions based upon least squares to unobserved selection and infer bounds on our estimates associated with reasonable assumptions on the extent of selection on unobservables.

The first contribution is simple – the impact of religious belief or faith on education is not well understood with much of the quasi-experimental literature focussed on distinct historical examples that may have limited external validity. We can examine the impact of faithfulness in the modern English education system. Our outcomes are extensive both in terms of the range of variables we examine but also the times at which they are measured with academic outcomes at ages 16, 18 and 25; wage rate at age 25; whether the individual attends university, a prestigious university, and their degree attainment; and future religious belief, again at age 25. Such a thorough examination has not previously been undertaken in the England, or in any other country.

Context is key. In England this means that faith school attendance must be controlled for if our estimates are to be considered credible. Faithfulness and faith school attendance are correlated

and, if we were to exclude it, we would expect our faithfulness estimates to be biased upwards. There is already an extensive literature on the effects of faith schooling, which is conflicted over the size and significance of the effect. Papers in this area generally omit mention of the religious belief of attendees and we have not found an example that controls for a measure akin to our faithfulness variable. Much of the research is based on US data and may lack external validity for England – in particular, much of the faith school literature is focussed on US Catholic schools, where they are exclusively in the private sector. English faith schools are funded and regulated by the state that makes them essentially comparable with their secular counterparts. Any effects of faith schooling are of particular importance in England because, although the empirical evidence for England is both sparse and undecided, there is a widespread perception among policymakers and parents alike that faith schools are institutions that improve the attainment of those who attend – more so than non-faith schools. Faith schools make up a large share of schools in England – 18.7% of secondary (high) schools – and are almost always oversubscribed, and politicians are inclined to expand them, if only for that reason (Andrews and Johnes, 2016).

Thus, our second contribution is to also estimate the effect of faith schooling whilst separating out the effects of faithfulness. In so doing we can show that not only are faith school effects minimal, but that it is faithfulness that is important. This has important consequences for the English education system. By schooling religious believers together, policy makers (artificially) generate the perception that faith schools are the best place for parents to send their child. This represents an inefficiency in the school choice system as well as raising the thorny question of whether government should be promoting sorting by religion in, what in England are, taxpayer funded schools.[1] We feel that this point (still) needs to be made clearly

---

[1] The government currently has a policy of increasing the number of "good" and "outstanding" school places (as rated by inspectors) by allowing faith schools, in particular, to expand. If the effect is driven by having more religious people in attendance at those schools rather than being driven by the

to policymakers and parents. Being able to offer an alternative explanation for what can be readily seen in school league tables, an explanation that is explicitly based on selection by faith (that is unobservable in most datasets), may help displace the faith school explanation.

Thirdly, our data combine rich survey responses and detailed administrative records which cover characteristics relating to the school the young person attends (e.g., size, ethnic mix, selectivity, specialism), aspects of their family life (e.g., parental religious affiliation and faithfulness, employment, income, and education), and the neighbourhood they live in (an index of 6 indicators of deprivation covering crime and economic opportunity among other elements), as well as extensive information about that young person including affiliation, attendance and faithfulness (and also gender and ethnicity, receipt of free school meals, and 18 separate non-cognitive skills and personality traits). These proxies leave less that is unobservable, which is essential in estimating credible effects of faithfulness. But the school level information that we have, combined with the strong similarities between faith and secular schools in England, means that the faith school effect (or lack thereof) that we find might be said to be that associated with the ethos of the institution rather than other characteristics.

Of course, individuals do not randomly sort into their beliefs just as they do not randomly sort into schools, so bias arising from unobserved characteristics is an issue. Like Altonji et al. (2005b) (hereafter, AET) we do not believe that a plausible quasi-experimental strategy exists for our case. This is frequently acknowledged in the faith school literature and, we believe this applies to religiosity too. Work such as Gruber (2005), which uses religious market density to instrument religious participation, supposes that the share of co-religionists impacts on one's belief, but on nothing else. This seems rather unlikely. Credible attempts to do so are possible, and Bryan et al (2020)'s RCT work is a case in point. But this is isn't possible in our context.

---

institutions themselves, as our results suggest, then this effect may get diluted as the schools expand and so admit a larger share of non-believers, undermining any case they may have for their expansion.

Instead, AET offer a sensitivity analysis, later formalised and expanded by Oster (2019) and we use her approach here. Although employed with increasing frequency in empirical studies (with over 300 Web of Science citations), the test is still relatively new, so we devote some space to explaining its implementation in what follows. We believe that this approach is well suited to the combination of rich survey and administrative data that we have. Nevertheless, we caution that Oster's test is **not** a silver bullet that enables causal inference, but it may substantially augment the usefulness and the credibility of OLS estimates.

This links to the remaining aspect of our contribution – we use the Oster (2019) test, facilitated by our rich dataset, to challenge the credibility of our least squares estimates. The rich data facilitates the test, and it allows us to go further than previous, mostly not quasi-experimental, work that has also examined the impact of religiosity. In particular, we can explore how our least squared estimates compare with bounds generated by reasonable assumptions - the approach is not used in Sullivan et al (2018), for example, who examine the impact of religion of upbringing on educational outcomes.

In terms of our results, we find little in the way of statistically significant effects among the whole sample across virtually all our outcomes, when we control for a rich range of covariates including religious affiliation.[2] But the story is different among the single-religion subsamples where we compare **either** Catholic **or** Protestant individuals of higher faith with those of lower faithfulness in the same faith. For these comparisons, significantly positive estimates of the impacts of faithfulness are found for the number of good passes that an individual attains at General Certificate of Secondary Education (GCSE, age 16 high stakes exit, from compulsory schooling, examinations), and for the likelihood of affiliating as a Christian at age 25, a decade

---

[2] In fact, in what follows, our sample consists of **only** Protestants, broadly defined, Catholics, and those that profess no religious affiliation. This is due to the small numbers with other affiliations in the data. In addition, it helps to have a sample that is as homogenous as possible for a strategy based on Oster. Thus, our estimates need to be interpreted as being conditional on not being of non-Christian faiths.

later. These estimates are robust to the inclusion of a large range of covariates, including an extensive range of non-cognitive skills, and are robust in the context of the Oster (2019) test. Some other associations are suggested: such as for A-level educational attainment at age 18, and the likelihood of attending university, but these are not found to be robust. Numerous dimensions of heterogeneity are examined, and we find larger effects for those of lower attainment in their primary school tests (at age11) - for GCSE passes and for future Christian belief.

Our data explores the role of non-cognitive skills, including 18 separate variables that seek to capture as much of the unobserved elements of an individual's personality as possible. These cover conscientiousness, extraversion, altruism, and neuroticism. Surprisingly, these skills are shown to <u>not</u> mitigate the effects of religion for Protestants and Catholics. Because we have administrative data from the English National Pupil Database (NPD) linked to our cohort study, we are also able to examine several dimensions of heterogeneity, beyond those of some other papers that explore educational outcomes - such as, by free school meal (FSM) status (an indicator of financial hardship), and faith school denomination.

Including faith schooling in the analysis allows us to speak to the effects of the ethos of where a young person is educated, versus the impacts of their own intrinsic beliefs. Our findings suggest that individual faithfulness is more important than the institution in which one is educated. Indeed, as is common in the faith schooling literature, we find that the impacts of faith schools are economically and statistically insignificant for virtually all of our outcomes. This is true regardless of whether faithfulness is included in specifications.

We provide lower bounds on our OLS estimates that are derived from the Oster test analysis. These are both surprisingly *tight* and tell a surprisingly consistent story - that the effects of faith schools on educational outcomes are really the effect of faithfulness, not the school. Indeed,

the effectiveness of faith schools is statistically insignificant for both the faithful, and those that are not. It is only future religious affiliation where we find a positive impact of faith schooling. This may be the implicit rationale of faith schools, and heightened faith persistence might be construed as evidence of a desirable outcome for religious parents who sent their child to a faith school. But this paper demonstrates the ineffectiveness of faith schools to add value (in terms of educational outcomes) and reveals that it is due to a selection issue associated with faithfulness.

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

1.2 Literature

The religiosity literature has a distinct strand relating to its connection to education. Hungerman (2014a) discusses religion in the context of club goods, where individuals have the option of religious consumption or secular consumption. The presence of potential free riders who want the benefits of religious consumption without necessarily conforming to certain practises and rules leads religious groups to emphasise certain behaviours to screen out the unfaithful. These behaviours may include an expectation of working hard, which has implications for educational attainment and for labour market outcomes. McCullough & Willoughby (2009) similarly suggest that religion modifies an individual's priorities so that they *want* to accord with the prescribed practices, such as hard work.

Endogeneity pervades the empirical analysis of the economic impacts of religion. Self-selection implies 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 & Mocan (2018) respectively). Moreover, the effect of education on religion may

differ across faiths (see McFarland et al. (2011)) – hence we explore the separate roles of Protestant and Catholic affiliation below.

Some work claims to identify exogenous variation. For example, Gruber (2005) uses 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 how spillovers could make this instrument invalid. Along more historical lines, Becker & 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 they find a positive and significant impact on literacy. Similarly, Spenkuch (2017) uses a 1555 treaty to engineer a fuzzy regression discontinuity design (RDD). Serfs followed the faith of their territorial lord (either Catholic or Protestant) creating a patchwork of religious populations that correlates strongly with the situation today. Protestants are found to work longer hours, and although they do not earn higher wages, they earn more as a result of being paid for more hours of work. This latter instrument is likely to suffer from concerns about cultural persistence.

In related work, Squicciarini (2020) finds evidence that religion slowed the spread of technological progress in 19th century France, with religious education provision being the key mechanism. Becker et al. (2017) find, using a unique dataset and fixed effect methods, a negative relationship between education and church attendance. Earlier education expansion affects church attendance, but the opposite is not true. A different, and nuanced, contribution is made in Glaeser & Sacerdote (2008) who note that education is associated with increases in church attendance overall, but this varies across religious groups – those affiliated to the most educated religious denomination (Episcopalian) attend less frequently than Baptists, who are the less educated. The answer lies in the suggestion that education acts to reduce belief (thus decreasing attendance) but also increase social skills and the utility to engaging with others

(and, so, increasing attendance). Brown & Taylor (2007) agree, in showing a negative relationship between education and church attendance using the UK National Child Development Study.

Other work also suggests a role for belief. In explaining the intensity of one's work ethic, Schaltegger & Torgler (2010) find that Protestant faith is statistically significant when it is interacted with both education and intensity of religious belief. While 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.

Finally, Adamczyk (2009) uses the same religiosity measure as we use here, and they have 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. We also note that Sander (2001) and Wadsworth & Walker (2017) find a positive relationship between Catholic school attendance in the US and subsequent religiosity.

Our paper also flies close to the larger literature on faith schools, which we briefly outline. Results in this literature are mixed – some notable examples attempt quasi-experimental approaches such as: Hoxby (1994), Evans & Schwab (1995), Neal (1997), Cohen-Zada & Elder (2009), West & Woessmann (2010) and Allen & Vignoles (2016).[3] All except the last focus on Catholic schools in the US. Naturally, there are reasons that these instruments may not be valid. For example, if culture and values are highly persistent then the historical population in an area

---

[3] Though this last article is more about school competition resulting from faith schools.

may still affect outcomes through wider cultural mechanisms rather than just through faith schooling.

Indeed, Altonji et al. (2005a) argue that there is unlikely to be *any* convincing exogenous variation that would facilitate an analysis of the casual impacts of faith schools. Although validity of instruments cannot be tested per se, the authors explore a number of ways to cast doubt on the instrumental variable strategies used in the literature. In addition, they employ their own innovative method, used in their earlier work and ultimately published as Altonji et al. (2005b). They use this to support an estimated positive impact of Catholic schools on the likelihood of attending college. It is this approach that is later developed by Oster (2019).

Convincing attempts are possible, though, such as Gihleb & Giuntella (2017). Their analysis begins with an OLS regression that finds positive effects of Catholic school attendance that reduce the probability of grade repetition. In a novel IV approach, they then exploit the rapid, and differential, decline (by more than 50%) in the supply of teaching nuns that led to closures of US Catholic schools, following the Vatican II reforms of 1962-65. This is more convincing than other IV attempts because it exploits both spatial and temporal variation - and they find *no effect* on grade repetition, contrary to their OLS results. Interestingly, and importantly for our work, they use the AET method to examine the robustness of their 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. This interesting application shows the potential for AET based methods to bring conclusions based on OLS in line with IV. We find this work compelling.

In the UK context, Gibbons and Silva (2011) also assess the impact of faith school – like us, without a quasi-experimental strategy. They too argue that no convincing source of exogenous variation for faith schools in England is likely to exist. Their focus is on primary schools rather

while we focus on secondary schools in this paper. The moderate results they identify are not robust in the context of the AET method, similar to our results.

At the intersection of these two literatures are those papers that examine religion and faith schooling effects together. As Cohen-Zada & Sander (2008), who re-examine the Neal (1997) work, point out, there are studies that control for religious affiliation in the estimation of Catholic school effects, although many do not control for religious groups other than Catholic.

Sullivan et al. (2018) do explicitly control for religious affiliation - because part of its research question is to examine the long-term impact of Britain's faith schooling whilst controlling explicitly for the individual's faith of upbringing. They use the British Cohort Study of children born in 1970 and examine the effect of faith of upbringing on the number of GCSEs, whether they have any A-level passes, and the highest education level attained by the age of 42. Being raised as "other Christian" correlates with GCSEs, being raised Protestant or Catholic correlates with getting any A-levels (age 18 high stakes tests) and being raised in any faith correlates with highest education level at age 42. Faith schooling significantly raises the number of GCSEs attained. This effect is bigger and more significant for Protestant (Church of England) than Catholic schools – the latter have only marginally significant effects on GCSEs and have no effect on the other two outcomes. However, their paper relies on correlations alone.

Our measure of religiosity is different: intensity of belief instead of faith of upbringing - a variable that likely captures wider cultural factors beyond religiosity. Our data provides a rich range of other school characteristics from administrative records to control for aspects of faith schooling that the Sullivan et al analysis does not. While we know the school size, selectivity, and gender, income and racial composition of students, Sullivan et al know only the school's denomination. This means we can better control for school ethos, including the composition of peers, beyond its other characteristics. Additionally, we know the prior test scores at age 11

21

from the matched administrative data, and we observe faithfulness *before* the exam outcomes. We also control for *parental* affiliation and faithfulness which likely captures faith of upbringing, meaning our estimates show the effect of intrinsic religiosity beyond the effect that Sullivan et al identify. Indeed, their use of "upbringing" is sufficiently vague that the answer could capture an individual's family background at a point contemporaneous to, or even after, the GCSE or A-level outcomes being assessed.

In addition, we have a wider range of outcome variables that are more granular (GCSE point score and A level point score) relative to the Sullivan paper which can only control for having any A-levels or not. Moreover, their setting is schooling in the 1970s – much has changed in the English school landscape over the following 20 years to our cohort.

1.3 Context and Data

1.3.1 Next Steps Data

This paper uses the *Next Steps* dataset (also known as the first Longitudinal Study of Young People in England, LSYPE1).[4] The dataset is a cohort study beginning in 2004 with the random sampling of Year 9 (age 13/14) pupils, from 647 randomly selected (approximately 20%) English junior high schools (for age 11-16), both state and "independent" (i.e. fee-paying/private that are often, and confusingly, referred to as "public" schools in England) - resulting in a sample of approximately 21,000 pupils (about 2.5% of the whole English cohort that *Next Steps* children belong to). Questions were asked of both the cohort member and their parents. Response rates were relatively high – 74 percent in the first wave and 85 percent in each wave thereafter (apart from a later booster sample of ethnic minorities which had a lower response rate).[5]

---

[4] The data and documentation can be found at University College London (2021b, 2021a). A description of the dataset and its history can be found at Centre for Longitudinal Studies 2018).

[5] There is evidently some attrition and so we make use of survey weights. Anders (2017) shows using survey weights is sufficient to deal with attrition in our dataset.

The study followed the cohort member through their education up until the age of 20 (Wave 7). They were all then re-contacted at age 25 (Wave 8) that enables longer term analysis to be conducted. As mentioned above, we keep only those who are Protestants, Catholics, or of no religion.[6] We also keep only those in state schools – meaning we drop the roughly 4% of the sample who are in independent schools. The measure of religiosity, discussed further below, and all of the control variables come from Wave 1, when the cohort member is aged 14 or 15. All outcomes are from age 16 or later, so all of the right-hand side variables used in analysis below are measured pre-treatment.

From *Next Steps*, we extract controls for personal characteristics, such as gender (male/female) and ethnicity (white/non-white) and for various characteristics about the individual's household, including parental reports of their employment, education, and religiosity, and other variables such as area-level deprivation. We also have access to the data from the National Pupil Database (UCL, 2021a), the UK government's administrative dataset for education in England, for *Next Steps* children. This gives us prior and subsequent academic attainment for the *Next Steps* cohort and allows the identification of the school's denomination (Protestant, Catholic, other or none). Beyond this, a large range of school characteristics, such as the gender of intake, the size of the school, and whether it has a "sixth form" (that provides for post-16 education without switching to another location) attached to it. The full list of control variables can be found in the descriptive statistic tables in section 1.3.4. Most importantly, we have the faith and faithfulness of the main parent (usually the mother) – this is a self-report from the main parent, not a cross report from the child.

---

[6] This leaves 80% of the sample. There are not many Muslims in faith schools in our data and very few Muslims identify as anything other than the highest faithfulness level – indeed, fewer than 20 Muslim individuals say they have the lowest level of faithfulness – making them very different to others in the sample and not suited to the binary faith treatment we will ultimately use. Our results are thus conditional on being Protestant, Catholic, or of no religion.

Next, one's own's family financial circumstances might well play a role in determining one's location in the first place. In the absence of being able to afford to live near a good school, a parent may well exploit the priority that is given to religious background to leverage a place at a faith school, especially when the local secular alternatives are poor. This would be a way of avoiding bad school peers - even though you cannot afford to live sufficiently close to a secular school with good peers, you are likely to be able to access a place at a faith school because faith trumps proximity in the hierarchy of admission criteria. Indeed, the child might use religion as a way of coping with stress, bullying etc, associated with attending a poor school in a poor area.

We also have KS2 from the administrative data that is merged to our survey data – this is an age 11 ability measure. The effect of ability on school choice is limited (only a small proportion of state schools have managed to retain their historical rights to do so – small enough to discount). But there is a suggestion in the literature that ability is associated with open-mindedness and inquisitiveness that might lead a young person to explore religion. KS2 is very highly correlated with KS4 (GCSE's at age 16). Additionally, we have parental education that may be correlated with school choice preferences and more educated parents might be inclined to place higher weight on school quality (league table position) and less on school proximity. Proximity is associated with higher property prices and rents – so income plays a role too. In addition, local deprivation is, at this age, likely to be important through neighbourhood peer effects - which might tempt a youngster away from the path of faithfulness.

We also include 18 variables (See Section 1.4.2) that pick up non-cognitive skills – we think that we span the OCEAN typology. These will be correlated with motivation, determination, self-control.

Religiosity in *Next Steps* is measured in accordance with recommendations in McAndrew & Voas (2011). The survey asks about: affiliation, the extensive margin of religiosity; practise,

derived from questions relating to religious class attendance; and belief, taken from the question "How important would you say your religion is to the way you live your life?", with responses of "Not at all", "Not very", "Fairly", and "Very" - what we refer to as *faithfulness*.

There are several reasons for our preference for this measure, that are rehearsed in the introduction. The first is that what we call faithfulness may capture *genuine* belief better than the other measures we have available. In the 1990 British Social Attitudes Survey weekly church attendance stood at around 12 percent (Natcen Social Research, 2021). By 2005, the year that the young people in our data are asked questions about their religiosity, this had fallen to under 10 percent, a decline of roughly a sixth; this figure had remained stable by 2015, when the respondents in our data were aged 25, in the 8th wave of the Next Steps survey. These numbers are slightly lower in *Next Steps* at around eight percent. Almost 40% of the young people in our data say their religion is very or fairly important to them, indicating a potential influence of religion on their behaviour even if they may not attend church that often. Moreover, some in our sample of 15-year-olds may not attend church by choice; instead, their parents might cajole them to attend, so affiliation would capture faithfulness but there is no reason to suspect that such unwilling affiliates would answer the faithfulness question in anything other than a truthful way.[7] Our *primary* focus is, therefore, on faithfulness.

The outcomes we have are varied. Educational outcomes start at age 16 with three outcomes defined by results on the GCSE high-stakes tests at 16 – high stakes because these contribute to accessing the post-compulsory academic track. Probably the most important benchmark in the English education system is whether the individual attained 5 "good" passes (defined as above grade C). We also compute the number of good passes, and the precise overall GCSE

---

[7] The numbers are lower than the church attendance responses in the British Social Attitudes Survey data for the same year, which may be due to our younger sample. *Next Steps* also asks about attendance at "religious classes", rather than church attendance, making comparability with other work more difficult, but we also provide estimates relating to religious classes in Appendix A.

point score. After that, the next outcome is A-Level points score – test scores resulting from exams taken at age 18 (if the individual attended the academic, rather than vocational, track of post-compulsory education).[8]

The remaining outcomes are taken from the Waves 6, 7, and 8 data (ages 19, 20, and 25). University attendance, a binary outcome, takes value 1 if the individual has attended university by age 25. Russell Group university attendance takes value 1 if the individual has attended one of the prestigious universities that make up that group, and 0 otherwise, that 24% of this cohort attended.[9] For graduates in the survey data, degree "class" attained at the end of the individual's university studies is recorded.[10] Two further outcomes remain: the hourly wage rate derived from weekly earnings and weekly hours of work at age 25; and religious affiliation at age 25.

1.3.2 Institutional Background

In England, a large majority of faith schools are state funded and regulated.[11] Faith schools have to adhere to the highly prescriptive National Curriculum; their funding arrangements are comparable to non-faith schools with the money following the pupils (although some of their buildings may be owned by the Church); the requirements for teachers are the same (although faith schools may require their staff to at least be sympathetic to religion and so may apply some degree of discrimination in hiring); and both types of school are regulated by the Office

---

[8] At GCSE, an 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 points. For A Levels, when converted to UCAS points – the point system used for university admissions – A* is worth 140, A is worth 120, B is worth 100, C is worth 80, D is worth 60, and E is worth 40.

[9] UK universities are charitable institutions that are almost exclusively funded through an income-contingent student loan system. This again means that the English context makes for a good laboratory.

[10] Most undergraduate degrees are "classified" into "First", "Upper second", "Lower second", and "Third" with the few remaining graduates being classified in a variety of ways below third class. The proportion of university drop-outs is very small. We exclude non-graduates from this outcome.

[11] We only include state schools in our analysis. Often private schools have a religious affiliation, either explicitly or implicitly. They provide schooling for a very small proportion of pupils – less than 5% at this age and time. Thus, we lack the power to analyse these private pupils separately, and we would be in breach of disclosure rules if we attempted to do so.

for Standards in Education (OfSTED). These similarities make England a good laboratory for testing the impacts of different school types.[12]

A faith school in England is defined as any that have an explicitly stated religious character. Faith schools can use religious belief as a criterion for admitting pupils, for up to half of their intake, if they are oversubscribed. The overwhelming majority of faith schools are Christian. Of these, the largest share is Roman Catholic (9.4% of all schools), with a smaller number being the established Church of England (6.1%), or of another Christian denomination (2.3%) which are broadly Protestant in nature. Much smaller numbers of schools of other faiths exist: Jewish since in 1732, Muslim since the 1950s, and Sikh and Hindu since 1999 and 2008, respectively (see Andrews & Johnes, 2016).[13]

1.3.3 Descriptive Statistics

Figure 1.1 shows faithfulness broken down by Christian denomination and by faith and non-faith school status. Those of no religion are, frustratingly, not asked how important their *lack of religion* is for the way they live their lives. As a result, we code them as "not at all" faithful - that is they are pooled with the affiliated individuals who are also respond "not at all" to the *live my life* question. This has the advantage that we can assess the impact of faith schools in the whole sample as well as just the sample of the affiliated, but it does mean that a, possibly heterogeneous, sample of individuals are grouped into the control group in terms of our treatment variable. Thus, we also show a Protestant and Catholic sub-sample column in each table.

---

[12] Despite their similarities faith schools have a (limited) ability to cherry-pick their intake in that they have marginally fewer students being from disadvantaged backgrounds or being high achievers, as Andrews and Johnes (2016) show.

[13] Every Local Authority (LA) in England has at least one faith school, though there is heterogeneity in exactly how many – some LAs have as much as 40% of school places being in faith schools (Andrews and Johnes, 2016). See Long & Danechi (2019) for an up-to-date summary of English faith school distinctiveness and suggest how faith schools may become more distinctive in the future.

In Figure 1.1 there are more devout Catholics than there are devout Protestants. As expected, those who attend non-faith schools are less likely to say their religion is important to the way they live their lives than those who attend faith schools. These patterns suggest the possibility of interaction effects, which are investigated in later analysis.

We have a second treatment of interest, apart from faithfulness, and that is faith schooling. Additionally, there is the prospect of analysing their interaction. Dwindling cell sizes as more combinations are analysed, together with our preference for using the Oster test on a binary treatment, leads us to collapse the top three faithfulness responses into one and

Figure 1.1 Importance of Religion by Religious Denomination and Faith School Status
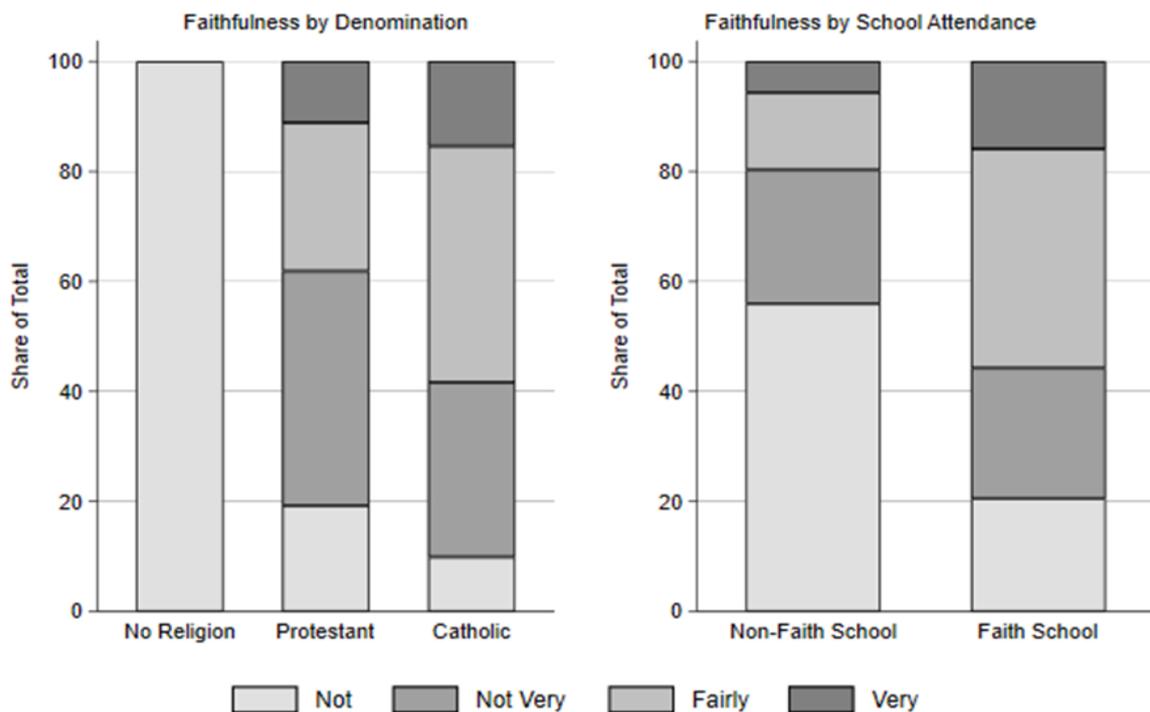
say that individuals are *not faithful* if they respond that their religion is *"not at all" important* to the way they live their life and say individuals are faithful if they respond to something more than that. Our decision on how to group the religiosity variable here is data driven – if

regressions are run including each faithfulness level, with "not at all" faithful as the reference category, we find that the coefficients on each level above this are not statistically significantly different from each other.

Descriptive statistics are given in Table 1.1, and in Appendix Tables A1 to A5. Specifically, Table A1 is individual characteristics, A2 parental and household characteristics, A3 school characteristics, and Tables A4 and A5 are the non-cognitive skill measures that are introduced in the Empirical Strategy section below. Briefly, these skills are grouped into two tables for ease of presentation, and into the OCEAN categories as best the data allows (where OCEAN is the abbreviation for Openness, Conscientiousness, Extraversion, Altruism and Neuroticism).

It is very clear, from Table 1, that the faithful generally have better educational outcomes than the non-faithful (see the first five rows, where the first, 5 GCSE passes, was effectively the gold standard of achievement at this time because it was important for progression to an academic post-compulsory track – where the difference is 18% of an SD). The same can be said for faith school pupils, relative to non-faith school pupils (where the difference for 5 GCSE passes is 24% of an SD). While this raw data suggests that the Faith * Faith School interaction may be positive, meaning that the faith schooling effect is larger for the faithful than the non-faithful, we find **no** evidence of this in the heterogeneity analysis later on.

Table 1.1   Summary Statistics – Outcomes (Mean/SD/N)

	Faithfulness			School Type		
	Non-Faithful	Faithful	Total	Non-Faith	Faith	Total
5 Good Passes at GCSE	0.46	0.55	0.50	0.49	0.61	0.50
	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)
	4528	4438	8966	7604	1279	8883
No. of Good GCSE Passes	5.64	6.65	6.14	6.02	6.96	6.16
	(4.27)	(4.10)	(4.22)	(4.25)	(3.88)	(4.21)
	4528	4438	8966	7604	1279	8883
GCSE Point Score	366.15	405.22	385.49	381.56	413.83	386.21
	(158.73)	(143.30)	(152.54)	(153.95)	(137.31)	(152.08)
	4528	4438	8966	7604	1279	8883
A Level Point Score	174.10	182.76	178.96	172.07	200.01	177.62
	(123.77)	(122.20)	(122.96)	(122.54)	(120.57)	(122.65)
	2459	3147	5606	4420	1095	5515
Attended University	0.41	0.55	0.48	0.45	0.62	0.48
	(0.49)	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)
	3906	4019	7925	6533	1272	7805
Good degree Class	0.73	0.72	0.72	0.72	0.73	0.72
	(0.45)	(0.45)	(0.45)	(0.45)	(0.44)	(0.45)
	508	681	1189	937	235	1172
Attended Russell Group	0.25	0.24	0.24	0.23	0.28	0.24
	(0.43)	(0.43)	(0.43)	(0.42)	(0.45)	(0.43)
	1474	2079	3553	2727	746	3473
Weekly Income (£)	411.22	428.88	420.25	413.57	443.89	418.45
	(268.40)	(271.93)	(270.33)	(264.11)	(293.84)	(269.31)
	2401	2511	4912	4072	781	4853
Hours Worked (All Jobs)	39.14	39.71	39.44	39.17	40.57	39.41
	(11.44)	(11.47)	(11.46)	(11.46)	(11.21)	(11.43)
	2065	2294	4359	3576	730	4306
Christian at Age 25	0.18	0.57	0.38	0.33	0.64	0.38
	(0.38)	(0.50)	(0.49)	(0.47)	(0.48)	(0.49)
	2469	2626	5095	4217	817	5034

Note: A good GCSE (school exit exams at 16) pass is defined as above grade C. Good degree is defined as 1 if the student was awarded a 1st class or upper 2nd class degree, and 0 otherwise.

1.4 Empirical Strategy

1.4.1 Specification

Our analysis uses Ordinary Least Squares (OLS) estimation of a linear specification

$$Y_{is} = \beta_0 + \beta_1 F_{is} + \beta_2 FS_{is} + \beta_3 X_{is} + \epsilon_{is} \qquad (1)$$

where Y_{is} is one of the outcomes listed above for individual i in school s. FS_{is} is a binary indicator that is zero if the individual attends a non-faith state school and one if the individual attends a faith state school. F_{is} is one if the individual is faithful (as defined above) and zero if they are not faithful. The $F_{is} * FS_{is}$ interaction is also examined later. X_{is} is a vector of individual, parental, and household characteristics including gender, ethnicity, parental employment, parental faithfulness and affiliation, and postcode level deprivation. An important control is attainment in primary school Key Stage 2 (KS2) primary exit exams, taken at age 11, prior to attending their current secondary school (until age 16). Including this in equation (1) allows it to be thought of as a value-added equation. Also included are school level characteristics: school size and the share of pupils on free school meals (FSM), a commonly used indicator of financial hardship. We also explore the role of non-cognitive skills in Section 1.4.2.

Religiosity is measured in Wave 2, when individuals are 15. Outcomes begin at age 16, ensuring that our treatment pre-dates the outcomes we are assessing. In an OLS framework such as this, it is important to narrow potential avenues for bias – imposing this timing reduces the possibility of reverse causality. All other controls are from Wave 2 or earlier for that reason.

Sensitivity analysis in empirical research is traditionally conducted by observing how treatment effect estimates change as additional control variables are included: if the list of covariates is fairly rich and the movement in the treatment effect minimal, then the threat of large amounts of unobserved selection that could nullify results might be said to be low. However, as pointed out in Oster (2019), this may not be sufficient to be informative about stability, and coefficient

31

movements need to be scaled according to the addition of covariates to the R^2. Hence, we make use of the Oster work, which points out that observed selection is only informative about unobserved selection if they are distributed similarly across treatment and controls.[14]

The test can be used to calculate a bias adjusted treatment effect (β^*).

$$\beta^* = \tilde{\beta} - \delta[\dot{\beta} - \tilde{\beta}]\frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}} \tag{2}$$

where parameters with the · accent relates to the bivariate regression of the outcome against the treatment only; while parameters with the ~ accent relate to a regression with observable characteristics included. In each case β is the treatment effect and R is the relevant R^2 value. δ is the degree of unobserved selection relative to observed selection, that is referred to as the coefficient of proportionality.

The approach rests on several assumptions. Some are given in Oster's own paper, but we opt for the list of five assumptions from De Luca et al. (2019). (1) the covariance between treatment and $\beta_3 X_{is}$ is nonzero. (2) The controls in X_{is} are uncorrelated with the error term. (3) The controls in X_{is} are mutually uncorrelated. (4) the ratio $\frac{\sigma_{1\epsilon}}{\sigma_\epsilon^2} = \delta\frac{\sigma_{1\nu}}{\sigma_\nu^2}$ holds for $\delta=1$, where ν is $\beta_3 X_{is}$. This is the equal selection relationship outlined in both Oster (2019) and De Luca et al (2019). (5) The elements of $\beta_3 = (\beta_{31}, \dots, \beta_{3n})$ are proportional to $\mu = (\mu_1, \dots, \mu_n)$, where μ is the vector of coefficients resulting from a regression of the treatment on the list of observed controls. Taken together the assumptions are strong, and they are difficult to verify.

In terms of the presentation of results, we are interested in two parameters – first we want to infer the degree of unobserved selection that would need to exist to reduce the magnitude of the treatment effect to zero, which is the parameter δ. These are shown in each regression table. The threshold for robustness in this case is generally set to one – that is, an equal amount of

---

[14] The test is implemented using the `psacalc` command in Stata.

unobserved selection to observed selection is assumed as a benchmark. Covariates are seldom included at random (unlike the assumption in the original AET work), but rather inclusion is based on theory and previous empirical work, so it is a reasonably high bar to have things that are not included account for more than that which is included, especially when the data are rich. Naturally, the higher is δ the better. The second way is to *bound* estimates assuming a particular degree of unobserved selection for δ – and this then allows us to solve for the bound, β^*, that would be consistent with the assumption. These bounds are given in Table 4. The Oster test is not a guarantee that inferences can be interpreted as causal but it does improve the credibility of OLS estimates.

Explaining *all* of the variation in an outcome, i.e., attaining an R^2 value of one, would be difficult – not least because of measurement error in observational data. Our short-term educational outcomes are based on administrative data and are likely to have low measurement error, compared to the self-reports that are much more commonly used in such research. Nonetheless, taking the amount of total variation *left* to be explained by unobserved selection could overstate the vulnerability of the estimates to selection on unobservables. Oster suggests using the observed R^2 value from the estimated regression multiplied by some number *larger* than one. Based on comparisons between observational studies and RCT estimates, suggests a rule of thumb of 1.3 would be appropriate. We also double the observed R^2.

It is important to articulate the sources of the expected omitted variable bias. The most obvious, given the context of education economics, are innate ability and family background. For example, if more able individuals are also more open minded and inquisitive, then they may be more likely to explore their faith. This could mean that both faithfulness and outcomes are higher – leading to *upward* bias in our estimate of the faithfulness effect. However, it is also possible that more able individuals are more likely to understate their faithfulness – as a result, faithfulness effect estimates could be biased downwards. We have controls for prior test scores,

which though imperfect should reduce the elements of prior (to age 16) ability that are unaccounted for. Similarly, families whose children attend faith schools could be different to those who do not. Family background may also be important. If higher family income is associated with a lesser likelihood of religious belief, as in recent evidence from Silveus & Stoddard (2020), then there could be a downward bias on our faithfulness coefficient by omitting some elements of family background that go together with great familial resources. Personality traits provide another source of potential omitted variable bias – if more conscientious individuals are more likely to be faithful and they also tend get higher grades, then our estimates would be biased upwards. Similar narratives can be formed around other traits. We control for the large number of such non-cognitive skills that are outlined below.

1.4.2 Non-cognitive Skills

We think of faithfulness as a trait. Indeed, it is common to think of other traits, such a work ethic, being associated with religious belief (for example, Weber (2001)). It is also thought that religion might act as a coping mechanism as outlined in Fruehwirth et al. (2019). Here, we include a battery of non-cognitive traits. Such traits, or skills, are not commonly observed in this context and it is natural to think that these might allow us to build a specification that further tightens the bound on OLS estimates. The list of 18 non-cognitive skills is given, grouped loosely under the OCEAN categories for which the *Next Steps* data has appropriate variables in Wave 2. To proxy for conscientiousness we include responses to: "I work as hard as I can in school"; "If I work hard I'll succeed"; "School work is a waste of time"; "School work is worth doing"; "If somebody is not a success in life, it is usually their own fault"; and "Working hard at school now will help me get on later in life". To proxy for extraversion – we include the following variables that align with Deming (2017) measures of sociability: whether the cohort member "attends clubs out of school hours"; the frequency of doing "group sport activities"; and how often the individual "sees friends either at home or elsewhere". To proxy

for altruism, we include frequency of "taking part in community work". To proxy for neuroticism we include responses to: "I can pretty much decide what will happen in my life"; "How well you get on in this world is mostly down to luck"; whether the cohort member has "recently lost sleep to worry"; their frequency of "feeling worthless recently"; to what degree they have "recently felt under strain"; how frequently they have "felt depressed recently"; and how much they have been "losing confidence in themselves recently".

The variables are included in three ways. First, all 18 are included separately in **X**; then they are grouped into the OCEAN categories given above; finally, because categorising the available data into the OCEAN framework is subjective (in terms of which category best suits a particular variable) we undertake a principal component analysis (PCA) and retain the most informative subset of the components for inclusion in the regressions. The components included are those 5 variables, out of 18, with an eigenvalue of greater than one.[15]

1.5 Results

1.5.1 OLS
OLS specifications are presented in Tables 2 and 3, and further outcomes are included in Appendix Tables A6 to A12 but are briefly summarized in the text below. These tables have three panels, and only the top (larger) one is relevant for the moment – the bottom two panels are explained later. Table 2 presents the findings for the number of GCSE passes, where we find robust effects of faithfulness but no statistically significant effects of faith schooling. Table 3 shows the effect on a Wave 8 outcome, "Being Christian at Age 25", which shows significant faithfulness *and* faith schooling effects. The pattern of controls, which come from Wave 2 prior

---

[15] This means that it passes the Kaiser-Guttman criterion, and it can be claimed to summarise more variation than any single variable (Guttman, 1954; Jackson, 1993). The screeplot in provided in Appendix Figure A1.

to any of our outcomes, is common across each of these tables.[16] Column (1) includes just the two treatment variables – faithfulness and faith school. Column (2) adds the gender, ethnicity, and religious denomination controls. Column (3) adds month of birth, KS2 score, index of multiple deprivation in one's area, and receipt of free school meals (FSM). Column (4) adds parental and household characteristics - region of residence, highest educational qualification in the family, mother's employment status, whether the individual lives in a single parent family, whether their parent was aged under 20 when the cohort member was born, the number of dependent children in the household, and the mother's religious affiliation and faithfulness level. Column (5) adds school characteristics – whether the school has a specialism, and if so what; the share of pupils on FSMs; whether the school is an academically selective "grammar" school; whether it has a "sixth form" that provides post-16 schooling; current school size; the share of pupils who do not have English as their first language; the share of pupils who are white; whether the school has a single-sex intake; and distance the individual travels to school.

Tables 1.2 and 1.3 report results *only* for the two outcomes that are the most robust. Columns (1) to (5) give results for the whole sample, including the non-affiliated, whilst columns (6) and (7) show sub-samples for Protestants and Catholics, respectively. In these subsamples Protestants (Catholics) of higher faithfulness are compared to Protestants (Catholics) who responded as "not at all" faithful. Table 2 shows the impact of faithfulness and faith school attendance on the number of good passes an individual receives at GCSE; defined as a pass at grade C or better. What is immediately evident is that faithfulness is statistically significant in all columns at the 1% level, although the magnitude drops as more controls are added.

---

[16] The results are broadly unchanged if you use the previous wave (Wave 1 at age 13) as the IZA working paper version of this paper does. Although slightly lower in general, the same outcomes display robust results.

The size of the coefficient is substantial – in Table 1.2, column (5), being faithful as opposed to not, is associated with an increase of almost 0.6 of an additional GCSE pass (about 10% of the mean). The effect is even larger in the Catholic subsample. R^2 values are also high. In contrast, the faith school effects in Table 2 are *never* statistically different from zero. The story is different in Table 3, where the outcome is future Christian affiliation. As in Table 2, the faithfulness effects are *always* significant at the 1% level (5% for the Catholic subsample). Although the fall in magnitude from column (1) to column (2) is reasonably large, the coefficient is stable thereafter. In this linear probability model, being faithful is associated with an increase in the probability that an individual is a Christian at age 25 of around 13 percentage points. Again, the effects are slightly larger for Catholics.

The important difference in Table 1.3, compared to Table 1.2, comes from the effect of faith schooling also appearing to be significant at the 1% level for the whole sample, at 5% level for Protestants, and 10% level for Catholics. The effect is broadly indistinguishable from the faithfulness effect at around 11 percentage points.

To summarise the Appendix tables for the additional short-term educational outcomes – Tables A6 and A7 are broadly similar to Table 1.2. These two outcomes are whether the individual attains the 5 good passes that is an important benchmark in English schooling and, GCSE point score. As these are just slightly different measures of GCSE attainment it is unsurprising that they are generally in agreement. In the whole sample faithfulness has positive associations whilst faith school does not. In Table A6, there is no significant effect in the Protestant subsample but in Table A7 there is – perhaps suggesting that an effect is present only among higher achievers who would get 5 good passes anyway.

Table 1.2   OLS Regression Results for Number of Good Passes at GCSE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Number of Passes at GCSE						
						Protestant	Catholic
Faithful	1.003***	0.813***	0.739***	0.597***	0.578***	0.526***	0.973**
	(0.124)	(0.188)	(0.131)	(0.136)	(0.132)	(0.144)	(0.380)
Faith School	0.346	0.429	0.239	0.091	0.153	0.269	0.054
	(0.267)	(0.286)	(0.185)	(0.187)	(0.195)	(0.259)	(0.305)
Female		0.809***	0.604***	0.627***	0.670***	0.669***	0.814***
		(0.133)	(0.077)	(0.075)	(0.076)	(0.104)	(0.188)
Non-White		-0.167	0.514***	0.397**	0.360**	0.410*	0.116
		(0.214)	(0.168)	(0.178)	(0.176)	(0.221)	(0.404)
Protestant		0.209	-0.147	-0.133	-0.164		
		(0.190)	(0.130)	(0.135)	(0.134)		
Catholic		0.024	-0.282	-0.296	-0.348*		
		(0.287)	(0.199)	(0.206)	(0.203)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	6,697	6,697	6,697	6,697	6,697	3,260	800
R^2	0.017	0.026	0.537	0.551	0.563	0.555	0.615
Oster Deltas – Faithful							
$R^2_{max} = 1$		0.008	0.352	0.260	0.261	1.341	2.684
$R^2_{max} = 2*R^2$		0.279	0.352	0.260	0.261	1.341	2.684
$R^2_{max} = 1.3*R^2$		0.738	1.000	0.703	0.672	3.563	5.543
Oster Deltas – Faith School							
$R^2_{max} = 1$		0.017	0.328	0.107	0.178	0.528	0.107
$R^2_{max} = 2*R^2$		0.626	0.328	0.107	0.178	0.528	0.107
$R^2_{max} = 1.3*R^2$		1.941	0.941	0.289	0.461	1.412	0.223

Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

Table 1.3   OLS Regression Results for Being a Christian at Age 25

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Christian at Age 25						
						Protestant	Catholic
Faithful	0.325***	0.163***	0.163***	0.131***	0.132***	0.118***	0.208**
	(0.016)	(0.027)	(0.027)	(0.028)	(0.028)	(0.030)	(0.095)
Faith School	0.217***	0.153***	0.154***	0.122***	0.108***	0.104**	0.129*
	(0.028)	(0.031)	(0.031)	(0.030)	(0.031)	(0.043)	(0.065)
Female		0.039***	0.039***	0.044***	0.045***	0.046*	0.095**
		(0.015)	(0.015)	(0.015)	(0.015)	(0.024)	(0.047)
Non-White		0.097***	0.096***	0.096***	0.105***	0.226***	0.152*
		(0.027)	(0.027)	(0.031)	(0.032)	(0.053)	(0.078)
Protestant		0.188***	0.188***	0.160***	0.162***		
		(0.027)	(0.026)	(0.028)	(0.028)		
Catholic		0.304***	0.303***	0.274***	0.274***		
		(0.041)	(0.040)	(0.041)	(0.041)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	3,650	3,650	3,650	3,650	3,650	1,837	437
R^2	0.162	0.183	0.193	0.212	0.221	0.093	0.224
Oster Deltas - Faithful							
$R^2_{max} = 1$		0.035	0.037	0.031	0.033	0.122	0.606
$R^2_{max} = 2*R^2$		0.152	0.151	0.115	0.115	1.116	2.021
$R^2_{max} = 1.3*R^2$		0.457	0.457	0.361	0.362	3.205	5.987
Oster Deltas - Faith School							
$R^2_{max} = 1$		0.093	0.099	0.078	0.061	0.080	0.221
$R^2_{max} = 2*R^2$		0.400	0.401	0.285	0.213	0.759	0.743
$R^2_{max} = 1.3*R^2$		1.213	1.222	0.902	0.687	2.361	2.256

Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

In Table A8, the story broadly repeats itself for the A-level outcome (usually taken at age 18) and is high stakes because of its role in determining progression into university. Although faithfulness is significant only at the 5% level for the whole sample, and 10% level for the Protestant subsample, it is insignificant for the Catholic subsample. Where significant associations are found they are meaningful in size at around 10 to 11% of a standard deviation. University attendance, in Table A9, is similar again – insignificant effects of faithfulness for Catholics, but for Protestants the likelihood of attending increases by around 5 percentage points if one is faithful. Instead, however, the faith school effect is significant and large for Catholics, showing an effect size of around 11 percentage points in the probability of attending.

Tables A10 and A11 relate to university attainment – elite Russell Group university attendance, and degree class (obtaining a "good" degree – meaning first or upper-second class) that both play an important role in progression into a graduate-level job. Table A12 is the effect on the log wage rate and shows no significant effects of faithfulness or of faith school; with one exception - wage rates appear to be lower for the Catholic subsample, but only at the 10% level.

Tables A14 and A15 show that the choice of treatment measure is important. Table A14 shows that if we were to focus only on affiliation, we might spuriously identify effects, as once faithfulness is included the significance associated with Protestant affiliation largely disappears. Finally, Table A15 shows that religious classes do appear to have significant effects that are broadly unchanged by the inclusion of faithfulness. Despite this, the magnitude is generally much smaller than the faithfulness effect which would lead to impacts being understated if our treatment was to focus on the measure of practice rather than belief.

1.5.2 Oster Testing

As it stands, a number of our estimates suggest significant impacts. Although the drop in magnitude, in some cases, as covariates are added *could* indicate a large degree of selection, the Oster test can help in this regard by saying how important unobserved selection would need

to be to make the estimated coefficient go to zero. At the bottom of Tables 1.2 and 1.3 (and A6 to A12) are two panels each containing three rows of Oster δs. The first row assumes a maximum attainable R^2 of 1. As explained, this may be unrealistic, so rows 2 and 3 set an alternative maximum attainable R^2 by scaling the observed R^2 of each regression by either 2 or 1.3.[17] A $\delta>1$ is an indication of robustness. In each table, δ values are given for each regression, barring the first. They can feasibly disagree as they assume unobserved selection takes the same form as observed selection – in which case the later columns should be given the greater weight.

In Table 1.2, column (5), for the final whole sample column, robustness of the faithful treatment is not met. It is close, however, in the third row, with 67% as much unobserved selection needed as observed selection to explain away results. For the subsamples, however, the standard is met even when the max R^2 is assumed to be 1. In the least conservative (third) δ row, for Protestants, unobserved selection would need to be 3.5 times bigger than observed selection to nullify effects; for Catholics it would need be to be 5.5 times bigger. This strongly suggests that the relationship between faithfulness and the number of GCSE passes is likely to be close to the unknown causal effect.

In Table 1.3, where both faithful and faith school treatments are statistically significant, robustness in the whole sample is again not supported. In the subsamples, although, the effects are found to be highly robust. For Protestants, the faithful coefficient would be robust if there were 3 times as much unobserved selection, whilst for Catholics coefficient would need 6 times as much unobserved selection to make the effect size go to zero. For the faith school coefficient both coefficients, for Protestants and Catholics, have δs of just over 2 in the third row of each Oster Delta panel.

---

[17] If scaling the observed R^2 would make R_{max} greater than 1, then R_{max} is set to one.

Across the Appendix outcomes, the whole sample meets the threshold for robustness only for the A-level point score, although it is frequently close even when it does not meet it – but, with the R^2 values being as high as they are, and the data as rich at is, these might nonetheless be said to be reasonably robust. It is difficult to be definitive as to why the *whole* sample *does not* show robust results but the sub-samples *do*. It seems likely that it arises *either* from the subsamples being more homogenous than the whole sample *or* to the fact that non-believers are not at all faithful by construction, meaning that we have artificially reduced the variation by assuming that the non-affiliated can be pooled with those who are affiliated but feel that it is "not at all" important. For Protestants the robustness threshold is met for GCSE point score, A-level point score, and university attendance. This is for the faithful coefficient only as, apart from the Christian at age 25 outcome, faith schooling effects are never significant even if they might, hypothetically, pass the Oster threshold.

For Catholics the threshold is met for the faithful coefficient in the cases of 5 good GCSE passes benchmark and the GCSE point score. In the case of the faith school effect, it is met for university attendance and the log wage rate.[18]

Table 1.4 shows the lower bounds on those estimates that were previously found to be significant and with Oster's δ >1. This makes use of a maximum R^2 assumed to be 1.3 times the observed R^2 and uses the fullest specifications from previous tables regardless of the sample being used. Their standard errors of the bounds are computed using bootstrapping with 1000 replications.

---

[18] False discovery might be a concern since we have many outcomes. But this is not as extreme as one might think. The way we approach our interpretation of the Oster test means that we ultimately focus on only the most robust coefficients (which are those that have the lowest p-values, anyway). Table A13 shows p-values for a range of ways of correcting for false discovery, for the sake of completeness. None of the alternative ways suggest that we might have misinterpreted the significance of the estimates, so we feel no need to change any of the inferences that we make.

Table 1.4   Bounded Estimates for Significant Coefficients in OLS Tables (in Main Body and Appendix)

Outcome	Original Coefficient	Lower Bound (Oster β)	N
Panel A - Whole Sample, Faithful Coefficient			
A Level Point Score	0.105**	0.178	3986
	(0.051)	(0.669)	
Panel B - Protestant Sample, Faithful Coefficient			
Number of Good Passes at GCSE	0.526***	0.401**	3260
	(0.144)	(0.160)	
GCSE Point Score	0.169***	0.146***	3260
	(0.034)	(0.039)	
A Level Point Score	0.097*	0.086	2055
	(0.055)	(0.059)	
University	0.048**	0.020	2797
	(0.023)	(0.026)	
Christian age 25	0.118***	0.091***	1837
	(0.030)	(0.031)	
Panel C - Protestant Sample, Faith School Coefficient			
Christian Age 25	0.0104**	0.068	1837
	(0.043)	(0.047)	
Panel D - Catholic Sample, Faithful Coefficient			
5 Good Passes at GCSE	0.114**	0.094*	800
	(0.049)	(0.051)	
Number of Good Passes at GCSE	0.973**	0.863*	800
	(0.380)	(0.454)	
GCSE Point Score	0.270***	0.252**	800
	(0.096)	(0.099)	
Christian age 25	0.208**	0.194***	437
	(0.095)	(0.059)	
Panel E - Catholic Sample, Faith School Coefficient			
University	0.109**	0.125**	686
	(0.045)	(0.060)	
Wage Rate	-0.020*	-0.025**	354
	(0.012)	(0.011)	
Christian Age 25	0.129*	0.106	437
	(0.065)	(0.067)	

Notes: Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls are as in Tables 2 and 3.

Some of the bounds are not significantly different from zero. The faithfulness effect for A-Level score in the whole sample, is *lower* than its lower bound, although imprecision is a problem here. The number of GCSE passes, its points score, and Christian affiliation at age 25 all remain significant for the Protestant subsample. The only faith school coefficient for the Protestant subsample that was significant earlier that does not have a lower bound that it statistically different is for Christian religion at age 25. For Catholics, the faithfulness lower bound remains significant in all cases. For the faith school effect on, university attendance and the wage rate are significant, this is not true for being Christian at 25. It is remarkable that the estimated lower bounds on the faithfulness effects binds the OLS estimates so tightly and are below the OLS estimates in all but one insignificant case.

1.5.3 Non-cognitive Skills

Up until this point the list of controls has *not* included non-cognitive skills, but they are included now. For the sake of brevity, only those results that were previously significant and robust according to the Oster δs in each OLS table (rather than the Oster βs in Table 1.4) are shown. The first column in Table 1.5 includes all non-cognitive skills as separate covariates and shows the coefficients on faithfulness and faith school, using the Protestant sample only.[19] The second includes the skills but collapsed into the OCEAN categories. The third column uses the first five principal components as discussed above. Table 5 shows the results for the Protestant Coefficient. Table A15 shows the A-level score that was previously the only outcome that was robust for the whole sample, whilst Tables A17 and A18 show the differing outcomes for the faithful and faith school coefficients, respectively, that were robust for the Catholic subsample.[20]

---

[19] The equivalent tables for Catholics, and for the whole sample can be found in Appendix A - A17 and A16.

[20] No separate table is shown for the faith school coefficient for Protestants as it is in Table A5. It is no longer significant once non-cognitive skills are included.

Table 1.5   Significant Faithful Coefficients from Earlier OLS Specifications Including Non-Cognitive Skills (Protestants Only)

	(1)	(2)	(3)
Panel A – Number of Passes at GCSE			
Faithful	0.320*	0.453**	0.383**
	(0.174)	(0.179)	(0.178)
Faith School	0.188	0.273	0.232
	(0.303)	(0.304)	(0.306)
N	2,094	2,094	2,094
R^2	0.574	0.543	0.556
Oster Delta	1.545	2.844	2.115
Panel B – GCSE Point Score			
Faithful	0.138***	0.168***	0.151***
	(0.043)	(0.044)	(0.043)
Faith School	0.078	0.087	0.075
	(0.081)	(0.082)	(0.081)
N	2,094	2,094	2,094
R^2	0.564	0.530	0.546
Oster Delta	2.986	4.818	3.783
Panel C – A Level Point Score			
Faithful	0.110	0.114*	0.113*
	(0.067)	(0.067)	(0.067)
Faith School	0.110	0.127*	0.123*
	(0.072)	(0.071)	(0.072)
N	1,399	1,399	1,399
R^2	0.390	0.354	0.358
Oster Delta	5.725	7.054	6.837
Panel D – University Attendance			
Faithful	0.022	0.041	0.034
	(0.027)	(0.027)	(0.027)
Faith School	-0.007	0.003	0.000
	(0.037)	(0.036)	(0.037)
N	1,821	1,821	1,821
R^2	0.346	0.302	0.309
Oster Delta	0.507	1.140	0.886
Panel E – Christian Religion Age 25			
Faithful	0.107**	0.121***	0.120***
	(0.042)	(0.040)	(0.040)
Faith School	0.079	0.077	0.083
	(0.059)	(0.056)	(0.056)
N	1,205	1,205	1,205
R^2	0.156	0.119	0.119
Oster Delta	2.365	3.171	3.118
Non-Cognitive Controls	All	Ocean Groups	PCA

Notes: Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls are as in Tables 2 and 3 with the addition of non-cognitive skills. There are 18 of these in total. These are given in the first paragraph of section 4.2. An maximum attainable R^2 value is set as 1.3 times the observed value for each specification.

In Table 1.5, whilst the GCSE and Christian belief at age 25 outcomes still have significant faithfulness effects associated with them in each column, the same is not true for A-level points or university attendance, where only columns (2) and (3) are significant for A-level points. For the GCSE and Christian religion outcomes, not only are coefficients broadly significant but the Oster values are still comfortably greater than one. In addition, the coefficients do not fall below the lower bounds of Table 1.4, or where they do, they are only fractionally below and not statistically different from the lower bound in question. In Table A16 the association with A level points remains significant, but only at the 10% level. In Table A17 the five good passes at GCSE outcome no longer has a significant faithful effect associated with it. Neither does the first column for number of good passes at GCSE - but the other outcomes remain significant. The GCSE point score and Christian religion at age 25 outcomes still have significant faithfulness associations at the 1% level.

1.5.4 Heterogeneity Analysis

The augmented specifications that include non-cognitive skills are used for the remaining analysis. In the main body, and in the remaining Appendix tables, a number of dimensions of heterogeneity are considered. Table 1.6 explores the possibility that being faithful becomes even more important when in a faith school so that the ethos of the person and the institution are multiplicative in terms of their effects. Table 1.6 suggests that this is *not* the case as the interaction effects are insignificant across the board.

Table 1.7 examines gender differences. Some notable differences exist – for example, the GCSE effects shown earlier in this paper seem only to exist for boys. The effect on Christian belief at age 25 is important and of similar magnitude for both subsamples. Interestingly, it appears that faith schools have large and significant effects for boys for the GCSE point score and degree class outcomes, though the latter may be an artifact of a reduced sample. Intriguingly, the GCSE points association is not mirrored in the other GCSE outcomes,

suggesting perhaps that it improves the score distribution for boys within grade, but does not impact the extensive margin.

Table 1.8 looks at heterogeneity by prior attainment. In the New Testament, the moral of the *Parable of the Talents* is that people should use the skills God gave them and put them to good use (Matthew 25:14-30 and Luke 19:11-27). As such we might expect those who arrived at high school with higher attainment would see the biggest faith effects as they proceed to work at their "talent" for studying. There is no support for this narrative. Those under the median KS2 score are little different to those above. Though the faithful effect might be said to be larger for GCSE passes for those under the median KS2 score the difference is not statistically significant according to the tests at the bottom of Table 1.8. This is perhaps more to do with imprecision of estimates as the raw data magnitudes do differ. The impact on future Christian religion is almost insignificant for those under the median KS2 score, whilst those above the median see larger effects that are significant at the 1% level.

Turning to the dimensions of heterogeneity in the appendix, little difference appears in the effects by school denomination in Table A18, where the type of faith school is split by the religious affiliation of school. Differences do occur in Table A19 where the sample is split into those on FSM and those not in receipt of FSM. Though imprecisely estimated, the impacts on GCSE outcomes are larger for those on free school meals (excluding the 5 good passes benchmark, which could be negatively impacted, although the sample is too small to be definitive). Effects on future religion are significant only for the larger group who do not get free school meals, though again this is likely to be driven by sample size. One further table, Table A20, shows the effects across all outcomes if those who are in the highest faithfulness category before the treatment was made binary are dropped. These estimated effects are slightly smaller but not statistically significantly so.

Table 1.6   Exploring Interaction Effects (Whole Sample)

VARIABLES	(1) 5+ Good Passes	(2) Number of Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Russell	(7) Degree Class	(8) Wage Rate	(9) Christian Age 25
Faithful	0.002	0.350**	0.146***	0.094	0.035	0.004	-0.070	-0.035	0.142***
	(0.020)	(0.159)	(0.039)	(0.063)	(0.026)	(0.036)	(0.057)	(0.049)	(0.037)
Faith School	0.001	-0.350	0.018	-0.072	-0.067	-0.002	0.058	-0.072	0.117
	(0.043)	(0.391)	(0.122)	(0.146)	(0.053)	(0.078)	(0.164)	(0.094)	(0.090)
Faithful*Faith School	0.019	0.504	0.023	0.193	0.076	0.021	-0.064	0.107	-0.022
	(0.047)	(0.375)	(0.106)	(0.160)	(0.058)	(0.083)	(0.182)	(0.130)	(0.100)
Protestant	-0.022	-0.240	-0.084**	-0.132**	0.005	-0.008	0.003	-0.002	0.135***
	(0.021)	(0.159)	(0.039)	(0.065)	(0.026)	(0.038)	(0.066)	(0.019)	(0.037)
Catholic	-0.029	-0.674***	-0.189***	-0.108	-0.004	0.016	-0.098	0.020	0.249***
	(0.029)	(0.229)	(0.058)	(0.082)	(0.036)	(0.049)	(0.100)	(0.037)	(0.050)
N	4,300	4,300	4,300	2,733	3,701	1,842	628	2,015	2,398
R^2	0.469	0.578	0.565	0.347	0.324	0.224	0.185	0.020	0.218

Notes: Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake. Non-cognitive skills are, grouped loosely under the OCEAN categories – conscientiousness: responses to "I work as hard as I can in school", "If I work hard I'll succeed", "School work is a waste of time", "School work is worth doing", "If somebody is not a success in life, it is usually their own fault", and "Working hard at school now will help me get on later in life". Extraversion – we include the following variables that align with Deming (2017)'s measures of sociability: whether the cohort member attends clubs out of school hours, the frequency of doing group sport activities, and how often the individual sees friends either at home or elsewhere. To proxy for altruism, we have frequency of taking part in community work. Neuroticism we include responses to: "I can pretty much decide what will happen in my life", "How well you get on in this world is mostly down to luck", whether the cohort member has recently lost sleep to worry, their frequency of feeling worthless recently, to what degree they have recently felt under strain, how frequently they have felt depressed recently, and how much they have been losing confidence in themselves recently.

Table 1.7   Heterogeneity by Gender (Whole Sample)

VARIABLES	(1) 5+ Good Passes	(2) Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Russell	(7) Degree Class	(8) Wage Rate	(9) Christian Age 25
Panel A - Female Sample Only									
Faithful	-0.004	0.223	0.076	0.120	0.067*	-0.042	-0.066	0.008	0.133***
	(0.030)	(0.211)	(0.050)	(0.086)	(0.037)	(0.048)	(0.096)	(0.012)	(0.047)
Faith School	0.024	0.006	-0.020	0.137*	0.021	0.047	-0.106	-0.004	0.102**
	(0.029)	(0.208)	(0.049)	(0.080)	(0.037)	(0.043)	(0.084)	(0.011)	(0.046)
N	2,065	2,065	2,065	1,406	1,808	997	354	1,053	1,292
R^2	0.476	0.565	0.532	0.343	0.300	0.240	0.237	0.096	0.270
Panel B - Male Sample Only									
Faithful	0.006	0.553***	0.201***	0.100	0.017	0.056	-0.011	-0.056	0.178***
	(0.028)	(0.205)	(0.047)	(0.084)	(0.034)	(0.050)	(0.101)	(0.270)	(0.049)
Faith School	0.015	-0.020	0.110**	0.062	-0.056	0.036	0.339***	0.032	0.143***
	(0.031)	(0.229)	(0.053)	(0.089)	(0.037)	(0.055)	(0.121)	(0.277)	(0.052)
N	2,235	2,235	2,235	1,327	1,893	845	274	962	1,106
R^2	0.464	0.576	0.570	0.347	0.340	0.258	0.314	0.040	0.227
Test of Female Coefficient - Male Coefficient = 0									
Faithful Coefficient	0.797	0.265	0.079	0.864	0.305	0.153	0.623	0.382	0.506
Faith School Coefficient	0.837	0.931	0.070	0.529	0.130	0.870	0.000	0.563	0.571

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1 The full list of controls is as in Table 6.

Table 1.8   Heterogeneity by Prior Attainment (Primary School/KS2 Scores, Whole Sample)

VARIABLES	(1) Five Good Passes	(2) Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Russell	(7) Degree Class	(8) Wage Rate	(9) Christian Age 25
Panel A - Under Median KS2									
Faithful	0.042	0.546**	0.152**	0.010	0.047	-0.010	0.035	-0.125	0.101*
	(0.034)	(0.267)	(0.059)	(0.145)	(0.041)	(0.050)	(0.325)	(0.413)	(0.061)
Faith School	0.007	0.127	0.107*	0.062	-0.040	0.030	-0.207	-0.021	0.068
	(0.036)	(0.282)	(0.063)	(0.129)	(0.043)	(0.051)	(0.311)	(0.393)	(0.062)
N	1,741	1,741	1,741	708	1,431	390	119	677	865
R^2	0.210	0.386	0.418	0.147	0.235	0.214	0.440	0.059	0.248
Panel B - Over Median KS2									
Faithful	-0.025	0.283*	0.118***	0.122*	0.026	0.003	-0.136*	0.003	0.183***
	(0.025)	(0.168)	(0.041)	(0.066)	(0.032)	(0.040)	(0.069)	(0.005)	(0.041)
Faith School	0.033	-0.090	-0.012	0.096	0.009	0.029	0.042	-0.003	0.120***
	(0.026)	(0.177)	(0.043)	(0.067)	(0.033)	(0.040)	(0.070)	(0.005)	(0.042)
N	2,559	2,559	2,559	2,025	2,270	1,452	509	1,338	1,533
R^2	0.217	0.309	0.367	0.297	0.239	0.226	0.199	0.092	0.259
Test of Panel A Coefficient - Panel B Coefficient = 0									
Faithful Coefficient	0.100	0.402	0.655	0.424	0.666	0.866	0.438	0.412	0.267
Faith School Coefficient	0.550	0.490	0.108	0.807	0.351	0.985	0.205	0.836	0.492

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The list of controls is as in Table 6.

1.6 Discussion and Conclusion

This paper presents robust evidence of a relationship between one's intrinsic religiosity (faithfulness) and a range of outcomes. The strongest results are found for GCSE attainment and Christian belief at age 25, but less robust associations exist for A-level point score and university attendance. Across the results presented, the most robust associations are found among Protestants. Faithfulness appears to be more important than both affiliation and frequency of attendance at religious classes (both common measures of religiosity). Overall, we feel that we have made a strong case for displacing the popular perception that it is faith schools that improve outcomes - with an explanation that is based upon faithfulness. When faith and secular schools are effectively constrained to the same things in the same way, which is a reasonable way of characterising the position in England, we would have been surprised to have found anything else.

The combination of a rich range of covariates from a unique English dataset supported by the Oster (2019) tests, and the addition of extensive measures of non-cognitive skills, all point to relationships that go beyond simple correlations. Our measure of religiosity 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 we believe that this evidence represents a convincing attempt to identify effects.

Possible explanations for the mechanism by which the faithfulness effect is operating are suggested by Hungerman (2014a) who argues that religion prescribes/proscribes certain behaviours that the faithful enact/avoid; and with McCullough and Willoughby (2009) who argue that faith provides a coping mechanism for stress. The former explanation should be accounted for in the non-cognitive skills analysis, at least in so far as work ethic is included, although it is quite conceivable that an individual does not see themselves as hard-working but, rather, as simply doing their duty.

A drawback of our approach is that we know nothing about the effects of the belief systems of those who do not profess to have a religion. Although these people do not have a religious faith, they may still have something that is recognisable as a secular faith – such as humanism. By the same token there might be said to be a distinction between spiritualism and religiosity, where spiritualism might be expressed as a person who feels closely linked to nature or similar. Future work might focus on these issues but, sadly, our data does not help us make progress in this area, and we suspect that this will need a custom survey to achieve.

In contrast to faithfulness where we find robust effects, whose lower bounds seem to be quite tight, we find that faith schooling has an association with future religious belief that is similar in magnitude to the faithfulness effect - but these associations do not seem to be robust – for any of our outcomes. This result reflects the literature, where many papers have found mixed effects of faith schools. Research that examines faith school effects in England generally does not observe significant impacts. For example, Gibbons and Silva (2011) find 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 *secondary* faith schooling. We expand on their analysis by explicitly examining the role of faithfulness, exploring a much broader range of outcomes, and exploring the role of non-cognitive skills.[21]

The policy environment around faith schools has recently taken on renewed importance. In the US, a recent Supreme Court ruling suggested that private school choice voucher programmes cannot exclude religious schools (SCOTUSblog, 2020). Evidence on the extent to which such schools cause better educational outcomes is, therefore, crucial – and our English evidence on the absence of impacts of faith schools seems relevant. Indeed, if the *only* effect of faith schools is to perpetuate faith, then it is difficult to see how subsidising them could be squared with

---

[21] This begs the question – why do parents still choose these schools if they do not improve outcomes? We explore this further in the IZA working paper version of this paper.

constitutional constraints. In the UK context, our findings also have relevance. We feel that they cast doubt on the soundness of current policy to allow faith schools to expand, especially at a time of falling rolls.

Despite the importance of faithfulness throughout the results here we have not been able to explain them away. Even a large battery of non-cognitive skills made no difference to our central findings. Thus, the effect of faithfulness remains a mystery. It has not proved possible to pin the effect down to peer effects since the peers-of-peers method will not work in our data since we do not observe the faithfulness (or even faith) of the primary peers of the secondary school peers, unlike Mendolia et al. (2018).

A   Appendix to Chapter One

Appendix Table A.1 Summary Statistics – Individual (Mean/SD/N)

	Faithfulness			School Type		
	Non-Faithful	Faithful	Total	Non-Faith	Faith	Total
Faithful	0.00	1.00	0.50	0.44	0.79	0.50
	(0.00)	(0.00)	(0.50)	(0.50)	(0.40)	(0.50)
	5053	4979	10032	8223	1579	9802
Faith School	0.07	0.26	0.16	0.00	1.00	0.16
	(0.25)	(0.44)	(0.37)	(0.00)	(0.00)	(0.37)
	4928	4874	9802	8286	1591	9877
Protestant	0.18	0.77	0.47	0.50	0.36	0.48
	(0.39)	(0.42)	(0.50)	(0.50)	(0.48)	(0.50)
	5053	4979	10032	8286	1591	9877
Catholic	0.02	0.23	0.13	0.06	0.49	0.13
	(0.16)	(0.42)	(0.33)	(0.23)	(0.50)	(0.33)
	5053	4979	10032	8286	1591	9877
Gender	0.46	0.52	0.49	0.49	0.52	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
	4971	4929	9900	8174	1570	9744
Ethnicity	0.09	0.22	0.15	0.14	0.24	0.16
	(0.29)	(0.41)	(0.36)	(0.35)	(0.43)	(0.36)
	5053	4979	10032	8286	1591	9877
KS2 Score	27.15	27.57	27.35	27.26	27.88	27.35
	(3.96)	(3.83)	(3.90)	(3.94)	(3.56)	(3.89)
	4803	4577	9380	7957	1345	9302
Receives Free School Meals	0.12	0.10	0.11	0.11	0.11	0.11
	(0.33)	(0.30)	(0.32)	(0.32)	(0.31)	(0.32)
	4458	4356	8814	7575	1187	8762

Appendix Table A.2 Summary Statistics - Parental and Household (Mean/SD/N)

	Faithfulness			School Type		
	Non-Faithful	Faithful	Total	Non-Faith	Faith	Total
Region of Residence	5.32	5.31	5.31	5.36	5.06	5.31
	(2.50)	(2.48)	(2.49)	(2.49)	(2.49)	(2.49)
	5051	4976	10027	8282	1590	9872
Highest Qualification in Family	1.90	2.15	2.03	1.96	2.33	2.02
	(1.28)	(1.29)	(1.29)	(1.27)	(1.36)	(1.29)
	4835	4806	9641	7973	1524	9497
Mother's Employment Status	1.74	1.77	1.75	1.75	1.77	1.75
	(0.44)	(0.42)	(0.43)	(0.43)	(0.42)	(0.43)
	4833	4786	9619	7939	1535	9474
Single Parent	0.13	0.10	0.12	0.12	0.12	0.12
	(0.34)	(0.30)	(0.32)	(0.32)	(0.32)	(0.32)
	5000	4926	9926	8207	1566	9773
Young Parent (Under 20 When Born)	0.13	0.08	0.10	0.11	0.07	0.10
	(0.33)	(0.27)	(0.30)	(0.31)	(0.25)	(0.30)
	5053	4979	10032	8286	1591	9877
Number of Dependent Children in HH	2.06	2.05	2.05	2.06	2.03	2.05
	(0.97)	(0.99)	(0.98)	(0.99)	(0.96)	(0.98)
	4277	4308	8585	7083	1370	8453
Mother's Religion	0.59	0.96	0.77	0.75	0.91	0.77
	(0.49)	(0.21)	(0.42)	(0.43)	(0.29)	(0.42)
	5053	4979	10032	8286	1591	9877
Mother's Faithfulness	1.61	2.76	2.18	2.06	2.83	2.19
	(0.84)	(0.98)	(1.08)	(1.04)	(1.06)	(1.08)
	4882	4847	9729	8026	1548	9574
Distance to School in KM	3.24	3.75	3.49	3.15	4.93	3.44
	(5.65)	(4.28)	(5.02)	(4.82)	(5.18)	(4.92)
	5045	4974	10019	8276	1589	9865
IMD Score	22.67	22.45	22.56	22.34	23.96	22.60
	(16.77)	(16.51)	(16.64)	(16.60)	(17.25)	(16.71)
	5043	4974	10017	8275	1588	9863

Appendix Table A.3 Summary Statistics – School (Mean/SD/N)

	Faithfulness			School Type		
	Non-Faithful	Faithful	Total	Non-Faith	Faith	Total
Protestant School	0.03	0.07	0.05	0.00	0.35	0.05
	(0.18)	(0.26)	(0.22)	(0.00)	(0.48)	(0.22)
	4877	4759	9636	8286	1425	9711
Catholic School	0.02	0.17	0.09	0.00	0.65	0.09
	(0.15)	(0.37)	(0.29)	(0.00)	(0.48)	(0.29)
	4877	4759	9636	8286	1425	9711
Share of Pupils on Free School Meals	14.23	14.50	14.36	15.02	12.14	14.56
	(12.34)	(14.06)	(13.23)	(13.16)	(13.32)	(13.22)
	4967	4939	9906	8286	1591	9877
School has Sixth Form	0.56	0.59	0.57	0.59	0.53	0.58
	(0.50)	(0.49)	(0.49)	(0.49)	(0.50)	(0.49)
	4967	4939	9906	8286	1591	9877
	1134.64	1114.15	1124.54	1140.84	1016.38	1123.92
Number of Pupils in Current School	(349.87)	(350.58)	(350.35)	(351.00)	(331.52)	(351.00)
	4842	4704	9546	8283	1303	9586
Number of Pupils in Primary School	325.67	320.98	323.39	326.25	304.59	323.31
	(134.21)	(141.35)	(137.74)	(138.03)	(135.72)	(137.91)
	4708	4447	9155	7875	1236	9111
Share of SEN pupils	1.55	1.54	1.55	1.40	2.20	1.51
	(9.11)	(8.29)	(8.71)	(9.55)	(1.36)	(8.89)
	4842	4704	9546	8283	1303	9586
Share of White Pupils	83.93	77.12	80.53	84.57	64.80	81.39
	(22.47)	(28.27)	(25.75)	(19.82)	(36.78)	(24.50)
	4967	4939	9906	8286	1591	9877
	6.11	9.52	7.81	7.93	7.78	7.91
Share with First Language not English	(11.28)	(15.95)	(13.91)	(14.13)	(13.23)	(13.99)
	4967	4939	9906	8286	1591	9877
Single Sex Intake	0.08	0.12	0.10	0.10	0.13	0.10
	(0.27)	(0.33)	(0.30)	(0.30)	(0.34)	(0.30)
	4842	4704	9546	8283	1303	9586

Appendix Table A.4 Summary Statistics - Non-cognitive Skills (Extraversion and Neuroticism) (Mean/SD/N)

	Faithfulness			School Type		
	Non-Faithful	Faithful	Total	Non-Faith	Faith	Total
Attend Clubs and Societies	1.63	1.84	1.73	1.69	1.92	1.73
	(0.96)	(1.06)	(1.02)	(0.99)	(1.12)	(1.01)
	4431	4437	8868	7331	1399	8730
Frequency of Group Sport	2.62	2.44	2.53	2.54	2.49	2.53
	(1.64)	(1.50)	(1.57)	(1.58)	(1.55)	(1.57)
	4964	4927	9891	8165	1570	9735
Frequency of Friends Visiting in Last Week	2.04	1.93	1.98	2.01	1.87	1.98
	(0.99)	(0.93)	(0.96)	(0.97)	(0.90)	(0.96)
	4955	4918	9873	8149	1569	9718
Frequency of Visiting Friends in Last Week	2.68	2.48	2.58	2.62	2.38	2.58
	(1.03)	(1.00)	(1.02)	(1.03)	(0.97)	(1.02)
	4957	4921	9878	8153	1569	9722
How Often Feeling Depressed Recently	1.90	1.90	1.90	1.90	1.91	1.90
	(0.98)	(0.95)	(0.96)	(0.97)	(0.93)	(0.96)
	4748	4715	9463	7787	1514	9301
How Much Losing Confidence in Oneself Recently	1.72	1.72	1.72	1.71	1.75	1.72
	(0.91)	(0.90)	(0.91)	(0.91)	(0.91)	(0.91)
	4769	4744	9513	7831	1524	9355
How Much Sleep Lost to Worry Recently	1.73	1.80	1.76	1.76	1.81	1.76
	(0.89)	(0.87)	(0.88)	(0.88)	(0.88)	(0.88)
	4796	4745	9541	7858	1519	9377
How Much Thinking Of Oneself as Worthless Recently	1.52	1.48	1.50	1.50	1.50	1.50
	(0.84)	(0.80)	(0.82)	(0.83)	(0.81)	(0.82)
	4751	4736	9487	7812	1516	9328
How Well you Get On is Down to Luck	2.94	3.00	2.97	2.96	3.01	2.97
	(0.74)	(0.72)	(0.73)	(0.73)	(0.73)	(0.73)
	4481	4471	8952	7372	1427	8799
How Often Feeling Constantly Under Strain Recently	1.97	2.01	1.99	1.97	2.06	1.99
	(0.95)	(0.92)	(0.94)	(0.94)	(0.92)	(0.94)
	4613	4643	9256	7605	1493	9098

Appendix Table A.5 Summary Statistics – Non-cognitive Skills (Conscientiousness and Altruism) (Mean/SD/N)

	Faithfulness			School Type		
	Non-Faithful	Faithful	Total	Non-Faith	Faith	Total
If I work Hard I'll Succeed in Later Life	3.41	3.50	3.45	3.44	3.50	3.45
	(0.61)	(0.59)	(0.60)	(0.61)	(0.58)	(0.60)
	4766	4769	9535	7859	1520	9379
I Work as hard as I can in School	2.87	3.02	2.94	2.94	2.98	2.95
	(0.72)	(0.69)	(0.71)	(0.71)	(0.70)	(0.71)
	4739	4725	9464	7802	1504	9306
School is waste of time	3.35	3.49	3.42	3.40	3.49	3.41
	(0.76)	(0.68)	(0.72)	(0.73)	(0.69)	(0.72)
	4758	4765	9523	7848	1524	9372
School work worth doing	3.21	3.29	3.25	3.24	3.29	3.25
	(0.77)	(0.79)	(0.78)	(0.78)	(0.79)	(0.78)
	4807	4808	9615	7918	1539	9457
I decide what happens in my life	2.77	2.81	2.79	2.78	2.85	2.79
	(0.76)	(0.78)	(0.77)	(0.77)	(0.76)	(0.77)
	4442	4381	8823	7257	1419	8676
If you Work Hard You'll Succeed	3.25	3.34	3.29	3.28	3.35	3.29
	(0.57)	(0.58)	(0.58)	(0.58)	(0.58)	(0.58)
	4766	4772	9538	7854	1525	9379
Lack of Success is Own Fault	2.76	2.75	2.76	2.77	2.71	2.76
	(0.77)	(0.75)	(0.76)	(0.76)	(0.72)	(0.76)
	4507	4423	8930	7360	1425	8785
Does Community Work	0.05	0.07	0.06	0.06	0.07	0.06
	(0.23)	(0.25)	(0.24)	(0.24)	(0.25)	(0.24)
	4969	4929	9898	8172	1570	9742

Appendix Figure A.1Screeplot of Non-cognitive Skills Principal Component Analysis
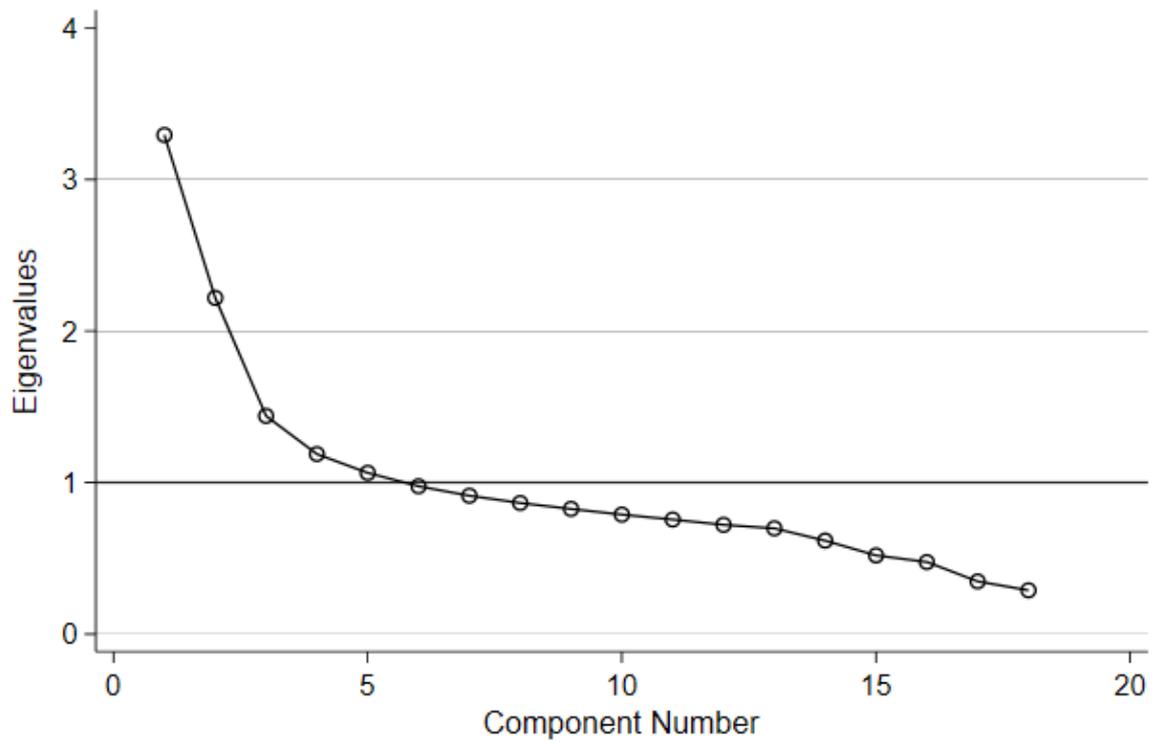
Appendix Table A.6　　　　OLS Regression Results for Attaining 5 Good Passes at GCSE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Attained 5 Good Passes at GCSE						
						Protestant	Catholic
Faithful	0.086***	0.071***	0.062***	0.039**	0.039**	0.024	0.114**
	(0.014)	(0.022)	(0.017)	(0.018)	(0.017)	(0.019)	(0.049)
Faith School	0.060*	0.062*	0.043*	0.027	0.035	0.045	0.023
	(0.031)	(0.033)	(0.022)	(0.021)	(0.022)	(0.030)	(0.038)
Female		0.066***	0.044***	0.047***	0.053***	0.052***	0.062**
		(0.015)	(0.010)	(0.010)	(0.009)	(0.013)	(0.028)
Non-White		-0.029	0.044**	0.010	0.007	0.027	0.009
		(0.025)	(0.021)	(0.022)	(0.022)	(0.030)	(0.051)
Protestant		0.014	-0.022	-0.011	-0.016		
		(0.023)	(0.017)	(0.018)	(0.018)		
Catholic		0.012	-0.019	-0.009	-0.018		
		(0.033)	(0.025)	(0.026)	(0.026)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	6,697	6,697	6,697	6,697	6,697	3,260	800
R^2	0.011	0.015	0.431	0.447	0.460	0.453	0.573
Oster Deltas - Faithful							
$R^2_{max} = 1$		0.005	0.208	0.117	0.122	0.336	1.681
$R^2_{max} = 2*R^2$		0.269	0.274	0.145	0.143	0.407	1.681
$R^2_{max} = 1.3*R^2$		0.724	0.908	0.483	0.475	1.354	4.116
Oster Deltas - Faith School							
$R^2_{max} = 1$		0.012	0.346	0.184	0.243	0.561	0.242
$R^2_{max} = 2*R^2$		0.714	0.457	0.228	0.285	0.678	0.242
$R^2_{max} = 1.3*R^2$		2.028	1.518	0.760	0.949	2.253	0.599

Notes: Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

Appendix Table A.7　　　　　　**OLS Regression Results for GCSE Point Score**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	GCSE Point Score						
						Protestant	Catholic
Faithful	0.259***	0.219***	0.204***	0.179***	0.177***	0.169***	0.270***
	(0.029)	(0.044)	(0.032)	(0.032)	(0.032)	(0.034)	(0.096)
Faith School	0.060	0.081	0.035	0.007	0.026	0.040	0.075
	(0.064)	(0.070)	(0.053)	(0.056)	(0.059)	(0.064)	(0.080)
Female		0.199***	0.151***	0.155***	0.162***	0.165***	0.209***
		(0.033)	(0.020)	(0.019)	(0.020)	(0.025)	(0.054)
Non-White		-0.023	0.142***	0.131***	0.129***	0.126**	-0.016
		(0.051)	(0.040)	(0.042)	(0.043)	(0.055)	(0.116)
Protestant		0.041	-0.046	-0.041	-0.043		
		(0.046)	(0.033)	(0.034)	(0.034)		
Catholic		-0.006	-0.080	-0.076	-0.088*		
		(0.068)	(0.050)	(0.052)	(0.052)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	6,697	6,697	6,697	6,697	6,697	3,260	800
R^2	0.019	0.029	0.536	0.549	0.557	0.562	0.559
Oster Deltas - Faithful							
$R^2_{max} = 1$		0.010	0.391	0.316	0.320	2.102	2.770
$R^2_{max} = 2*R^2$		0.295	0.391	0.316	0.320	2.102	2.770
$R^2_{max} = 1.3*R^2$		0.752	1.111	0.858	0.842	5.376	7.110
Oster Deltas - Faith School							
$R^2_{max} = 1$		0.015	0.197	0.036	0.133	0.355	1.390
$R^2_{max} = 2*R^2$		0.506	0.197	0.036	0.133	0.355	1.390
$R^2_{max} = 1.3*R^2$		1.615	0.567	0.098	0.352	0.923	3.636

Notes: Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

Appendix Table A.8 OLS Regression Results for A Level Point Score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A Level Point Score						
						Protestant	Catholic
Faithful	0.073**	0.118**	0.150***	0.103**	0.105**	0.097*	0.161
	(0.035)	(0.053)	(0.049)	(0.051)	(0.051)	(0.055)	(0.166)
Faith School	0.072	-0.005	0.063	0.030	0.072	0.103	0.065
	(0.063)	(0.069)	(0.053)	(0.050)	(0.055)	(0.070)	(0.108)
Female		0.120***	0.185***	0.200***	0.213***	0.206***	0.311***
		(0.038)	(0.028)	(0.027)	(0.027)	(0.039)	(0.069)
Non-White		-0.066	0.197***	0.120**	0.109**	0.054	-0.029
		(0.062)	(0.050)	(0.052)	(0.053)	(0.082)	(0.097)
Protestant		-0.086	-0.115**	-0.110**	-0.116**		
		(0.058)	(0.051)	(0.053)	(0.053)		
Catholic		0.078	0.018	0.019	-0.012		
		(0.084)	(0.072)	(0.072)	(0.073)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	3,986	3,986	3,986	3,986	3,986	2,055	529
R^2	0.003	0.009	0.276	0.305	0.320	0.319	0.430
Oster Deltas - Faithful							
$R^2_{max} = 1$		0.005	0.425	0.226	0.248	0.894	-6.522
$R^2_{max} = 2*R^2$		0.560	1.113	0.514	0.526	1.908	-8.635
$R^2_{max} = 1.3*R^2$		1.474	3.655	1.704	1.742	6.309	-28.550
Oster Deltas - Faith School							
$R^2_{max} = 1$		0.000	0.259	0.096	0.290	0.817	-0.253
$R^2_{max} = 2*R^2$		-0.031	0.680	0.219	0.615	1.744	-0.335
$R^2_{max} = 1.3*R^2$		-0.103	2.259	0.728	2.045	5.783	-1.113

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

Appendix Table A.9   **OLS Regression Results for University Attendance (Age 25)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	University Attendance						
						Protestant	Catholic
Faithful	0.132***	0.110***	0.098***	0.062***	0.060***	0.048**	0.072
	(0.014)	(0.022)	(0.020)	(0.021)	(0.021)	(0.023)	(0.067)
Faith School	0.036	0.034	0.026	0.001	0.017	0.017	0.109**
	(0.029)	(0.031)	(0.023)	(0.022)	(0.022)	(0.033)	(0.045)
Female		0.099***	0.090***	0.095***	0.101***	0.100***	0.144***
		(0.016)	(0.012)	(0.012)	(0.011)	(0.018)	(0.031)
Non-White		0.094***	0.161***	0.125***	0.118***	0.157***	-0.031
		(0.025)	(0.026)	(0.028)	(0.028)	(0.042)	(0.058)
Protestant		0.018	0.001	0.013	0.012		
		(0.023)	(0.020)	(0.022)	(0.022)		
Catholic		0.010	-0.008	0.002	-0.008		
		(0.035)	(0.030)	(0.031)	(0.031)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	5,686	5,686	5,686	5,686	5,686	2,797	686
R^2	0.020	0.032	0.258	0.294	0.304	0.291	0.398
Oster Deltas - Faithful							
$R^2_{max} = 1$		0.010	0.106	0.066	0.067	0.203	0.425
$R^2_{max} = 2*R^2$		0.279	0.302	0.159	0.153	0.494	0.643
$R^2_{max} = 1.3*R^2$		0.735	0.978	0.526	0.505	1.632	2.124
Oster Deltas - Faith School							
$R^2_{max} = 1$		0.012	0.091	0.004	0.058	0.088	1.138
$R^2_{max} = 2*R^2$		0.366	0.262	0.009	0.132	0.215	1.714
$R^2_{max} = 1.3*R^2$		1.187	0.871	0.030	0.440	0.717	5.521

Notes: Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

Appendix Table A.10 **OLS Regression Results for Russell Group University Attendance (Age 25)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Russell Group University Attendance						
						Protestant	Catholic
Faithful	-0.020	-0.018	-0.011	-0.023	-0.027	-0.017	-0.035
	(0.018)	(0.029)	(0.028)	(0.029)	(0.029)	(0.032)	(0.095)
Faith School	-0.006	-0.045	-0.035	-0.054**	-0.015	-0.042	0.056
	(0.033)	(0.031)	(0.030)	(0.027)	(0.029)	(0.034)	(0.065)
Female		-0.029	-0.011	-0.000	0.010	-0.015	0.058
		(0.021)	(0.018)	(0.018)	(0.017)	(0.023)	(0.056)
Non-White		-0.033	0.048	0.022	0.009	-0.041	0.162**
		(0.029)	(0.030)	(0.033)	(0.034)	(0.049)	(0.065)
Protestant		-0.008	-0.001	0.019	0.019		
		(0.031)	(0.029)	(0.031)	(0.031)		
Catholic		0.077*	0.061	0.084**	0.069*		
		(0.042)	(0.041)	(0.042)	(0.042)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	2,614	2,614	2,614	2,614	2,614	1,385	369
R^2	0.001	0.005	0.129	0.168	0.191	0.196	0.348
Oster Deltas - Faithful							
$R^2_{max} = 1$		0.002	0.028	0.097	0.138	0.263	2.137
$R^2_{max} = 2*R^2$		0.346	0.190	0.476	0.583	1.078	3.999
$R^2_{max} = 1.3*R^2$		1.088	0.634	1.582	1.938	3.586	13.299
Oster Deltas - Faith School							
$R^2_{max} = 1$		-0.008	-0.382	-0.347	0.320	0.216	-0.217
$R^2_{max} = 2*R^2$		-1.251	-2.560	-1.704	1.353	0.880	-0.405
$R^2_{max} = 1.3*R^2$		-3.006	-8.441	-5.574	4.504	2.907	-1.333

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

Appendix Table A.11 **OLS Regression Results for Degree Classification**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First Class or Upper Second Degree Classification						
						Protestant	Catholic
Faithful	0.003	-0.046	-0.040	-0.037	-0.033	-0.068	.
	(0.031)	(0.047)	(0.046)	(0.049)	(0.050)	(0.060)	
Faith School	-0.049	-0.009	0.005	0.007	-0.010	-0.007	.
	(0.049)	(0.051)	(0.051)	(0.051)	(0.054)	(0.071)	
Female		0.035	0.052*	0.057*	0.062*	0.046	.
		(0.030)	(0.030)	(0.030)	(0.032)	(0.037)	
Non-White		-0.169***	-0.110*	-0.120*	-0.130*	-0.401***	.
		(0.059)	(0.061)	(0.073)	(0.075)	(0.091)	
Protestant		0.070	0.072	0.028	0.030		.
		(0.048)	(0.047)	(0.054)	(0.055)		.
Catholic		0.000	-0.013	-0.047	-0.043		.
		(0.078)	(0.075)	(0.083)	(0.084)		.
							.
Individual Controls	No	No	Yes	Yes	Yes	Yes	.
Parental Controls	No	No	No	Yes	Yes	Yes	.
School Controls	No	No	No	No	Yes	Yes	.
N	867	867	867	867	867	480	.
R^2	0.001	0.013	0.063	0.108	0.132	0.207	.
Oster Deltas - Faithful							
$R^2_{max} = 1$		-0.022	-0.136	-0.308	-0.491	0.500	.
$R^2_{max} = 2*R^2$		-1.549	-1.999	-2.523	-3.222	1.897	.
$R^2_{max} = 1.3*R^2$		-4.518	-6.510	-8.322	-10.667	6.156	.
Oster Deltas - Faith School							
$R^2_{max} = 1$		0.002	-0.004	-0.009	0.019	0.073	.
$R^2_{max} = 2*R^2$		0.145	-0.059	-0.077	0.123	0.279	.
$R^2_{max} = 1.3*R^2$		0.481	-0.196	-0.257	0.409	0.928	.

Notes: Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Wage Rate at Age 25						
						Protestant	Catholic
Faithful	-0.023	0.028**	0.019	0.012	0.011	0.008	-0.003
	(0.045)	(0.013)	(0.014)	(0.012)	(0.013)	(0.010)	(0.014)
Faith School	-0.024*	-0.022	-0.024	-0.023	0.003	-0.031	-0.020*
	(0.014)	(0.014)	(0.016)	(0.019)	(0.027)	(0.042)	(0.012)
Female		-0.024	-0.022	-0.026	-0.027	0.018	0.019**
		(0.044)	(0.043)	(0.049)	(0.051)	(0.021)	(0.009)
Non-White		0.088	0.086	0.080	0.073	0.143	0.012
		(0.107)	(0.104)	(0.088)	(0.072)	(0.147)	(0.014)
Protestant		-0.061	-0.054	-0.008	-0.006		
		(0.057)	(0.050)	(0.011)	(0.011)		
Catholic		-0.073	-0.057	-0.009	-0.012		
		(0.053)	(0.039)	(0.020)	(0.026)		
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes
Parental Controls	No	No	No	Yes	Yes	Yes	Yes
School Controls	No	No	No	No	Yes	Yes	Yes
N	3,002	3,002	3,002	3,002	3,002	1,508	354
R^2	0.000	0.001	0.005	0.008	0.012	0.046	0.212
Oster Deltas - Faithful							
$R^2_{max} = 1$		0.000	-0.001	-0.001	-0.001	0.013	0.066
$R^2_{max} = 2*R^2$		-0.210	-0.172	-0.111	-0.102	0.263	0.244
$R^2_{max} = 1.3*R^2$		-0.627	-0.569	-0.370	-0.340	0.876	0.813
Oster Deltas - Faith School							
$R^2_{max} = 1$		0.001	0.004	0.006	-0.001	-0.063	1.000
$R^2_{max} = 2*R^2$		0.734	0.880	0.766	-0.046	-1.316	3.556
$R^2_{max} = 1.3*R^2$		2.267	2.896	2.537	-0.152	-4.360	10.475

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake.

Appendix Table A.13   P-values corrected for false discovery

	Five Good Passes	Passes	GCSE Score	A Level Score	University	Russell	Degree Class	Wage Rate	Christian Age 25
Whole									
Original coeff	0.039**	0.578***	0.177***	0.105**	0.060***	-0.027	-0.033	0.011	0.132***
Original P -value	0.027	0.000	0.000	0.035	0.005	0.361	0.504	0.382	0.000
Hochberg	0.137	0.000	0.000	0.139	0.031	0.504	0.504	0.504	0.000
Sidak	0.221	0.000	0.000	0.272	0.046	0.982	0.998	0.987	0.000
Bonferroni	0.246	0.000	0.000	0.312	0.047	1.000	1.000	1.000	0.000
N	6,697	6,697	6,697	3,986	5,686	2,614	867	3,002	3,650
Protestant									
Original coef	0.024	0.526***	0.169***	0.097*	0.048**	-0.017	-0.068	0.008	0.118***
Original P-value	0.214	0.000	0.000	0.064	0.037	0.599	0.254	0.445	0.000
Hochberg	0.599	0.002	0.000	0.318	0.222	0.599	0.599	0.599	0.001
Sidak	0.885	0.003	0.000	0.446	0.288	1.000	0.928	0.995	0.001
Bonferroni	1.000	0.003	0.000	0.572	0.333	1.000	1.000	1.000	0.001
N	3,260	3,260	3,260	2,055	2,797	1,385	480	1,508	1,837
Catholic									
Original coeff	0.114**	0.973***	0.270***	0.161	0.072	-0.035	.	-0.003	0.208**
Original P-value	0.022	0.011	0.005	0.355	0.282	0.717	.	0.001	0.029
Hochberg	0.109	0.066	0.036	0.709	0.709	0.717	.	0.005	0.115
Sidak	0.161	0.085	0.040	0.970	0.929	1.000	.	0.005	0.209
Bonferroni	0.174	0.088	0.041	1.000	1.000	1.000	.	0.005	0.231
N	800	800	800	529	686	369		354	437

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Wage rate is omitted for Catholics as the number of degrees of freedom is too low to pass disclosure rules set by the UKDS, through whom we access our data. Controls are as in Tables 2 and 3.

Appendix Table A.14   **Coefficients on Religion Including and Excluding Faithful Treatment (Whole Sample)**

VARIABLES	(1) Five Good Passes	(2) Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Russell	(7) Degree Class	(8) Wage Rate	(9) Christian Age 25
Panel A - Excluding Faithful									
Protestant	0.011	0.245**	0.084***	-0.036	0.054***	-0.003	0.004	0.003	0.262***
	(0.012)	(0.096)	(0.024)	(0.037)	(0.015)	(0.021)	(0.041)	(0.011)	(0.019)
Catholic	0.011	0.089	0.048	0.072	0.038	0.047	-0.068	-0.001	0.378***
	(0.021)	(0.172)	(0.046)	(0.061)	(0.025)	(0.034)	(0.073)	(0.020)	(0.033)
N	6,743	6,743	6,743	4,005	5,726	2,625	868	3,019	3,676
R2	0.461	0.562	0.554	0.320	0.302	0.190	0.132	0.012	0.215
Panel B - Including Faithful									
Protestant	-0.016	-0.164	-0.043	-0.116**	0.012	0.019	0.030	-0.006	0.162***
	(0.018)	(0.134)	(0.034)	(0.053)	(0.022)	(0.031)	(0.055)	(0.011)	(0.028)
Catholic	-0.018	-0.348*	-0.088*	-0.012	-0.008	0.069*	-0.043	-0.012	0.274***
	(0.026)	(0.203)	(0.052)	(0.073)	(0.031)	(0.042)	(0.084)	(0.026)	(0.041)
Faithful	0.039**	0.578***	0.177***	0.105**	0.060***	-0.027	-0.033	0.011	0.132***
	(0.017)	(0.132)	(0.032)	(0.051)	(0.021)	(0.029)	(0.050)	(0.013)	(0.028)
N	6,697	6,697	6,697	3,986	5,686	2,614	867	3,002	3,650
R2	0.460	0.563	0.557	0.320	0.304	0.191	0.132	0.012	0.221
Test of Panel A – Panel B									
	0.037	0.000	0.000	0.046	0.010	0.310	0.388	0.483	0.000
	0.101	0.020	0.003	0.529	0.059	0.425	0.200	0.703	0.011

Notes: Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls are as in Tables 2 and 3 and in Appendix Tables A6 to A12.

Appendix Table A.15 **Religious Classes Coefficient Including and Excluding Faithful Treatment (Whole Sample)**

VARIABLES	(1) Five Good Passes	(2) Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Russell	(7) Degree Class	(8) Wage Rate	(9) Christian Age 25
Panel A - Excluding Faithful									
Attend Classes	0.035*	0.303*	0.067*	0.094*	0.010	0.013	-0.086	0.060	0.192***
	(0.021)	(0.155)	(0.035)	(0.052)	(0.024)	(0.030)	(0.054)	(0.057)	(0.032)
N	4,097	4,097	4,097	2,600	3,516	1,764	589	1,875	2,295
R2	0.467	0.556	0.547	0.322	0.294	0.196	0.184	0.037	0.123
Panel B - Including Faithful									
Attend Classes	0.038*	0.296*	0.061*	0.089*	0.011	0.017	-0.084	0.060	0.178***
	(0.021)	(0.156)	(0.036)	(0.052)	(0.023)	(0.030)	(0.054)	(0.058)	(0.033)
Faithful	0.032*	0.556***	0.180***	0.088*	0.048**	-0.027	-0.022	0.010	0.127***
	(0.018)	(0.136)	(0.032)	(0.052)	(0.022)	(0.030)	(0.055)	(0.007)	(0.028)
N	4,051	4,051	4,051	2,581	3,476	1,753	588	1,858	2,269
R2	0.465	0.556	0.551	0.321	0.295	0.198	0.184	0.038	0.129
Test of Panel A – Panel B									
	0.455	0.495	0.105	0.318	0.805	0.138	0.487	0.591	0.002

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Controls are as in Tables 2 and 3 and in Appendix Tables A6 to A12.

Appendix Table A.16   Significant FAITHFUL Coefficients from Earlier OLS Specifications Including Non-Cognitive Skills (Whole Sample)

VARIABLES	(1)	(2)	(3)
	A Level Point Score		
Faithful	0.111*	0.117*	0.110*
	(0.063)	(0.062)	(0.063)
Faith School	0.046	0.079	0.068
	(0.059)	(0.058)	(0.058)
N	2,733	2,733	2,733
R^2	0.365	0.338	0.346
Oster Delta	2.504	2.838	2.533
Non-Cognitive Controls	All	Ocean Groups	PCA

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Controls are as in Table 6.

Appendix Table A.17   Significant FAITHFUL Coefficients from Earlier OLS Specifications Including Non-Cognitive Skills (Catholics Only)

	(1)	(2)	(3)
Panel A - Five Good Passes at GCSE			
Faithful	-0.001	0.047	0.036
	(0.063)	(0.061)	(0.063)
Faith School	-0.000	-0.009	-0.018
	(0.051)	(0.045)	(0.045)
N	509	509	509
R^2	0.675	0.616	0.623
Oster Delta	-0.013	1.267	0.914
Panel B - Number of Good Passes at GCSE			
Faithful	0.657	1.185***	1.096**
	(0.450)	(0.433)	(0.440)
Faith School	-0.145	-0.194	-0.328
	(0.362)	(0.365)	(0.365)
N	509	509	509
R^2	0.731	0.649	0.669
Oster Delta	1.319	4.349	3.666
Panel C - GCSE Point Score			
Faithful	0.247***	0.380***	0.330***
	(0.092)	(0.102)	(0.100)
Faith School	0.031	0.024	-0.010
	(0.088)	(0.092)	(0.094)
N	509	509	509
R^2	0.697	0.587	0.609
Oster Delta	2.045	6.116	4.297
Panel D - Christian Religion Age 25			
Faithful	0.302***	0.292***	0.291***
	(0.115)	(0.101)	(0.107)
Faith School	0.057	0.131*	0.132*
	(0.096)	(0.076)	(0.077)
N	278	278	278
R^2	0.532	0.357	0.359
Oster Delta	2.640	3.970	3.557
Non-Cognitive Controls	All	Ocean Groups	PCA

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Controls are as in Table 6.

Appendix Table A.18   **Heterogeneity by School Denomination (Whole Sample)**

VARIABLES	(1) Five Good Passes	(2) Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Russell	(7) Degree Class	(8) Wage Rate	(9) Christian Age 25
Faithful	0.003	0.389**	0.151***	0.109*	0.040	0.004	-0.068	-0.029	0.142***
	(0.020)	(0.157)	(0.038)	(0.063)	(0.025)	(0.036)	(0.057)	(0.043)	(0.036)
Protestant School	0.019	0.108	0.079	0.077	-0.016	-0.055	0.122	0.037	0.096*
	(0.039)	(0.445)	(0.131)	(0.086)	(0.037)	(0.049)	(0.079)	(0.051)	(0.058)
Catholic School	0.011	0.008	-0.002	0.086	0.017	0.048	-0.061	-0.006	0.089
	(0.029)	(0.242)	(0.072)	(0.081)	(0.039)	(0.050)	(0.086)	(0.029)	(0.055)
N	4,253	4,253	4,253	2,702	3,658	1,822	619	1,989	2,370
R^2	0.468	0.578	0.564	0.346	0.322	0.222	0.194	0.020	0.219
Test of Protestant School Coefficient – Catholic School Coefficient = 0									
	0.867	0.840	0.579	0.937	0.536	0.114	0.120	0.433	0.928

Notes: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake. Non-cognitive skills are, grouped loosely under the OCEAN categories – conscientiousness: responses to "I work as hard as I can in school", "If I work hard I'll succeed", "School work is a waste of time", "School work is worth doing", "If somebody is not a success in life, it is usually their own fault", and "Working hard at school now will help me get on later in life". Extraversion – we include the following variables that align with Deming (2017)'s measures of sociability: whether the cohort member attends clubs out of school hours, the frequency of doing group sport activities, and how often the individual sees friends either at home or elsewhere. To proxy for altruism, we have frequency of taking part in community work. Neuroticism we include responses to: "I can pretty much decide what will happen in my life", "How well you get on in this world is mostly down to luck", whether the cohort member has recently lost sleep to worry, their frequency of feeling worthless recently, to what degree they have recently felt under strain, how frequently they have felt depressed recently, and how much they have been losing confidence in themselves recently.

Appendix Table A.19   **Heterogeneity by Free School Meal Status (Whole Sample)**

VARIABLES	(1) Five Good Passes	(2) Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Wage Rate	(7) Christian Age 25
Panel A – Gets FSM							
Faithful	-0.074	1.116	0.627***	0.125	0.072	0.030	0.222
	(0.089)	(0.760)	(0.195)	(0.500)	(0.114)	(0.076)	(0.221)
Faith School	0.118	0.029	0.013	0.367	-0.070	0.049	0.123
	(0.076)	(0.650)	(0.167)	(0.329)	(0.096)	(0.047)	(0.163)
N	353	353	353	143	279	106	151
R^2	0.497	0.549	0.530	0.485	0.430	0.578	0.477
Panel B – Not on FSM							
Faithful	0.003	0.358**	0.115***	0.106*	0.037	-0.035	0.151***
	(0.021)	(0.149)	(0.034)	(0.060)	(0.026)	(0.132)	(0.034)
Faith School	0.015	0.009	0.049	0.095	-0.008	0.020	0.120***
	(0.022)	(0.159)	(0.037)	(0.061)	(0.027)	(0.134)	(0.036)
N	3,947	3,947	3,947	2,590	3,422	1,909	2,247
R^2	0.454	0.558	0.542	0.331	0.311	0.021	0.230
Test of Panel A Coefficient - Panel B Coefficient = 0							
Faithful Coefficient	0.274	0.188	0.005	0.964	0.726	0.249	0.635
Faith School Coefficient	0.201	0.974	0.830	0.320	0.481	0.500	0.982

Notes: Cluster robust standard errors in parentheses, clustered by school; *** p<0.01, ** p<0.05, * p<0.1. Note: Russell Group University Attendance and Degree Class outcomes are omitted from this table as the degrees of freedom for the sample who receive FSM were too low to satisfy disclosure rules. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake. Non-cognitive skills are, grouped loosely under the OCEAN categories – conscientiousness: responses to "I work as hard as I can in school", "If I work hard I'll succeed", "School work is a waste of time", "School work is worth doing", "If somebody is not a success in life, it is usually their own fault", and "Working hard at school now will help me get on later in life". Extraversion – we include the following variables that align with Deming (2017)'s measures of sociability: whether the cohort member attends clubs out of school hours, the frequency of doing group sport activities, and how often the individual sees friends either at home or elsewhere. To proxy for altruism, we have frequency of taking part in community work. Neuroticism we include responses to: "I can pretty much decide what will happen in my life", "How well you get on in this world is mostly down to luck", whether the cohort member has recently lost sleep to worry, their frequency of feeling worthless recently, to what degree they have recently felt under strain, how frequently they have felt depressed recently, and how much they have been losing confidence in themselves recently.

Appendix Table A.20   Highest Faithfulness Category Removed (Whole Sample)

VARIABLES	(1) Five Good Passes	(2) Passes	(3) GCSE Score	(4) A Level Score	(5) University	(6) Russell	(7) Degree Class	(8) Wage Rate	(9) Christian Age 25
Faithful	0.006	0.388**	0.146***	0.104*	0.043*	0.007	-0.069	-0.036	0.132***
	(0.020)	(0.158)	(0.038)	(0.063)	(0.025)	(0.035)	(0.057)	(0.049)	(0.036)
Faith School	0.009	-0.058	0.012	0.059	-0.019	0.021	0.006	0.010	0.110***
	(0.024)	(0.234)	(0.071)	(0.062)	(0.029)	(0.038)	(0.059)	(0.027)	(0.039)
N	4,048	4,048	4,048	2,536	3,487	1,696	586	1,907	2,265
R^2	0.466	0.578	0.564	0.345	0.321	0.224	0.185	0.021	0.211

Notes: Cluster robust standard errors in parentheses, clustered by school; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Individual controls include for month of birth, month of interview, KS2 score, postcode level deprivation, free school meal status. Parental characteristics includes region of residence, highest educational qualification in the household, parental employment, whether the parent is a single parent, whether at least one parent was aged under 20 when the cohort member was born, number of dependent children under the age of 18 in the household, parent's religion and faithfulness. School characteristics are the school's specialist subject, the share of pupils on free school meals, whether the school was a grammar school, whether the school has a sixth form attached, school size, the share of pupils who are white, the share of pupils who speak English as a first language, the distance the household lives to the school, and whether the school is single gender intake. Non-cognitive skills are, grouped loosely under the OCEAN categories – conscientiousness: responses to "I work as hard as I can in school", "If I work hard I'll succeed", "School work is a waste of time", "School work is worth doing", "If somebody is not a success in life, it is usually their own fault", and "Working hard at school now will help me get on later in life". Extraversion – we include the following variables that align with Deming (2017)'s measures of sociability: whether the cohort member attends clubs out of school hours, the frequency of doing group sport activities, and how often the individual sees friends either at home or elsewhere. To proxy for altruism, we have frequency of taking part in community work. Neuroticism we include responses to: "I can pretty much decide what will happen in my life", "How well you get on in this world is mostly down to luck", whether the cohort member has recently lost sleep to worry, their frequency of feeling worthless recently, to what degree they have recently felt under strain, how frequently they have felt depressed recently, and how much they have been losing confidence in themselves recently.

2   Chapter Two – Paying Students to Stay in School: Short- and Long-term Effects of a Conditional Cash Transfer

2.1   Introduction

The Education Maintenance Allowance (EMA) was a conditional cash transfer (CCT) that was available in the UK to those in post-compulsory full-time education or training. The level of entitlement depended on family income. EMA had been introduced in 2004 based on favourable evaluation of its pilot studies (Chowdry et al., 2008; Dearden et al., 2009; Middleton et al., 2004). The purpose was to encourage participation in post-compulsory education by pupils from low-income families. It was abolished in England in 2011 on the grounds that most students would have continued their studies after compulsory schooling ended anyway. That is, EMA represented a deadweight loss to the taxpayer (Bolton, 2011).[22]

Where CCTs have been implemented, they have generally been shown to incentivize particular activities. Some of the most notable examples are Brazil's *Bolza Escola* and Mexico's *Progresa* – both of which aimed to improve school attendance for low-income households, and both have been favourably evaluated (see for example Attanasio et al., 2012; Glewwe & Kassouf, 2012; Schultz, 2004)). Indeed, both schemes have developed into more extensive programs with wider objectives beyond education. But examples of CCTs in developed countries are relatively scarce. Given the enthusiasm for CCTs in developing countries, EMA is an important and rare example for a developed economy.[23]

Unlike many CCTs, that often focus on younger students than EMA did, the money was paid directly to the young person, bypassing their parents. It was conditional on a single behaviour

---

[22] The compulsory leaving age at the time was 16, it is now 18.

[23] Another example, from the US, is the *Opportunity NYC: Family Rewards* program that was introduced from 2007 (see the MDRC evaluations by Greenberg et al. (2011); Riccio et al. (2010); Riccio & Miller (2016).

– attendance in education or training after age 16.[24] A student in receipt of the full EMA grant – £30 per week – would receive almost £1200 per year; for a family at the threshold of eligibility for that entitlement (family income of £20,810) this would be an increase of roughly 5 percent in such a household's finances. As such, the potential impact might be expected to be large. Effects were broadly positive in the pilot studies – improving participation at ages 17 and 18. There appears to have been less of an impact on later outcomes (for example university attendance) and the estimated impacts on educational attainment were mixed. In its final full year in England the 643,000 young people who received EMA represented 32% of all 16–18-year-olds (or 47% of those in full-time post-compulsory education) at an annual cost of over £500 million (Bolton, 2011).

The contribution of this paper is that it explores avenues for further exploration that previous EMA work has left undone. Firstly, by the time of its demise, EMA had existed for approaching a decade, with ample opportunity for the nature of effects identified (for example in Dearden et al (2009)) to have changed if, for example, some early success of EMA normalised participation in education after age 16. Secondly, the pilots were in specific areas and, whilst the evaluation analysis was convincingly implemented based on matched control areas, the effect, once the programme was rolled out nationally, may be different for general equilibrium reasons. Moreover, previous analysis looked at only a limited range of short-term outcomes. The CCT literature has increasingly focused on long-term outcomes as more time passes since these schemes were implemented.[25] Finally, previous work predated the widespread adoption of machine learning methods that offer considerably greater flexibility that traditional methods

---

[24] Students were required to prove their family income was below the threshold for eligibility and to show that they had their own bank account to be credited.

[25] See Millán et al. (2019b)

and allow sources of heterogeneous effects to be revealed that would have previously been overlooked.

The *Next Steps* dataset, used in this paper, follows a single cohort of young people in England, some of whom were eligible for EMA, and provides an opportunity to look at the EMA's impact when it was no longer a novelty. Moreover, the rich nature of the data enables an analysis of the effect of the CCT on risky behaviours, where there may be potential unintended consequences of giving adolescents relatively large sums of money – for example, in the form of alcohol and cannabis consumption. The data also enables the examination of long-term labour market outcomes at age 25 – fully 8 years after first receipt of EMA.

In the absence of a quasi-experimental strategy, I proceed to identify effects under the assumption of unconfoundedness. In practice, this means that, after estimating linear specifications, I first estimate average effects using Inverse Probability Weighting Regression Adjustment (IPWRA). I then use Causal Forests, a machine learning approach that is new to economic analysis, to examine heterogeneity in treatment in a flexible and systematic way (Athey et al., 2019; Athey & Imbens, 2016; Breiman, 2001). This rapidly developing methodology has already seen use in economics; recently and prominently in Davis & Heller (2020). EMA is a good application for the method, especially as the original research based on the pilots employed fully interacted linear models and examined several dimensions of heterogeneity. Causal Forests can improve upon these traditional methods and in so doing help improve the external validity of programme evaluation. Knowing where the largest treatment effects are to be found can inform the design of future interventions.

Several statistically significant effects are identified. Results of a similar magnitude to the pilot studies (possibly slightly larger at around eight percentage points) are found on retention in full-time education and training. Positive impacts are also found on university attendance by

age 25. Attainment and degree subject choice are not impacted. Indeed, neither are any other outcomes other than the probability of being on a zero hours contract at age 25 which is reduced for those on EMA at age 17. In reducing insecure work, EMA likely has a positive impact on welfare; whilst these contracts suit some, they do not suit all, and the opportunities provided by EMA might tip the balance away from the zero hours option.

The paper proceeds as follows. Section 2 covers the relevant literature, section 3 the background to EMA and the data I use, section 4 outlines the empirical strategy, whilst sections 5 and 6 present the results and then discuss and conclude.

2.2 Related Literature

Analysis of EMA's pilots (see Chowdry et al. (2008); Dearden et al. (2009); Middleton et al. (2004)) suggested that post 16 school participation might rise substantially. The Dearden et al. work examined the pilot schemes that took place in 1999/2000 in mostly more-deprived areas of England. It found improvements of around 4.5 percentage points to participation in the first year of post-16 education (year 12) and 6.7 percentage points in the second year of post-16 education (year 13). Along with EMA participation, the Department for Education report from the pilots also examine outcomes such as attendance at university and employment, though the effects are not found to be significantly different from zero (Middleton et al., 2004).

The range of other outcomes examined has not been extensive, although Feinstein & Sabatés (2005) find reductions in crime in the EMA pilot areas when it was introduced. Nor have many studies examined the impacts of EMA outside of the pilot schemes once the scheme was rolled out nationally from 2004 onwards. One of the few examples is Holford (2015) that finds, using the same *Next Steps* dataset used in this paper, that EMA reduced the labour supply of EMA eligible teenagers by around 13 percentage points. Longer term effects have not been investigated.

Britton & Dearden (2015) conducted analysis into the 16-19 Bursary Fund that partially replaced EMA. This scheme began in 2011/12 and was targeted at the lowest income students and was mediated through schools and colleges who had discretion in who received it. Their difference-in-differences analysis uses those who were never eligible for EMA (i.e. had parental income that was too high) as the control group compared to a treatment group of those who would have been eligible but no longer received it. Participation in full-time education post age 16 fell under the new scheme – suggesting that EMA itself had a positive impact on participation.

Analysis of EMA fits into an established literature on the use of CCTs in education; although examples of CCTs in developed nations do exist, it is more common to see them used in developing countries. Developing country examples include Mexico (Attanasio et al., 2012; Schultz, 2004), Nicaragua (Gitter & Barham, 2009), and Brazil (Glewwe & Kassouf, 2012; Peruffo & Ferreira, 2017). Alternatively, cash transfers might be unconditional but are implicitly linked to education as in the Moroccan "Labelled Cash Transfer" in (Benhassine et al., 2015).

In developed nations there have been more limited use of CCTs. A CCT like EMA has been used in Australia and was shown to have increased post-compulsory schooling and attendance at university (Dearden & Heath, 1996). In contrast, a small scheme in New York City yielded modest effects at best (Greenberg et al., 2011; Riccio et al., 2010; Riccio & Miller, 2016). In the US state of Georgia, student aid conditional on attainment has been used successfully to promote higher education participation (Dynarski, 2003). In contrast, in Denmark Humlum & Vejlin (2013), the results of promoting higher education participation through a CCT proved to have no statistically significant effect despite very precisely estimated parameters.

Generally, studies relating to conditional cash transfers in education have found positive impacts on attendance, but mixed results for attainment (Fiszbein & Schady, 2009). This indicates that CCTs improve what they are conditioned on – but not necessarily other outcomes. In the case of EMA, which is conditional on school attendance, we may not expect to see improvements in other outcomes. This is the case in the Middleton et al. (2004) work on the EMA pilots where, although full-time education attendance improved, later university attendance did not change, and attainment at A-level did not either (although GCSE attainment did slightly). Although, as attendance has improved there will be those who now have post-compulsory schooling qualifications who would not have had them before. As time passes, the analysis of longer-term effects becomes possible (see Millán et al. (2019a) – a literature that my results also speak to).

Across CCT studies there is a degree of heterogeneity in who the transfer targets (i.e., who is eligible) and who physically receives the payment (the young person being educated or their parents). Often programmes are aimed at young people still in compulsory education and the money goes to their mother, as in PROGRESA in Mexico or, indeed, one arm of the EMA pilots. The money could be paid to the father instead, as in the Moroccan case in Behassine et al (2015). There is less heterogeneity around the conditions for continuing to receive the CCT. Morais de Sa e Silva (2015) finds that 80% of the 43 CCTs surveyed required continued attendance to keep receiving the transfer.

2.3 Background and Data

EMA was rolled out nationally in 2004 having been piloted from 1999. 55 local authority areas were included in the pilot scheme across two different waves. The choice of pilot areas was not random; they were generally more deprived (Fletcher, 2000). The aim of the policy was to encourage people to stay in education after age 16 – which was then the compulsory schooling age. Cash payments were made directly to pupils (rather than their parents as in the case of

many CCTs across the world and one of the EMA pilots) during school term. Amounts varied with household income – students could receive £10, £20, or £30 per week depending on their household income being below £30,810, £25,522, or £20,818 respectively, contingent on remaining in full-time education or training.[26] The relatively high threshold for eligibility means that around 32% of all 16-18-year-olds, and close to half of those 16-18-year-olds in education, received some amount of EMA. Consequently, at its 2009-10 peak, the annual cost of EMA was £580 million.

Receipt was based on continued attendance in full-time education, with eligible courses being those studied in a further education college or some apprenticeships.[27] Anecdotally, it seems that this was not monitored closely. It was simple and easy to prove eligibility – all that was needed was a statement of family income, and proof of young person having their own bank account.[28] The young person could keep working part-time and their earnings would not impact entitlement for EMA. Similarly, them receiving EMA would not impact their parents benefit entitlement. Although the guidelines were set out at a national level, the administration was performed by schools and colleges.[29] When EMA ended in England in 2011 it continued in Scotland, Wales, and Northern Ireland.

2.3.1 The *Next Steps* Dataset

The data employed are from *Next Steps* (UCL, 2021b)*,* a cohort study from England, also known as the Longitudinal Study of Young People in England (LSYPE1). The study began in

---

[26] These thresholds were not adjusted over time.

[27] By "some" I mean – Learning Skills Council (LSC) funded Entry to Employment (E2E) courses, which are a work-based learning route that some might do prior to an apprenticeship, or Programme-led apprenticeships (PLAs) that are largely classroom based – 15 hours week was needed in the classroom for these apprenticeships to qualify.

[28] The EMA website (now archived) is accessible here.

[29] This is different to the new 16-19 Bursary, the less generous successor to EMA (analysed by Brittan and Dearden (2015)), the distribution of which is essentially at the discretion of colleges with minimal national guidance.

2004, randomly sampling 650 schools before randomly sampling around 30 Year 9 (age 13-14) pupils from each. These individuals were then resurveyed each year across seven waves until they were age 20. A further, eighth wave, was undertaken in 2015 when respondents were 25. The study is similar in character to the well-known US National Longitudinal Survey of Youth (NLSY), albeit for a single cohort, in that it contains a very detailed array of characteristics.

These data are linked to the National Pupil Database (NPD) (UCL, 2021a) – the administrative dataset for education in England. This means that prior attainment, at primary school (known as Key Stage 2 or KS2 at age 11) and secondary school (GCSE point score at age 16), can be controlled for and attainment effects of EMA (on A-level achievement, at age 16) can be estimated. The NPD also contains information on the lower super output area (LSOA) that the individual lived in at age 14 – this means post-code (zip-code) level data on deprivation can be matched into the data. Moreover, the precise local authority that the individual lived in at age 14 is known, too, which means those who lived in an EMA pilot area can be identified. It is plausible that pilot areas may have become more adept at advertising EMA to their students as they have more experience of it, so it is important to control for this.

My sample ultimately contains only those who were eligible for EMA. This mirrors the previous work done on the pilot studies. Eligibility is determined by household income the year prior to attending college as that is likely the income statement used to prove eligibility. This yields a maximum possible sample of 4,859, of whom 66.1 percent receive EMA.

Individuals leave school at age 16. This coincides with Wave 3 of *Next Steps* (in 2006). Wave 4, when individuals are 17, is therefore the first year that individuals can receive EMA. Wave 4 is the first wave from which outcomes appear in this paper. These are risky behaviours: frequency of drinking alcohol and whether the individual has ever tried cannabis. Wave 5

provides information on the main activity of the young person – the outcome that is constructed takes value 1 if the individual is in full-time education or training.[30] This is the measure of retention. Wave 5 also gives the attainment measure – UCAS score. These are the grades achieved in the top 3 A-levels converted into a point score that is used by universities when judging applications. Waves 6 to 8 yield information on university attendance and degree subject choice, where we are interested to see if the latter is impacted if a group of young people from disadvantaged backgrounds choose potentially more lucrative degrees in terms of future earnings. Finally Wave 8 allows examination of labour market outcomes – these are earnings, hours worked, whether the individual has ever been employed (a measure of long-term unemployment), and whether they are currently working in insecure work, defined as working on a zero hours contract – a type of employment contract in the UK that does not guarantee a minimum number of hours of work in any given week. These outcomes are measured when the individual is age 25 and are shown in Table 2.1, where they are broken down by receipt of EMA.

Controls come from Wave 4 and earlier. The full list of controls is given in Table 2.2, again broken down by EMA receipt. The second column shows the age at which the responses were collected. These covariates span across personal characteristics such as gender and ethnicity, to household characteristics like parental employment status, and highest educational attainment. A number of these variables are included specifically to account for selection into EMA. Prior knowledge of EMA is important – being able to apply for something is helped along substantially by knowing about it; it may also account for more driven individuals who have been planning to go to college for some time prior to attending. Local authorities that were part of the EMA pilots may be more experienced in advertising and administrating the

---

[30] Including some apprenticeships as detailed in footnote 27.

grant and so an individual living in those areas may have higher likelihood of take up. Similarly, EMA enrolment in your area will account for peer effects that make individuals who have friends who are applying more likely to apply. Other control variables like parental employment and education will also impact (likely positively) selection into EMA. The variables are largely binary (such as free school meal status) or have relatively small numbers of responses (such as general health or English region of residence). This is for purposes of overlap.

In both Tables 2.1 and 2.2 it is evident that differences exist between those who are on EMA and those who are not. For the outcomes, in Table 2.1, part of this may be impacts of the CCT – this would indicate large potential impacts of EMA, such as in the case of University Attendance. In the control variables, all of which are pre-treatment, these reflect differences in the type of person who applies and ultimately receives EMA. Naturally, these differences will likely extend to unobserved characteristics as well. Insofar as these unobserved characteristics matter for selection into treatment and for outcomes their impact will be lessened if overlap is good, assuming that unobserved variables follow a similar distribution to those which are included.

As discussed below, good overlap is an important assumption for methods based on unconfoundedness, but in selecting the list of controls that yields good overlap concerns may arise of choosing controls that give the most favourable impact on the outcome (in terms of what a research might like to find). I combat this by beginning the analysis by judging which covariates deliver the best overlap whilst using a randomly generated outcome variable. This enables me to judge overlap without knowing what the effect will be. Figures 2.1 and 2.2 show that good overlap and covariate balance are achieved. Figure 2.2 is generated from a propensity score matching estimation using the Stata command teffects, but the modelling of the first stage is analogous to IPWRA. The chart is intended to show how effective matching methods appear

to be in this application – and this is shown in Figure 2.2 where the right and left graphs are visually very similar. This same list of covariates is then used in the full OLS specification, too. Balance tables are not shown but are available on request. Population weights are used throughout.

Table 2.1   Outcomes by EMA Receipt

Variable	Age	Does Not Receive EMA			Receives EMA		
		N	Mean	Std Dev	N	Mean	Std Dev
Retention	18	1,408	0.551	0.498	2,916	0.700	0.458
UCAS Score	18	643	-0.029	1.036	2,098	0.029	0.990
University Attendance	25	1,412	0.331	0.471	2,881	0.519	0.500
STEM Degree Subject	25	435	0.379	0.486	1,382	0.413	0.493
Alcohol	17	1,289	0.081	1.000	2,086	-0.073	0.99
Cannabis (Ever Tried)	17	1,619	0.368	0.482	3,139	0.297	0.457
Earnings	25	600	5.827	0.615	1,293	5.812	0.718
Ever Employed	25	847	0.953	0.212	1,864	0.922	0.268
Hours	25	695	38.193	11.284	1,479	36.686	10.724
Zero Hours	25	698	0.079	0.27	1,476	0.065	0.247

Note: Retention is a dummy variable taking value 1 if the young person is still in full-time education (FTE) in Wave 5. UCAS Score is the score attached to the top 3 grades an individual achieved and is standardised to be mean 0. University is a dummy variable taking value 1 if the individual attends university by age 25. STEM degree subject takes value 1 if an individual is studying a STEM degree, and 0 otherwise. Alcohol is frequency of alcohol consumption, standardised. Cannabis consumption is a dummy variable taking value 1 if the individual has ever tried cannabis. Earnings is log employment earnings from all jobs. Ever employed is a binary variable that is 1 if the individual has had any job by age 25, and so is a measure of long-term unemployment in some sense. Hours worked is hours in one's main job. And finally, Zero Hours is a binary variable that is 1 if the individual is on an insecure labour contract that does not guarantee a minimum number of hours in a given week. These are unweighted figures, though survey weights are used in analysis below.

Table 2.2 Covariates by EMA Receipt

|		Does Not Receive EMA			Receives EMA		
Variable	Age	N	Mean	Std Dev	N	Mean	Std Dev
Gender	14	1,623	0.455	0.498	3,179	0.534	0.499
Non-White	14	1,645	0.249	0.433	3,208	0.405	0.491
Special Educational Need	14	1,629	0.209	0.407	3,166	0.173	0.378
On Free School Meals	14	1,604	0.129	0.335	3,155	0.273	0.446
Quartile of KS2 score	11	1,551	2.526	1.100	3,065	2.586	1.124
Quartile of GCSE Points	16	1,620	2.423	1.087	3,180	2.71	1.093
Pilot Area	14	1,648	0.333	0.471	3,211	0.455	0.498
EMA Enrollment in LA	17	1,648	0.416	0.141	3,211	0.474	0.153
Aware of EMA	15	1,613	0.626	0.484	3,178	0.702	0.457
General Health	17	1,638	3.408	0.677	3,170	3.422	0.659
Main Parent Age	17	1,580	44.87	6.032	3,149	45.215	6.643
Single Parent	17	1,571	0.105	0.307	3,138	0.15	0.357
Step Family	17	1,631	0.128	0.334	3,192	0.089	0.284
Main Parent Employment	17	1,584	1.753	0.816	3,172	2.059	0.859
Main Parent Education	17	1,586	3.788	1.742	3,163	4.294	1.935
No. Dep Children in HH	17	1,631	1.505	0.969	3,195	1.897	0.848
Housing Tenure	17	1,588	1.157	0.641	3,142	1.25	0.717
IMD in Year 11	16	1,648	24.89	16.59	3,208	31.142	18.367
Urban or Rural	16	1,648	3.706	0.739	3,209	3.771	0.669
Region	16	1,648	5.345	2.444	3,208	5.056	2.414

Note: Gender, Non-white, Special Educational Needs, and Free School Meals are all binary variables that are 1 if the individual is female, not white, has special educational needs or is on free school meals, respectively. KS2 score is the test taken at the end of primary (junior) school, GCSEs are high-stakes test taken at the end of secondary (high) school. Pilot Area is a dummy variable that is 1 if the individual lives in an area that was an EMA pilot area when EMA was piloted in the late 1990s. Aware of EMA is a binary variable that takes value 1 if the individual had heard of EMA in the penultimate year of secondary (high) school – that is prior to applying for EMA. General Health is a 4-point scale describing one's health in the last 12 months, ranging from "not at all good" to "very good". Main parent age is the age of the parent who responded to the parental section of the survey' single parent is a dummy that is 1 if the individual lives in a single-parent household; Main Parent Employment and Education are the parent's occupation and highest educational qualification; Number of dependent children in the household is a count of the number of people aged under 19 who live with the young person; Housing Tenure indicates whether the house the young person lives in is owned or rented; IMD is the index of multiple deprivation, a composite of six measures (outlined further in the data section of chapter 1); Urban or Rural is over several levels from small village to large city; and finally, region is the NUTS1 statistical regions of England.

Figure 2.1   Overlap
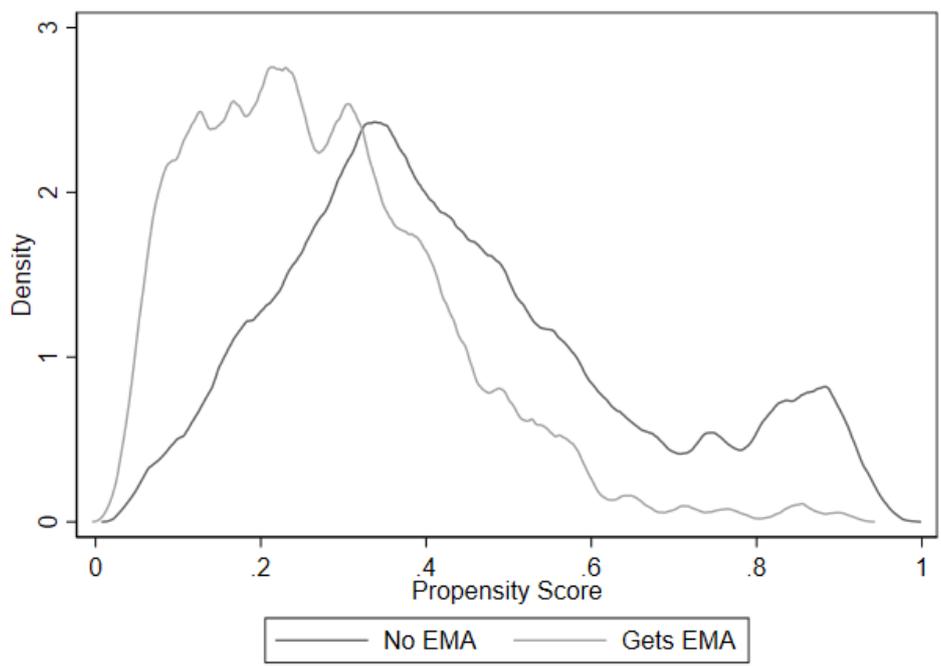
Figure 2.2   Covariate Balance
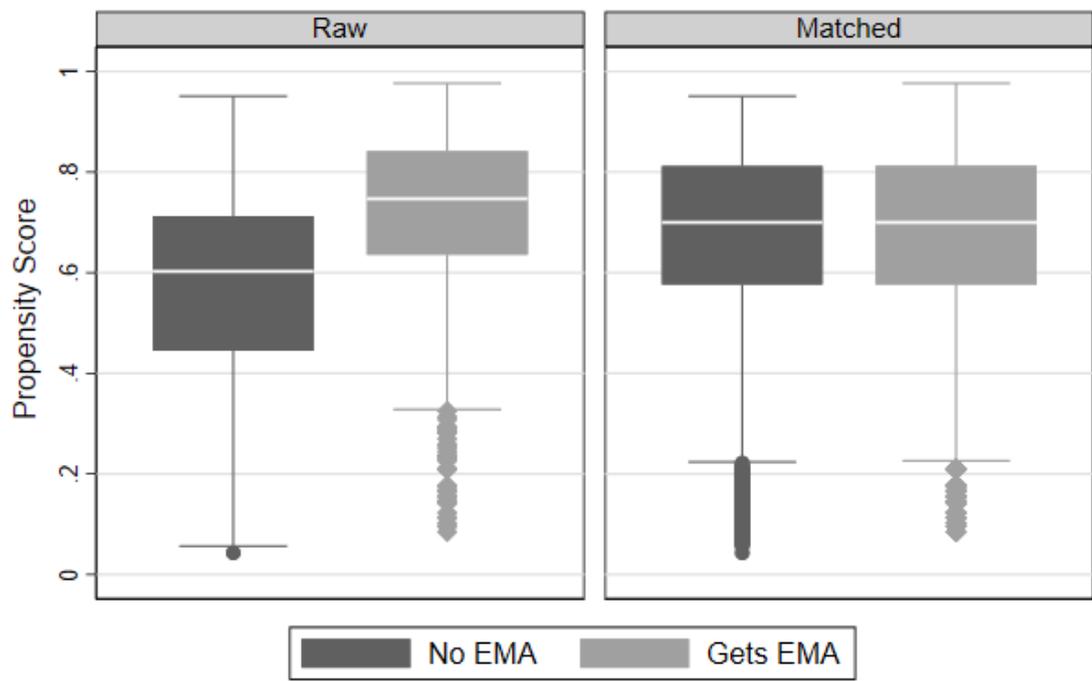
2.4 Empirical Strategy

The treatment effect of interest (τ_i) is that associated with receiving the Education Maintenance Allowance.[31] In a linear regression set up:

$$Y_i = \beta + \tau_i W_i + \boldsymbol{X}_i \gamma_i + \epsilon_i \tag{3}$$

Where Y_i is the outcome of interest (e.g., attendance in full-time education), W_i is the treatment (EMA receipt), $\boldsymbol{X}_i$ is a vector of control variables, and ϵ_i is a conventional error term.

In the light of the missing counterfactual, the paper proceeds by identifying estimates under the unconfoundedness (or ignorability) assumption – that is, that treatment assignment is unrelated to potential outcomes conditional on observed covariates.[32] Specifically, I estimate average effects using Inverse Probability Weighting Regression Adjustment (IPWRA).[33] Below, W_i is the binary treatment indicator, $Y_i(0)$ is the outcome of individual i in the absence of the treatment, $Y_i(1)$ is the outcome if the individual is treated, and X_i contains the collection of observed characteristics. The treatment is independent of the outcome, conditional on covariates.

$$W_i \perp\!\!\!\perp (Y_i(0), Y_i(1)) | X_i \tag{4}$$

Essentially, there is a trade-off between internal validity, due to the strong assumptions that underpin unconfoundedness, and external validity. As suggested by Athey & Imbens (2017), empirical approaches in economics should consider a range of methods rather than simply rely on those with the greatest internal validity (e.g., regression discontinuity, difference-in-differences (DiD), and instrumental variable (IV) methods). A second requirement is the

---

[31] The i subscript is relevant in for the later Causal Forest analysis, as individual level treatment effects can be estimated using that method.

[32] The thresholds for eligibility might facilitate a regression discontinuity design, but there are a few reasons that we do not proceed in this way. First, the data do not facilitate it – continuous measures of household income are only available two years prior to the time at which EMA is received, the discrete data after that is not appropriate for an RDD. Additionally sample sizes are too small near the cut-offs.

[33] See Imbens & Wooldridge (2009) for a more in-depth description of the method.

overlap assumption – that any given individual has a non-extreme probability (close to zero or one) of being in the treatment or control group.

$$0 < \Pr(W_i = 1 | X_i = x) < 1 \qquad (5)$$

IPWRA models both the treatment (EMA receipt) and the outcome in two separate equations. A propensity score is estimated that models 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 (observed) selection into treatment from the outcome equation. Based on the unconfoundedness and overlap assumptions, IPWRA is likely to be closer to causal effects than OLS by accounting for two levels of selection – in treatment and outcome. IPWRA exhibits the so-called "double robustness" property that means it produces consistent estimates even if one of the two equations is incorrectly specified. IPWRA also requires good overlap. As outlined above – I generate the overlap "blind" by using a randomly generated "x" as the outcome variable to avoid any temptation of picking the specification that yields a particular outcome.

Beyond average effects, I am interested in heterogeneity, because it is reasonable to assume that treatment will vary by at least some observable characteristics. To examine this, I employ Causal Forests (CFs).[34] CFs are a recent, and still rapidly developing, innovation in the application of machine learning for causal inference (Athey, 2017; Athey et al., 2019; Athey & Imbens, 2016; Mullainathan & Spiess, 2017; Varian, 2014; Wager & Athey, 2018). Recent developments have meant that valid statistical inference can now be made from these methods in the context of estimating causal effects rather than prediction – the more usual setting for machine learning techniques.

---

[34] Using the R package grf by Tibshirani et al. (2020), available from the CRAN project using the following link: https://CRAN.R-project.org/package=grf.

Causal forests developed from random forests whose trees make splits based on different variables (e.g., male/female, aged under 40/aged over 40) within a dataset to enable the prediction of an individual's outcome (e.g., will they default on a bank loan) based on their characteristics (Breiman, 2001). In essence it is an alternative to nearest neighbour matching in that each decision node (or leaf) defines the set of nearest neighbours for a given observation. Instead of choosing the k closest points to an observation based on distance, close points are defined as those that occupy the same leaf (Wager & Athey, 2018). The resulting forests are improvements on nearest neighbour matching in terms of bias and variance.

When used for prediction, the regression tree algorithm makes splits that optimise performance relative to some metric – commonly, minimising the mean squared error (MSE). A problem arises when it comes to causal inference. In any given causal tree, the MSE cannot be used as one never observes both $Y_i(0)$ and $Y_i(1)$ for every individual. This is different to prediction when the actual Y that the algorithm must predict is known for the training sample – in causal inference we do not have the counterfactual. Instead, developing work by others, Athey & Imbens (2016) minimises the *expected* MSE (EMSE) of predicted treatment and, further, show that this is equivalent to splitting based on the characteristics that yield the biggest differences in treatment effect plus a penalty parameter for within-node (or leaf) variation.

Causal Forests are non-parametric. This is appealing as, unless the underlying data generation process happens to be linear, the linear model may fail to identify the true effect by making assumptions about the true functional form. Finally, Causal Forests are estimated 'honestly'. Honest forests attempt to minimise the risk of spuriously identifying effects by only using any given data sample for either estimating treatment effects or for where to make splits. Not both.

In using Causal Forests, I follow the application outlined in Athey & Wager (2019). I grow initial regression forests of 2000 trees to provide out-of-bag predictions of the propensity score

(of treatment) and the main effect. Doing this enables the causal forest to focus on those features identified as most important in these initial forests rather than wasting splits on variables that are unimportant for heterogeneity. These values are inputs into the Causal Forest, which, when grown, has 10,000 trees. The Average Treatment Effect (ATE), Average Treatment Effect on the Treated (ATT), and the Average Treatment Effect on the Control (ATC) can be estimated using the Causal Forest. Conditional Average Treatment Effects (CATE) can also be estimated for each individual.

Arguably the greatest value of Causal Forests is in identifying heterogeneity in treatment effects. As Davis & Heller (2020) point out, testing for heterogeneity generally involves interacting treatment with various covariates where each additional hypothesis test may spuriously identify effects. If several interactions at the same time, or nonlinear functions of covariates are important, then traditional approaches may miss heterogeneity. Causal Forests flexibly and systematically model heterogeneity "based on high-dimensional nonlinear functions of observables" (Ibid., p665). I test for heterogeneity by comparing the average treatment effect for those individuals above the median treatment effect to those below. If this difference is statistically different from zero, there is evidence of heterogeneity. Treatment effect heterogeneity can also be charted.

The previous work on EMA has employed methods that, as in this paper, are not based on quasi-experimental methods. Dearden et al (2009), for example, employ fully interacted ordinary least squares (OLS) and probit techniques alongside PSM (as well as two difference-in-difference based sensitivity checks). In this context "fully interacted" means interactions of their treatment of being in an EMA pilot area with all other control variables. In contrast to this earlier work, I can flexibly and systematically analyse heterogeneity using Causal Forests with a much-reduced possibility of identifying sources of differential effects spuriously.

2.5 Results

2.5.1 Average Effects

Table 2.3 presents the first set of results; these relate to educational outcomes. Column (1) gives a simple OLS regression of the outcome on receipt of EMA (referred to as Gets EMA in the tables). Column (2) includes the full list of covariates, column (3) does the same but adds a school fixed effect to judge stability of results to past schooling. Column (4) then gives the average treatment effect (ATE) from an IPWRA specification, and Column (5) gives the average treatment effect on the treated (ATT) from the same specification.

In Table 2.3 two outcomes display robust and statistically significant effects – retention and university attendance. In the case of retention, once controls are added, this effect is around 8 percentage points. This is a little higher than in the pilot schemes where Dearden et al (2009) report retention effects of around 6.7 percentage points, though in statistical terms the estimate is not different to theirs. For university attendance the impact is similarly large; results appear a little less stable but are statistically the same across columns. There does not seem to be any impact on attainment which is always statistically insignificant. Some of this appears to be precision – as the point estimates are in fact large – in the case of the IPWRA estimates, they are around 8 percent of a standard deviation higher for those in receipt of EMA conditional on attendance. The final outcome in Table 2.3 is STEM degree subject. The IPWRA estimates differ from the others – in some sense they may be more credible – as IPWRA models both treatment and outcome where OLS models only outcome.

Moving to Table 2.4, which shows risky behaviours, after the simple specifications there are no estimates that are statistically different from zero. This is positive news for EMA in some sense - the fact that giving young people large sums of money does not seem to increase alcohol consumption or the likelihood of them having tried cannabis are encouraging findings. This implies relatively little in the way of negative externalities resulting from EMA.

Table 2.3   **Impact of EMA Receipt on Educational Outcomes**

	(1) OLS SIMPLE	(2) OLS FULL	(3) OLS FULL + School FE	(4) IPWRA ATE	(5) IPWRA ATT
Panel A - Retention in FTE/Training					
Gets EMA	0.141***	0.086***	0.088***	0.084***	0.079***
	(0.019)	(0.019)	(0.021)	(0.019)	(0.020)
N	3,722	3,722	3,692	3,722	3,722
R-squared	0.019	0.087	0.257		
Panel B - UCAS Point Score (Standardised)					
Gets EMA	0.021	0.072	0.061	0.081	0.085
	(0.055)	(0.048)	(0.055)	(0.050)	(0.053)
N	2,260	2,260	2,179	2,260	2,260
R-squared	0.000	0.353	0.551		
Panel C - University Attendance					
Gets EMA	0.170***	0.090***	0.077***	0.092***	0.089***
	(0.018)	(0.016)	(0.018)	(0.016)	(0.018)
N	3,714	3,714	3,692	3,714	3,714
R-squared	0.030	0.332	0.463		
Panel D - STEM Degree Subject					
Gets EMA	0.031	0.042	-0.006	0.063**	0.065*
	(0.033)	(0.033)	(0.041)	(0.032)	(0.035)
N	1,591	1,591	1,466	1,591	1,591
R-squared	0.001	0.065	0.408		

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, p<0.1. Column (1) includes no covariates, each other column uses the full list provided in Table 2. IPWRA specifications use the full list of covariates in both stages.

Table 2.5 shows labour market outcomes at age 25. Broadly, results are not forthcoming. The one exception to this is interesting, however. Whilst earnings, likelihood of ever having been employed, and hours worked are not impacted by getting EMA, the likelihood of being on an insecure "zero hours" employment contract is reduced by a statistically significant 4 percentage points, and this effect is stable across columns. Zero hours contracts are preferred by those whose conventional labour market opportunities are limited. The novel suggestion here is that EMA generates better conventional labour outcomes and thus reduces the probability of choosing zero hours employment.[35]

---

[35] I also examined mental health, as measured by GHQ score. This is measured at both age 17, when EMA is first received, and at age 25. There were no significant effects on this outcome.

Table 2.4   Impact of EMA on Risky Behaviours

	(1) OLS SIMPLE	(2) OLS FULL	(3) OLS FULL + School FE	(4) IPWRA ATE	(5) IPWRA ATT
Panel A - Frequency of Drinking Alcohol (Standardised)					
Gets EMA	-0.157***	-0.057	0.004	-0.062	-0.062
	(0.038)	(0.041)	(0.046)	(0.041)	(0.044)
N	2,954	2,954	2,885	2,954	2,954
R-squared	0.006	0.072	0.271		
Panel B - Ever Tried Cannabis					
Gets EMA	-0.054***	0.002	0.012	-0.007	-0.010
	(0.018)	(0.018)	(0.019)	(0.018)	(0.019)
N	4,101	4,101	4,083	4,101	4,101
R-squared	0.003	0.106	0.281		

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Column (1) includes no covariates, each other column uses the full list provided in Table 2. IPWRA specifications use the full list of covariates in both stages.

Table 2.5   Impact of EMA on Long-Term Labour Market Outcomes

	(1) OLS SIMPLE	(2) OLS FULL	(3) OLS FULL + School FE	(4) IPWRA ATE	(5) IPWRA ATT
Panel A - Log Earnings					
Gets EMA	0.015	0.017	0.056	0.004	-0.013
	(0.040)	(0.035)	(0.047)	(0.033)	(0.034)
N	1,681	1,681	1,565	1,681	1,681
R-squared	0.000	0.166	0.501		
Panel B - Ever Employed					
Gets EMA	-0.042***	-0.019	-0.012	-0.021	-0.024
	(0.012)	(0.013)	(0.015)	(0.014)	(0.016)
N	2,357	2,357	2,294	2,357	2,357
R-squared	0.006	0.121	0.401		
Panel C - Hours Worked					
Gets EMA	-1.393*	-0.977	-0.555	-1.087*	-0.991
	(0.733)	(0.711)	(0.836)	(0.659)	(0.658)
N	1,924	1,924	1,826	1,924	1,924
R-squared	0.003	0.142	0.435		
Panel D - Zero Hours Contract					
Gets EMA	-0.035**	-0.044**	-0.038**	-0.044**	-0.038**
	(0.018)	(0.017)	(0.018)	(0.017)	(0.019)
N	1,926	1,926	1,827	1,926	1,926
R-squared	0.004	0.068	0.387		

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Column (1) includes no covariates, each other column uses the full list provided in Table 2. IPWRA specifications use the full list of covariates in both stages.

Before moving on to systematically examine heterogeneity using causal forests we examine, as Dearden et al (2009) do, heterogeneity by gender. This is shown in Table 2.6. Significant differences are not generally forthcoming – but with two important exceptions. First, EMA has a large and highly significant effect on STEM subject choice for women but not for men. Second, the effect on hours of work seems to be entirely confined to men which is a remarkably large effect suggesting that EMA may have facilitated quite different employment for men that would have otherwise been the case.

Finally, when faced with a large number of outcomes, as arises here, it is natural to be concerned with the possibility of false discovery – that some results are significant simply by chance. Appendix Table A1 takes the IPWRA specifications from above (as they account for selection in both treatment as control unlike the other two methods) and adjusts the p-values for false discovery – using the method in Benjamini & Hochberg (1995). As is clear, almost every outcome that was significant before is still significant, the exception being hours worked. UCAS Points was marginally significant before (p=0.101) but the adjusted value is not (p=0.202). In fact, the significance level does not change on three of the results – retention, university attendance and working on a zero hours contract remain significant at the 1%, 1%, and 5% levels.

2.5.2 Causal Forests, Average Effects, and Treatment Heterogeneity

Table 2.7 reports the average treatment effect (ATE), the average treatment effect on the treated (ATT), and the average treatment effect on the control (ATC) for each outcome, using the full set of controls used in the IPWRA equations in Table 2.5 above. It shows broad agreement with the previous specifications – retention, university attendance, STEM degree subject, and the probability of being on a zero hours contract continue to show statistically significant effects. The effect sizes are very similar to those in earlier tables.

Table 2.6   Heterogeneity by Gender (IPWRA ATE Specifications)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A - Educational Outcomes								
	Retention		UCAS Points		University		STEM Subject Choice	
	Male	Female	Male	Female	Male	Female	Male	Female
Gets EMA	0.088***	0.094***	0.153**	0.081	0.100***	0.095***	0.007	0.145***
	(0.027)	(0.027)	(0.061)	(0.064)	(0.023)	(0.022)	(0.049)	(0.036)
N	1,798	1,924	1,019	1,241	1,784	1,930	676	915
Panel B - Risky Behaviours								
	Alcohol Consumption		Cannabis Ever		.	.	.	.
	Male	Female	Male	Female				
Gets EMA	-0.032	-0.086	0.001	-0.010	.	.	.	.
	(0.056)	(0.060)	(0.028)	(0.024)	.	.	.	.
N	1,244	1,254	2,009	2,092	.	.	.	.
Panel C - Labour Market Outcomes (Age 25)								
	Log Earnings		Ever Employed		Hours Worked		Zero Hours Contract	
	Male	Female	Male	Female	Male	Female	Male	Female
Gets EMA	-0.045	0.065	-0.022	-0.012	-3.162***	0.728	-0.028	-0.019
	(0.040)	(0.049)	(0.016)	(0.020)	(0.873)	(0.978)	(0.020)	(0.019)
N	733	948	1,030	1,327	738	876	733	879

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2.7   Average Estimates from Causal Forest Estimations

	(1) ATE	(2) ATT	(3) ATC	(4) Heterogeneity	(5) 95% CI
Retention	0.089***	0.086***	0.096***	0.033	+/- 0.087
	(0.019)	(0.019)	(0.022)		
UCAS	0.075	0.076	0.072	0.051	+/-0.199
	(0.051)	(0.054)	(0.048)		
University	0.092***	0.089***	0.098***	-0.039	+/- 0.059
	(0.015)	(0.016)	(0.016)		
STEM	0.064*	0.076**	0.030	0.058	+/- 0.137
	(0.034)	(0.035)	(0.034)		
Alcohol	-0.098**	-0.095**	-0.102**	-1.081	+/- 0.157
	(0.040)	(0.040)	(0.046)		
Cannabis	-0.017	-0.017	-0.016	0.077	+/- 0.074[++]
	(0.018)	(0.018)	(0.021)		
Earnings	0.022	0.020	0.027	0.072	+/- 0.143
	(0.035)	(0.035)	(0.038)		
Ever Employed	-0.028**	-0.030**	-0.024**	0.058	+/- 0.072
	(0.013)	(0.015)	(0.011)		
Hours	-0.623	-0.496	-0.839	-0.839	+/- 1.252
	(0.706)	(0.711)	(0.761)		
Zero Hours	-0.044**	-0.045**	-0.043**	-0.058	+/-0.073
	(0.019)	(0.020)	(0.017)		

Cluster robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Sample sizes are as in Table 6 for each outcome. Column (5) shows an heuristic test for heterogeneity – whether the difference between the ATE above and below the median is different from zero. + symbols denote significant differences. Included covariates are those in Table 2.

However, two new effects are identified – one on alcohol consumption, and another on the likelihood of ever having been employed. In the case of the latter, it appears that the estimates are similarly precise, but the coefficients are a little larger meaning that effects are significant, where before they were not. In the case of alcohol consumption, however, the effect sizes are themselves larger. Table 2.7 also includes a measure of heterogeneity – this takes the average individual treatment effect above and below the median treatment effect and tests the difference between them. The confidence interval is provided. By this heuristic, it is only cannabis consumption appears to have substantial amounts of heterogeneity with respect to EMA receipt.

Table 2.8 reports the variables most important for heterogeneity in each outcome. This measure essentially counts how frequently the causal forest algorithm makes splits using a given variable and weights that by the depth at which these splits are made. Earlier splits (higher up the tree) imply that treatment effect heterogeneity was maximised best by choosing that variable at that point, so shallower splits should be given more weight. The most important dimensions of heterogeneity are identified in a separate regression forest estimated before the causal forest is estimated. An initial regression forest helps to calibrate the causal forest and ensure that splits are not being made that are unimportant for heterogeneity. Athey and Wager (2020)'s application prioritises those variables that are more important than the average variable when it comes to making the splits. Table 2.8 reports the top five variables judged by this metric. Some variables have fewer than five; some variables appear frequently – the IMD and quartiles of KS2 and GCSE score for example, others appear less often. The latter is true for gender, region of residence, and parental education.

Below, Figures 2.3 to 2.5 show the overall distribution of conditional average treatment effects for each outcome, arranged in the same groups as the initial OLS/IPWRA tables above. Essentially, these are frequency plots of individual level treatment effects. Figure 3 gives the educational outcomes. In the case of retention, a small number of individuals experience negative (though close to zero) impacts of EMA, whilst the modal bin is between 0 and 5 percentage points. For some, effects are as large as 30 percentage points. UCAS score displays a stranger distribution with peaks at different points, with the minimum effect being clustered around zero, but the maximum being around 0.1 of a standard deviation. University attendance displays a similar pattern to retention in that most individuals have effects close to zero or around ten percentage points but a small number see much larger impacts. STEM degree subject choice, like retention, most resembles normality.

Table 2.8   Most Important Variables for Heterogeneity as Identified by Causal Forests

Outcome	Most Important	2nd	3rd	4th	5th
			Rank		
Retention	LA-EMA	GCSE	Dependent Children	Non-white	.
UCAS points	KS2	GCSE	.	.	.
University	GCSE	KS2	Non-white	.	.
STEM	LA-EMA	IMD	Main Parent Age	Region	GCSE
Alcohol	IMD	Gender	Non-white	.	.
Cannabis	GCSE	General Health	.	.	.
Earnings	IMD	GCSE	KS2	Gender	.
Ever Employed	GCSE	IMD	LA-EMA	Main Parent Age	KS2
Hours	IMD	KS2	GCSE	Gender	.
Zero Hours	LA-EMA	IMD3	Parental Education	GCSE	KS2

Note: Table displays the most important dimensions of heterogeneity for each outcome. Variable importance is determined by weighting the number of splits across tress in the Causal Forest by the depth at which those splits occur. A greater number of splits that occur by a particular variable, the more important it is for heterogeneity. Those with a variable importance about the mean are included. LA-EMA is the overall level of EMA enrolment in an LA in Wave 4, GCSE is quartiles of GCSE attainment, KS2 is quartiles of Key Stage 2 attainment, IMD is the Index of Multiple Deprivation.

For risky behaviours (Figure 2.4), alcohol consumption (Panel A) is fairly concentrated around negative eight percentage points. Cannabis consumption is different – it seems to have few who experience a zero effect of EMA and instead some who experience positive and some who experience negative effects, with the balance tilting towards more individuals experiencing negative effects. In Figure 2.5 labour market outcomes are displayed. The conditional average treatment effects of EMA on log earnings (Panel A) have a broad distribution. Though similar to STEM degree subject in range (from around -0.05 to 0.10) the frequencies at any given point are less concentrated. The opposite is true of the likelihood of ever being employed (Panel B) which is highly concentrated and seems to present very little heterogeneity, though there is a long tail of larger negative effects. Hours worked looks reasonably heterogeneous spanning a negative impact of three fewer hours worked at age 25 for somebody on EMA to an additional

two hours. Finally, zero hours contracts display a similar sort of distribution to cannabis consumption but are exclusively negative.

Figure 2.3   Distribution of Conditional ATE on Retention, UCAS Score, University Attendance, and STEM Degree Subject Choice
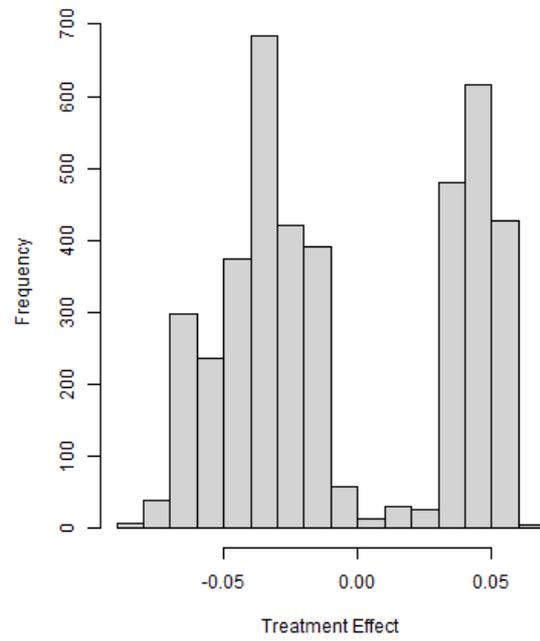

(A) Retention

(B) UCAS Score

(C) University

(D) STEM Degree
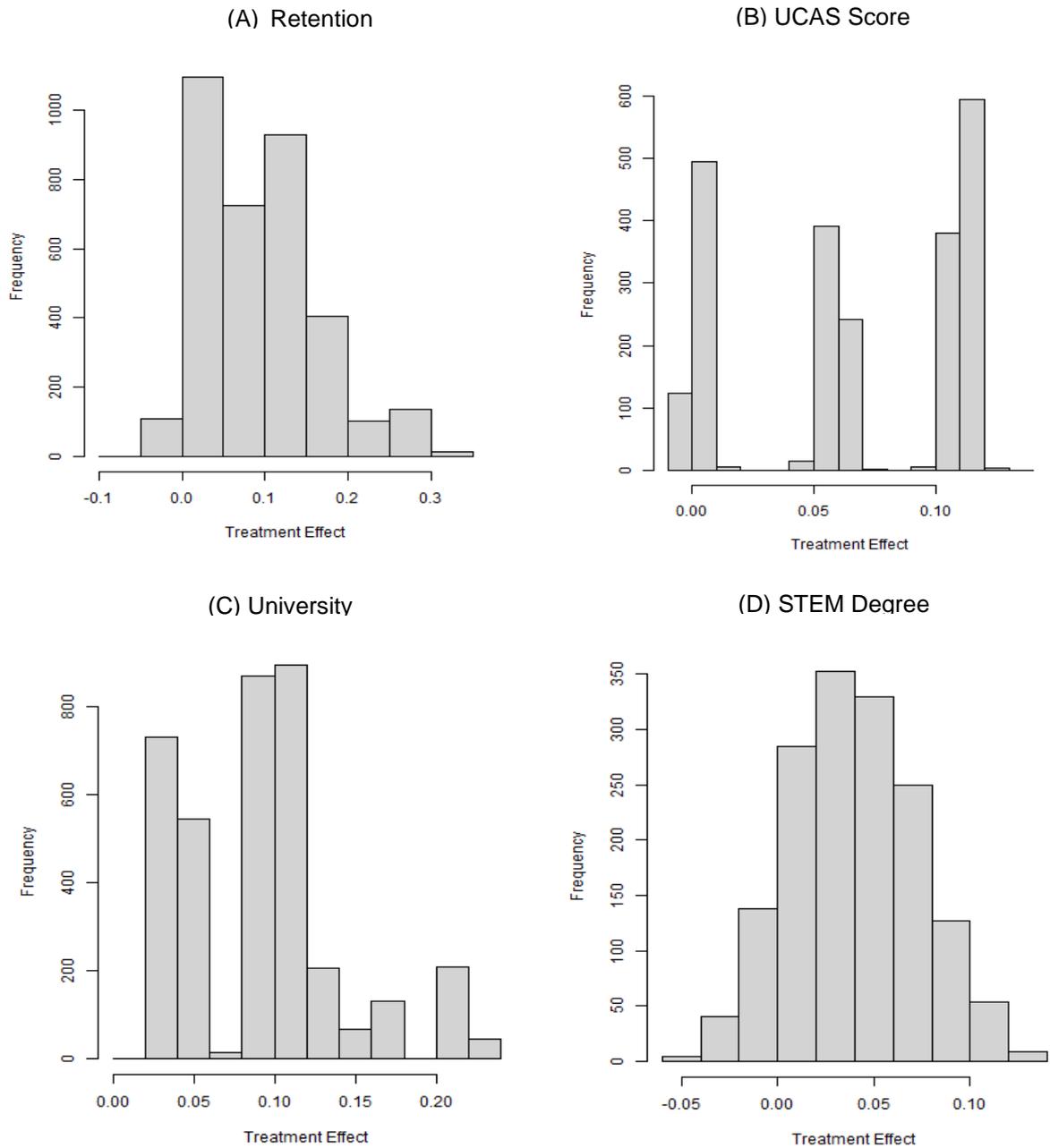
Figure 2.4   Distribution of Conditional ATE on Alcohol and Cannabis Consumption
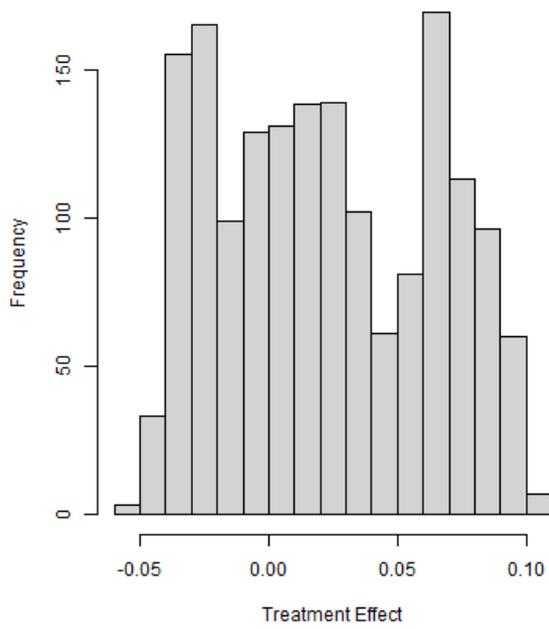
(A) Alcohol Consumption
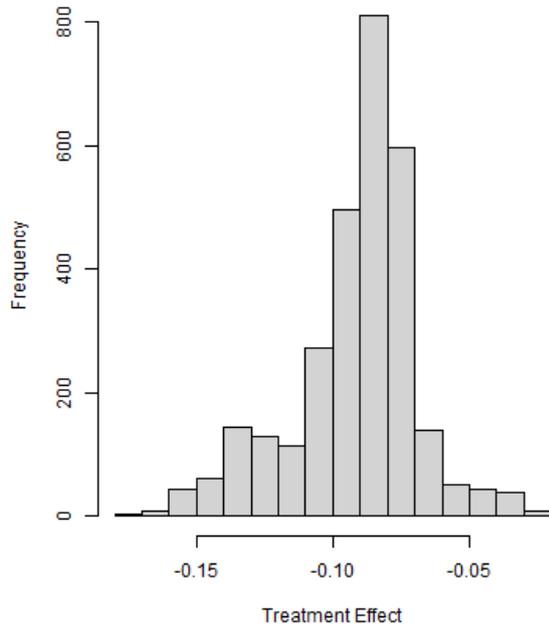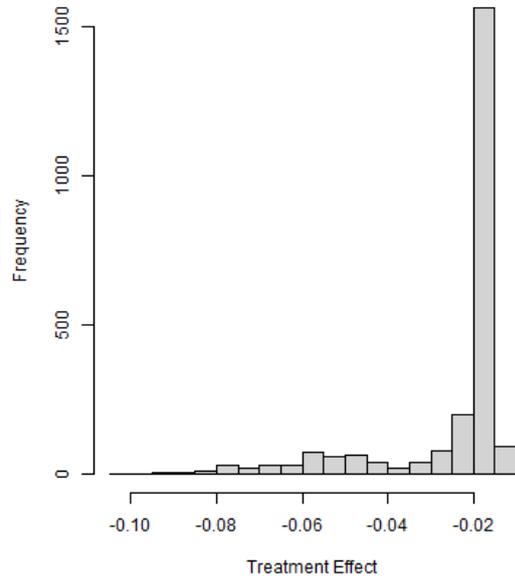
(B) Cannabis Consumption

Figure 2.5   Distribution of Conditional ATE on Log Earnings, Employed Probability, Hours Worked, and Zero Hours Contract Probability
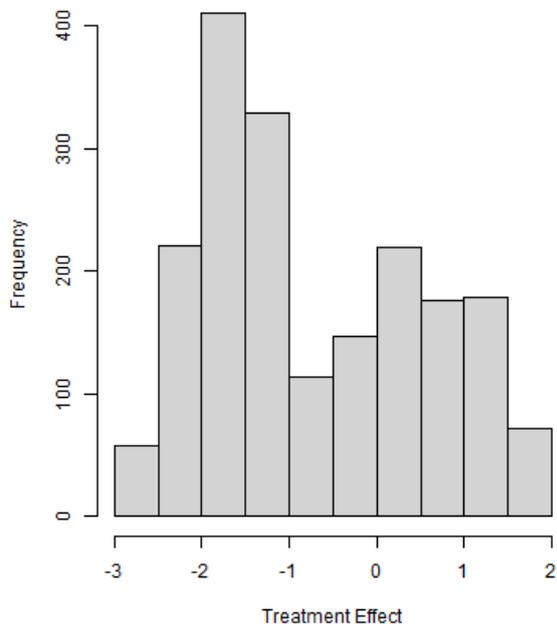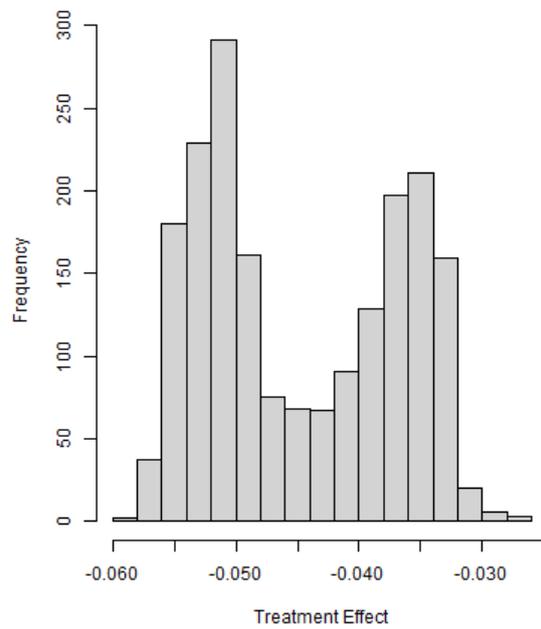

(A) Log Earnings

(B) Ever Employed

(C) Hours Worked

(D) Zero Hours

Figures 2.6 to 2.11 chart the two most important dimensions of heterogeneity for a selection of outcomes; these dimensions come from the 1st and 2nd most important variables listed in Table 2.8. The equivalent charts for the remaining outcomes (excluded from the main body for ease of presentation) are given in the Appendix. The charts are a mix of scatter and box-and-whisker plots that plot the treatment effect on the y axis and the characteristic by which it is varying on the x axis. There are some interesting results. For retention (Figure 2.6), EMA's effect seems largest for those in the lowest quartile of prior GCSE attainment (Panel B); higher proportions of people in your LA also being in receipt of EMA is important according to Table 2.8, but this is less clear when charted. Prior attainment at KS2 and GCSE are important dimensions for heterogeneity in the impact on attainment at A-level (Figure 2.7). Those who performed better in the past see smaller improvements as a result of EMA. In Figure 2.8, university attendance seems to be characterised mostly by greater variance in the effects of EMA by prior attainment rather than seeing the kind of pictures as for retention and A-level attainment. This seems to be true in Figure 2.9, too, where the effect of STEM degree subject seems to have lower variance as the proportion of those on EMA in your LA and the deprivation of your postcode rise.

In terms of risky behaviours, Figures B1 and B2 are interesting. For alcohol consumption (Figure B1) it seems that individuals partake less frequently when on EMA if they live in more deprived areas (as measured by IMD). Men and women in the second Panel of Figure B1 has similar averages, but for females there seems to be a greater chance of greater reductions; in essence, there is greater variability. For cannabis, the outcome where heterogeneity was identified by the heuristic in Table 2.7, there seems to be substantial variation. Interestingly, those who self-report higher general health seems to be more likely to take cannabis that those with lower self-reported health if they are in receipt of EMA. This is interesting because one may not have thought to examine this dimension of heterogeneity, but the causal forests show that it is important.

Figure 2.6   Two Most Important Dimensions of Heterogeneity for Retention

(A) EMA Enrolment in LA   (B) Quartile of GCSE Score
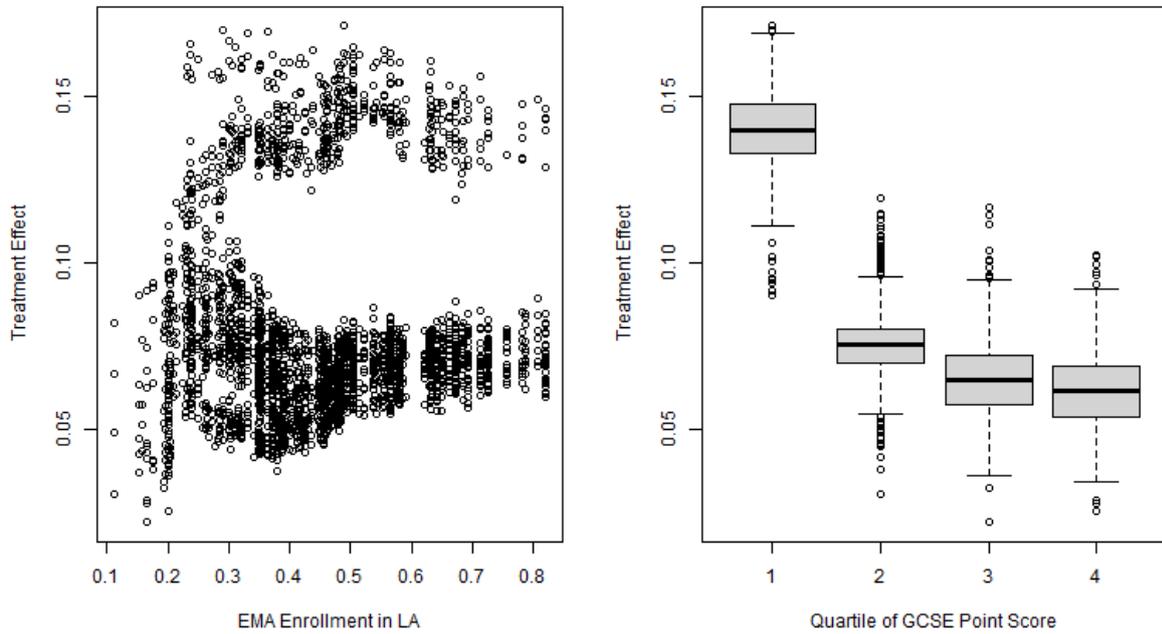
Figure 2.7   Two Most Important Dimensions of Heterogeneity for UCAS Point Score

(A) Quartile of KS2 Score   (B) Quartile of GCSE Score
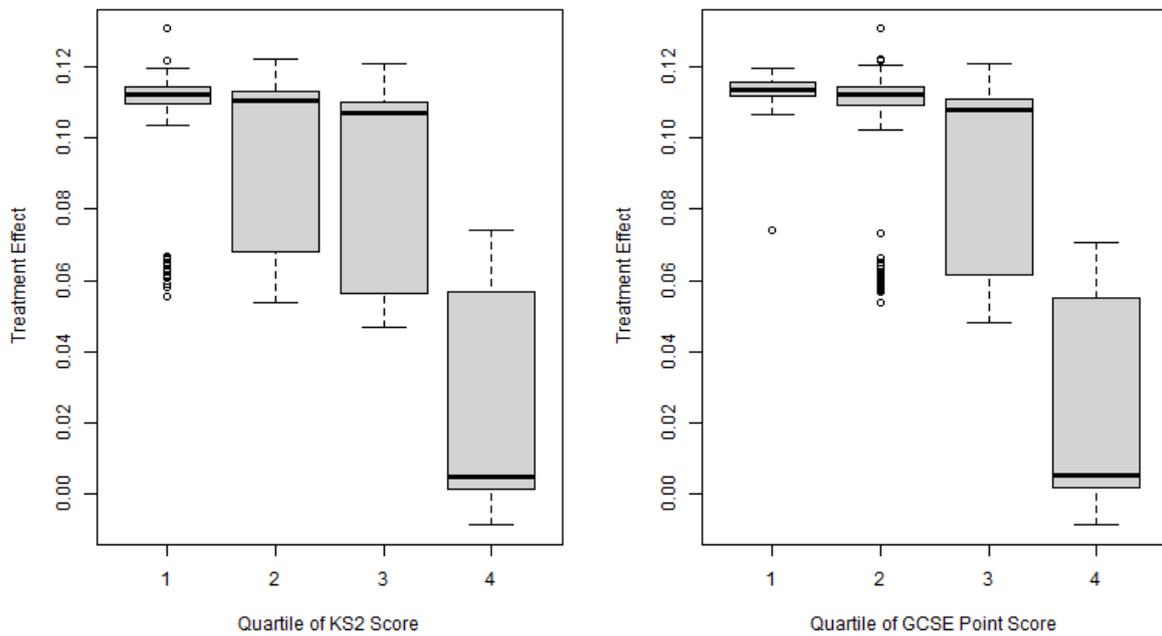
Figure 2.8   Two Most Important Dimensions of Heterogeneity for University Attendance

(A) Quartile of GCSE Score   (B) Quartile of KS2 Score
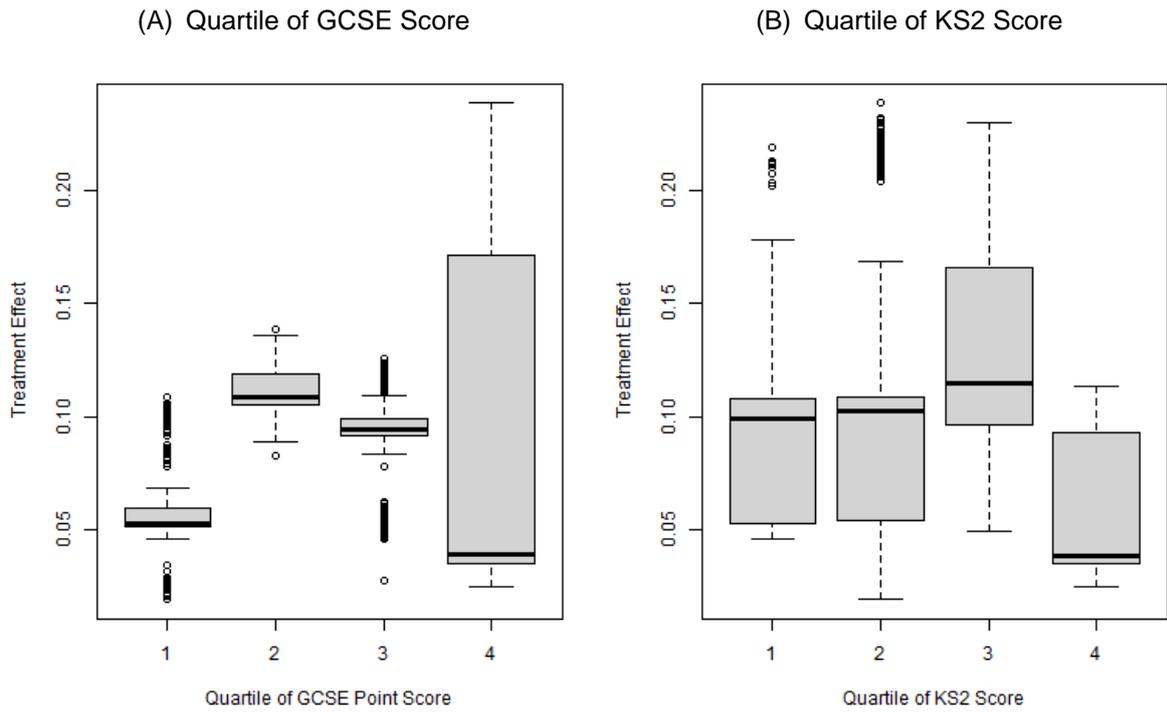
Figure 2.9   Two Most Important Dimensions of Heterogeneity for Stem Degree Subject

(A) EMA Enrolment in LA   (B) IMD
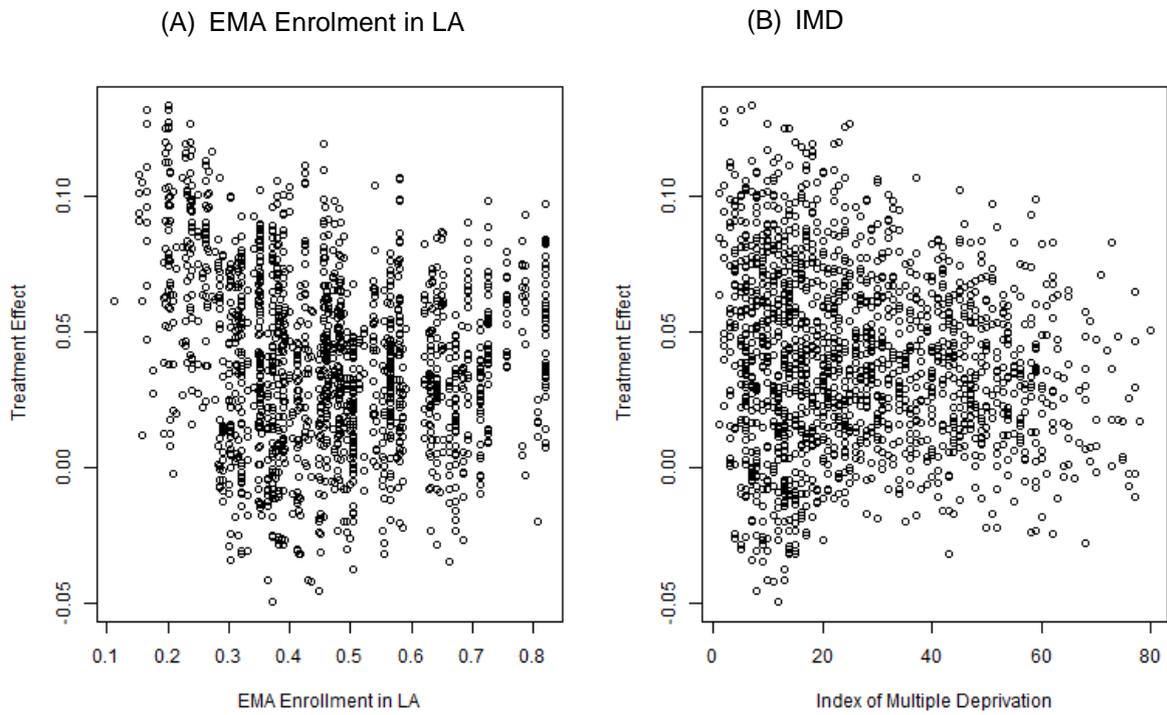
Figure 2.10   Two Most Important Dimensions of Heterogeneity for Ever Being Employed by Age 25
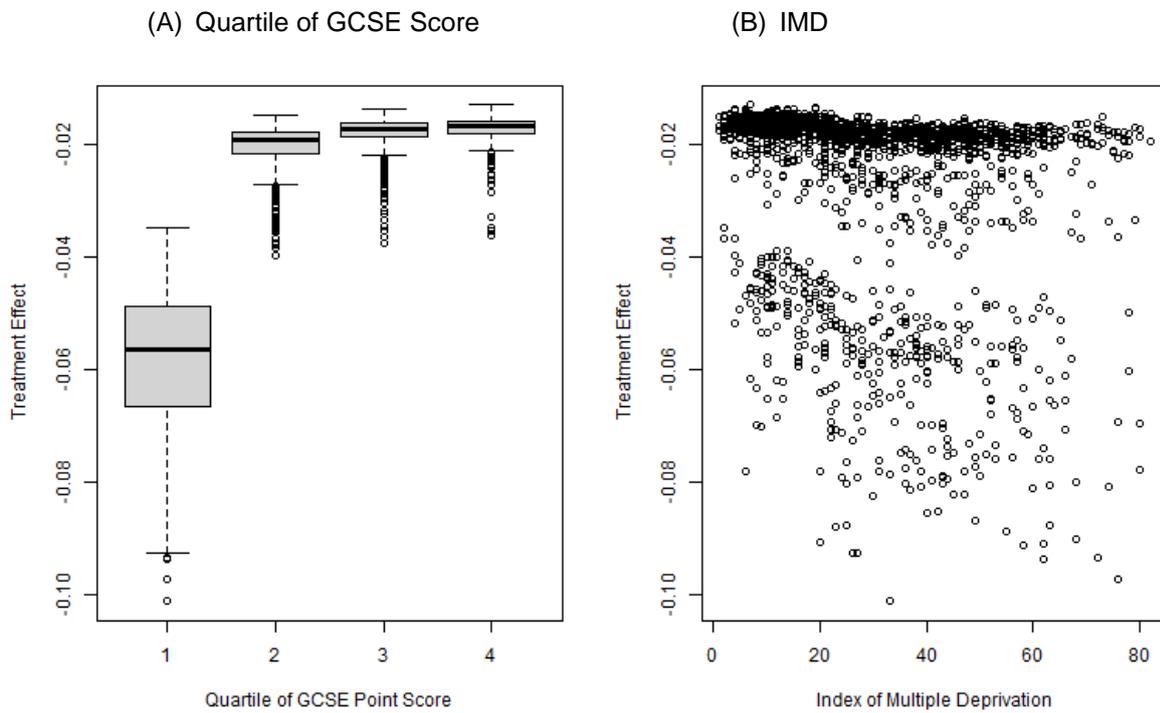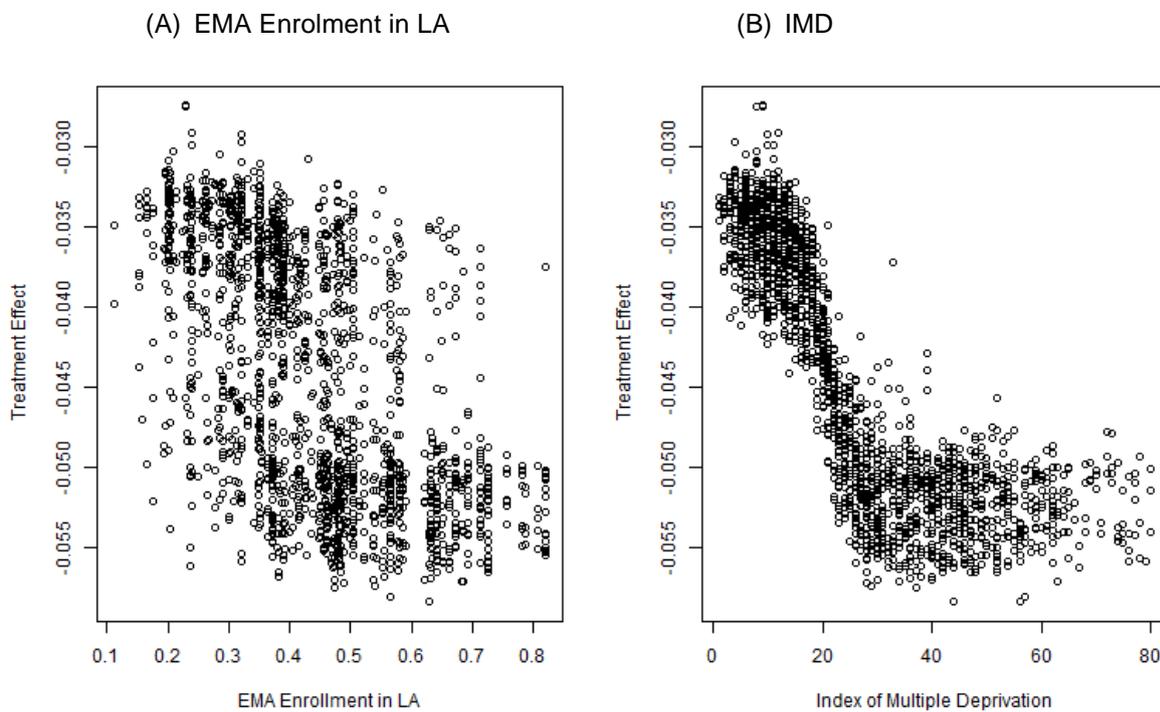
(A) Quartile of GCSE Score   (B) IMD

Figure 2.11   Two Most Important Dimensions of Heterogeneity for Being on a Zero Hours Contract at Age 25

(A) EMA Enrolment in LA   (B) IMD

Finally, in terms of labour market outcomes – there appear to be greater earnings benefits (Figure B3) for those in more deprived areas if they are on EMA. For some in less deprived areas the effects are actually negative. The likelihood of ever having had a job by age 25 (Figure 2.10) is roughly zero for the top three quartiles of prior GCSE attainment, but for the lowest EMA seems to mean a negative likelihood of being employed. This is interesting – and perhaps relates to EMA incentivising some people into career paths they were not best suited for. It seems being in a more deprived area means you see slightly longer working hours (Figure B4) – though this picture is less than clear. The effects seem to most reliably negative for those with higher attainment at KS2. Lastly, having more people in your LA on EMA and living in a more deprived postcode seem to make it more likely EMA has a more negative impact on the probability of working on a zero hours contract at age 25 (Figure 2.11). This is seen more clearly and easily in the case of the IMD (the right-hand panel of Figure 2.11).

Taken together, though it was initially indicated that there was not substantial heterogeneity in the effects of EMA by the heuristic in Table 2.7, it appears that some dimensions exist where there are differences in effects. This may be because in a number of cases it is more the case that some groups see greater variance in the effects rather than different averages. Some dimensions are clear – and few are starker than the large negative effect on the likelihood of working on a zero hours contract at age 25 (Figure 2.11) for those in more deprived areas.

2.5.3   Amount of EMA received

Up to this point, EMA has been treated as a binary treatment. This is because most recipients receive the full amount – this is shown in Table 2.9. Now, though, I examine the results by varying amounts of EMA receipt. Figure 2.12 shows the overlap in the whole sample by amount of EMA. As in Figure 2.1, the overlap is good.

Table 2.9   Distribution of Amounts of EMA Received

EMA Payment Amount (£)	Frequency	Percentage
0	1,168	31.38
10	244	6.56
20	298	8.01
30	2,012	54.06
Total	3,722	100

Figure 2.12   Overlap by EMA Amounts
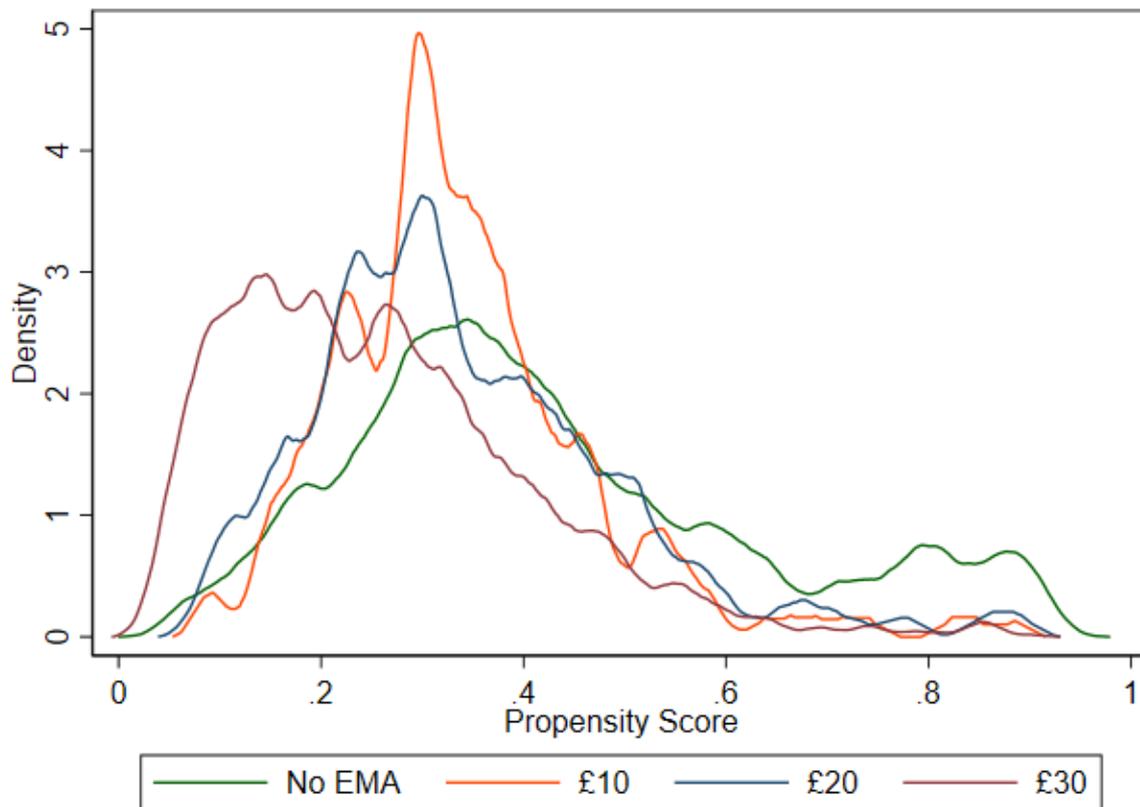
Table 2.10 takes the IPWRA average treatment effect specifications from above, but the treatment estimates are for the £10, £20, and £30 amounts instead of the binary EMA treatment variable. Two rows are blank – UCAS Points and STEM degree subject. In the case of the former, the overlap was not as good as in Figure 12 (though for all other outcomes, the overlap is very similar to that in Figure 2.12). For STEM degree subject the IPWRA command yielded no output due to issues of concavity. OLS estimates, of the same specification as the "OLS

Full" specifications above, for these two outcomes (not shown) showed insignificant effects for all EMA amounts for UCAS points whilst only the £10 and £20 brackets were insignificant for STEM degree subject. The £30 bracket was associated with a 6.8 percentage point increase in the likelihood of an individual picking a STEM degree subject at university, significant at the 10 percent level.

In terms of the coefficients in Table 2.10 the effects vary by outcome. For retention all levels have significant effects. The effect of £10 is larger in magnitude than the others but not statistically different to them. University attendance displays positive and significant effects of the £20 and £30 brackets. Alcohol and cannabis consumption each show negative effects of the £10 and £20 brackets whilst the probability of ever being employed seems only to be significantly impacted for those on the £10 bracket. Finally, hours worked and the likelihood of being on a zero hours contract are only (negatively) impacted for those on the £30 bracket.

This suggests that there is some evidence of heterogeneity in effect by amount received for risky behaviours and later employment. Hours worked and likelihood of being on a zero hours contract are perhaps easiest to understand – the larger amounts of money (given to those with the lowest family incomes) improve outcomes where smaller amounts do not. The other cases – alcohol and cannabis consumption and the likelihood over ever being employed by age 25 – are trickier. In the case of cannabis consumption, it is possible that there is a diminishing desirable (i.e., negative) effect as the amount of money handed out increases – on £10 or £20 the money is used to get the young person to college or to buy resources to learn with. Beyond that some may increase consumption with the additional money. The same reasoning does not apply to alcohol, where, as seen in the heterogeneity charts above, the impacts are either zero or negative, with little evidence of alcohol consumption increasing for any group.

Table 2.10 - IPWRA ATE Estimates for EMA Amounts

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcomes	£10	Std. Error	£20	Std. Error	£30	Std. Error	Observations
Retention	0.132***	(0.042)	0.097**	(0.039)	0.081***	(0.021)	3,722
UCAS							
Points	.	.	.	.	.	.	.
University	0.035	(0.034)	0.113***	(0.040)	0.094***	(0.017)	3,712
Stem	.	.	.	.	.	.	.
Alcohol	-0.125*	(0.075)	-0.132*	(0.076)	-0.009	(0.047)	2,497
Cannabis	-0.092***	(0.036)	-0.109***	(0.037)	0.008	(0.020)	4,098
Log Pay	0.052	(0.056)	0.017	(0.041)	0.016	(0.036)	1,416
Ever							
Employed	-0.204***	(0.030)	0.013	(0.026)	-0.020	(0.014)	2,355
Hours	-0.537	(0.928)	-0.268	(0.796)	-1.814**	(0.745)	1,612
Zero							
Hours	0.027	(0.026)	0.004	(0.026)	-0.034**	(0.014)	1,610

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
Note: UCAS Points, unlike the other outcomes displayed poor overlap when it came to the amounts and so is omitted. STEM subject choice is omitted as the IPWRA estimation did not produce any output due to issues of concavity. Sample sizes vary when compared to above as the few cases where no overlap existed were omitted, too.

2.6   Discussion and Conclusion

This paper uses a variety of methods to show robust positive impacts of receiving EMA on retention in full-time education or training and university attendance, and robust negative impacts on the likelihood of being on an insecure (zero hours) employment contract. There is a suggestion, too, that the likelihood of studying a STEM degree subject at university also increases. No other outcomes, including educational attainment, risky behaviours (frequency of alcohol consumption and ever having tried cannabis), and long-term labour market outcomes, are significantly impacted.

The positive impact on retention mirrors that identified in the pilot studies on EMA; indeed, it is slightly larger (at around 8 percentage points instead of 6.7 percentage points). The effect identified in the pilot studies is, therefore, similar once the programme has been rolled out and has been in existence for around 5 years. The university attendance impact is positive too and

differs to previous estimates of the impact of EMA. The more novel element of this analysis is in showing that there is no measurable impact on risky behaviours, and that insecure work at age 25 appears to decline. Indeed, the only significant effects on risky behaviours identified (in the causal forest analysis) are negative. It may have been easy to imagine that giving young people large amounts of cash might lead to negative outcomes. That alcohol and cannabis consumption is no higher for those receiving EMA suggests that this drawback is not present, and by implication the associated spillover effects of such behaviours do not increase. In terms of insecure work, it is likely the case that increased time in education (through retention and university attendance) explains the occurrence of this. But it is noteworthy that other labour market outcomes do not improve whilst this one does.

Beyond average effects, I examine heterogeneity systematically using causal forests. Across the range of outcomes, it appears that there are a number of cases where the effect of EMA differs by characteristics. This is true for both retention and attainment as A level (measured by UCAS score) – both are most positive for those in the lowest quartiles of prior attainment; alcohol consumption, which seems most negatively impacted by EMA in more deprived areas; cannabis consumption, which seems reliably to be reduced by EMA for those with lower self-reported health but appears to be more likely to increase for those with high self-reported health (though there is a wide distribution); the likelihood of ever having been employed, which is most negative for those with low GCSE attainment; and the likelihood of being on a zero hours contract at age 25 which varies starkly by deprivation where the individual lives. Of these, it is only the likelihood of ever being employed being worse for lower achievers which could be considered a drawback of EMA and is in need of further investigation.

The debate around EMA is still relevant in policy circles, for example in the United States where many states have compulsory school leaving ages below the UK's 18. But even though the UK now mandates education, where it used to incentivise it through EMA, that does not

leave the matter settled. The absence of negative impacts on risky behaviours combined with desirable impacts on university attendance and on the probability of insecure work make a reappraisal of EMA as a policy seem compelling. As attention worldwide turns to recovering lost education during COVID-19 lockdowns, EMA seems less like the deadweight loss government described back in 2010/11 than an idea ripe for potential revival.

B   Appendix to Chapter Two

Appendix Table B.1 IPWRA Specifications with p-values Adjusted for False Discovery

Outcome	Estimate	Std Error	N	p-value	q-value
Educational					
Retention	0.084***	(0.019)	3,722	0.000	0.000
UCAS Points	0.081	(0.050)	2,259	0.101	0.202
University	0.092***	(0.016)	3,714	0.000	0.000
STEM	0.063*	(0.032)	1,591	0.056	0.080
Risky Behaviours					
Alcohol	-0.062	(0.041)	2,954	0.135	0.451
Cannabis	-0.007	(0.018)	4,101	0.689	>0.999
Labour Market					
Earnings	0.004	(0.033)	1,681	0.900	>0.999
Ever Employed	-0.021	(0.014)	2,357	0.116	0.290
Hours	-1.087	(0.659)	1,924	0.099	0.165
Zero Hours	-0.044**	(0.017)	1,926	0.010	0.013

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Asterisks relate to the q-value. The q-value is the p-value adjusted in line with Benjamini & Hochberg (1995).

Appendix Figure B.1 Two Most Important Dimensions of Heterogeneity for Alcohol Consumption

(A) IMD   (B) Gender
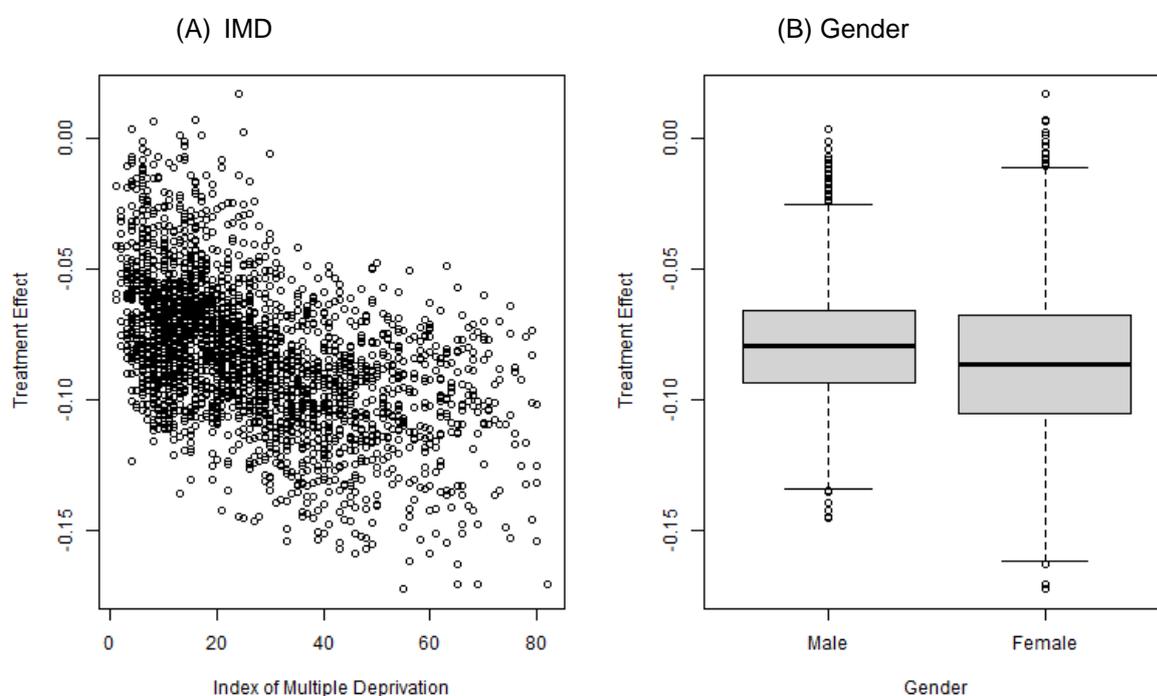
Appendix Figure B.2   **Two Most Important Dimensions of Heterogeneity for Cannabis Consumption**

(A) Quartile of GCSE Score   (B) Level of General Health
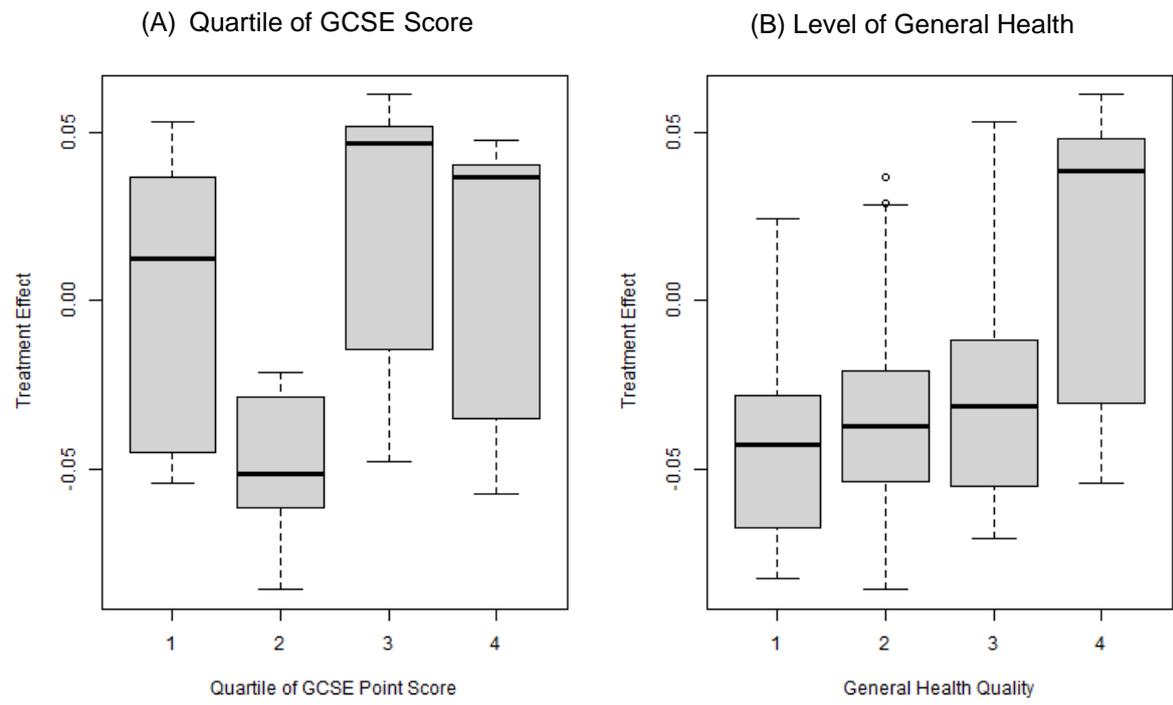
Appendix Figure B.3   **Two Most Important Dimensions of Heterogeneity for Log Earnings at Age 25**

(A) IMD   (B) Quartile of GCSE Score
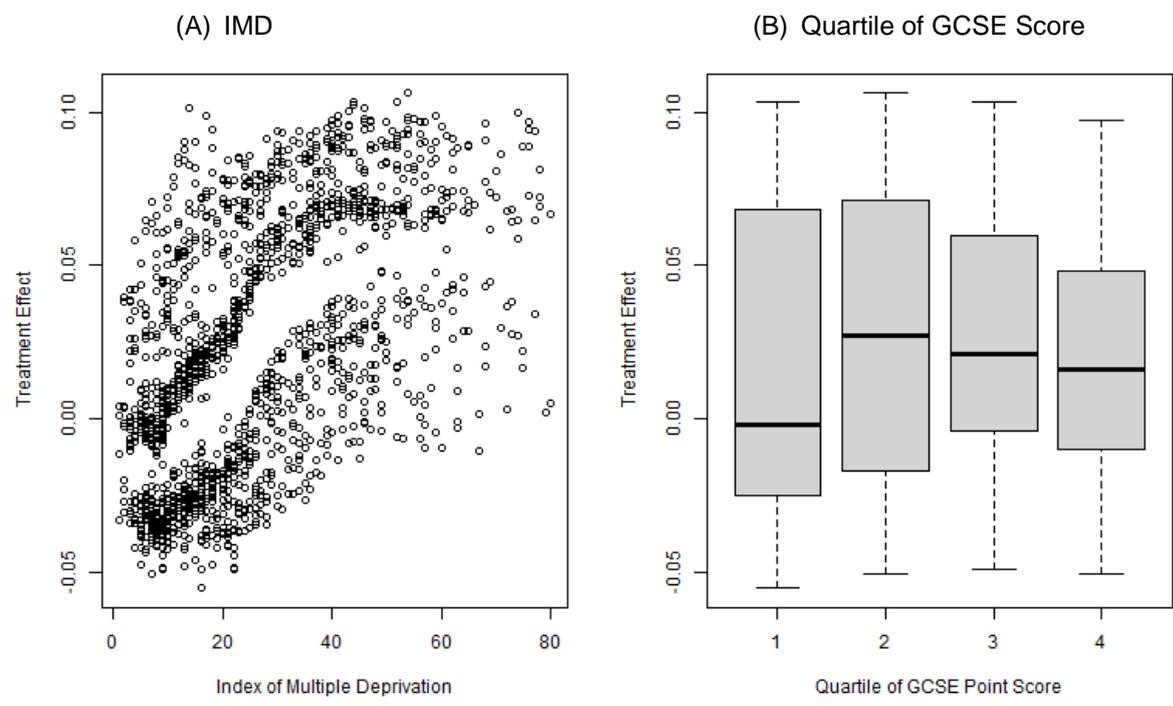
Appendix Figure B.4   Two Most Important Dimensions of Heterogeneity for Hours Worked at Age 25

(A) IMD

(B) Quartile of KS2 Score
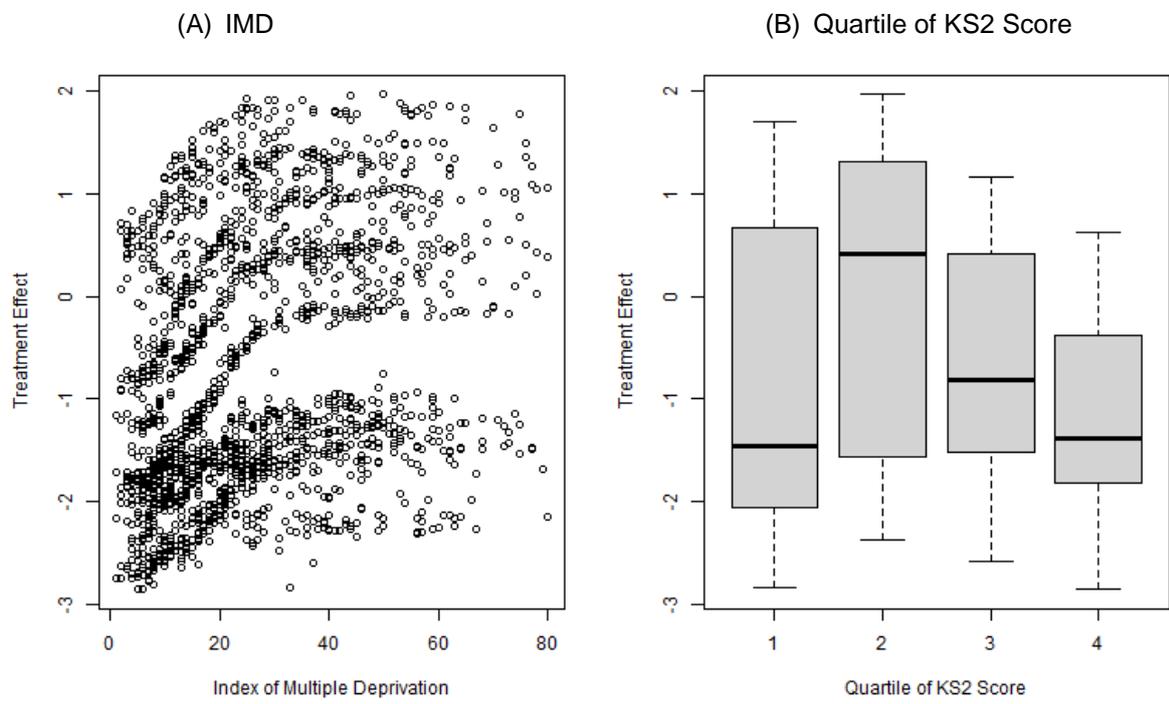
3 Chapter Three – Can Compulsory Schooling Build a Nation? The Causal Effect of Education on National Identity

3.1 Introduction

Identity is increasingly becoming a focus of economic inquiry due to its potential explanatory power across a wide range of outcomes of economic interest. How an individual identifies might impact how long they choose to study, the job they choose to apply for, and how happy their relationships are.[36] The study of identity covers, though is not limited to, gender, nationality, race, and religion. Such analysis can be broadly grouped into two categories – how identity is formed, and the effects of sentiments, feelings, and beliefs. In this paper, I focus on the formation of national identity.

Since the 2016 referendum on the United Kingdom's membership of the European Union, it has been increasingly common in the media, and in wider discourse, to implicitly link education with one's beliefs and views. A 2017 headline evidences this view – "Brexit caused by low levels of education, study finds" (Stone, 2017). The study in question, Zhang (2018), was not, in fact, causal – and did not claim to be – but university attendance was the best predictor of an individual's voting behaviour in the referendum.[37] The discussion around this invites a causal analysis of education's impact on identity. In broad agreement with Zhang (2018), Kunst et al. (2020), using changes to compulsory schooling laws across 4 European nations, examine how Euroscepticism is impacted by education, though they find little evidence of an effect. In contrast, this paper, though motivated by the link drawn in the context of the Brexit referendum between education and sentiments, focuses instead on British identity and is the first to exploit changes in compulsory schooling laws to examine the impact of education on national identity.

---

[36] For example, choice of what to study (Porter & Serra, 2020) , employment choice (Battu & Zenou (2010), and happiness in relationships (Bertrand et al., 2015).

[37] The timing discrepancy in the Zhang (2018) citation and the reporting of it are as a result of the online publication of the Zhang paper in August 2017.

Understanding identity – both its formation and effects – is of crucial importance to economists. An individual's identity is the lens through which they view the actions of others – how they may interact with economic agents and interpret the externalities created by them. Endogeneity clouds the empirical analysis of identity. A person may not identify as British and leave school early due to some unobserved characteristic – the result would be omitted variable bias. Similarly, not identifying as British may leave you wanting to spend less time in a British institution giving rise to the possibility of reverse causality. In this paper, I use the well-known raising of the school leaving age (RoSLA) reform, a compulsory schooling law change in the UK in 1972 that meant individuals could no longer leave school at 15 but instead had to wait until they were 16, as an instrument for number of years of education. As a result, the causal effect (or lack thereof) on national identity can be estimated.

The contribution of the paper is straightforward – to provide causal estimates of the impact of education on identity. This has been done for religious identity in Canada (Hungerman, 2014) and in cross country analysis of Europe (Mocan and Pogorelova, 2017) but not, to the best of my knowledge, for national identity. I examine two different national identity outcomes[38] – one (referred to as British identity) that takes value 1 if the individual identifies as British or 0 otherwise; and a second (referred to as *unionist* identity) that takes value 1 if the individual identifies as British and 0 if they identify as one of the constituent nations of the UK.[39,40] This distinction is important – one examines simply the willingness of individuals to report their identity as British regardless of the other identities they may possess and report. The second excludes those with dual identities and those identifying as "other" to examine being British

---

[38] The question wording is given in the data section of the paper.

[39] That is, England, Scotland, or Wales. Those with dual identities – say both English and British – are removed in this second national identity measure.

[40] Unionist is used here to distinguish the two outcomes. To identify as British, even if one would not call oneself a unionist explicitly, is to identify with the political union that exists between England, Scotland and Wales on the island of Great Britain (the largest island of the British Isles).

versus being English, Scottish, Welsh, or Northern Irish.[41] As the RoSLA impacts those born after August of 1957 and the data begins in 2001, the individuals under examination will have been out of education for some time – as a result, these estimates constitute an analysis of the long-term impact of education on identity and contribute to the question of the role of education in building a nation.

This analysis focuses specifically on those born in Great Britain and the findings are equivocal. In the Quarterly Labour Force Survey (QLFS), OLS estimates indicate statistically significant effects of education on all forms of identity. By comparison, IV estimates are insignificant for British identity (the first national identity measure), but negative and significant for unionist identity (the second national identity measure). The magnitude of IV estimates is generally larger than OLS estimates, but the degree of imprecision is such that this does not always matter for statistical significance. In a second dataset, the British Social Attitudes Survey (BSA), whilst OLS estimates are positive and significant (though of slightly smaller magnitude than the QLFS results), none of the IV estimates are significant, though they are of the same sign. It is possible that this is a local average treatment effect and the OLS represents the population average effect but articulating a story as to why this might be so is not simple.

The paper proceeds as follows. Section 3.2 explores the relevant literature; section 3.3 describes the data used for analysis. Section 3.4 outlines the RoSLA reform. Section 3.5 presents the empirical strategy. And sections 3.6 and 3.7 present results, and the discussion and conclusion, respectively.

---

[41] Those currently resident in Northern Ireland are not included in analysis for a number of reasons. Primarily, unionism or identifying with Britain means something very different to the other UK nations given the political context in Northern Ireland at the time of the RoSLA. Another reason is that there is no "Northern Ireland" response till the question until 2011, so it is unclear how those observations should be treated. But it remains that those resident in other UK nations can identify with Northern Ireland – they are few in number but are included in the outcome variables. Removing them makes no difference to results.

3.2 Related Literature

It took some time for economist to turn their attention to issues of identity. The seminal papers are those of Akerlof & Kranton (2000, 2002).[42] The latter deals specifically with education, formalising the sociological literature in defining different groups that are present in schools (e.g., the 'Jocks' and the 'Nerds') whose social groups may influence the seriousness with which they approach their schooling – in essence identity shapes educational engagement and attainment. Work is also referenced that may suggest the opposite; for example, Bowles & Gintis (1976) argue that US schools were driven to create compliant workers, whilst Kremer & Sarychev (2000) equate Western schools to factories for democratic values.

Recent work by Bandiera et al. (2018) provides a clear justification for examining the nation building impact of education. They provide qualitative evidence that the intention of policy makers in passing compulsory schooling laws in the US was precisely in order to instil American civic values. They go on to conduct survival analysis on the timing of compulsory schooling laws across US states, showing that the share of European-born migrants in those states affects the speed of adoption. They find that a larger share of migrants results in compulsory schooling laws being passed sooner – an increase of one standard deviation in the percentage of migrants resulting in a doubling of the hazard of a schooling law being brought into effect.

Clots-Figueras & Masella (2013) examine the impact of language on identity by looking at the shift to compulsory teaching of the Catalan language in Catalonia, Spain. Using the reform as an instrument for time exposed to compulsory language teaching, they find that more years of Catalan language training leads, causally, to an individual identifying more strongly as Catalan. In a similar vein, Fouka (2019) finds a German language ban in Ohio and Iowa after the First

---

[42] Though, the economics of religion, a crucial part of the identity of millions of people across the globe, can be traced further back e.g., Iannaccone (1998).

119

World War resulted in those affected being more likely to marry within the German community and provide their offspring with a German name – unintended consequences of the law.

Some papers have used compulsory schooling reforms for similar outcomes to this paper. Milligan et al. (2004) find positive impacts in the United States on voting participation, but do not find effects in the UK. Bommel & Heineck (2020) examine education's impact on political participation and interest using compulsory schooling changes in West Germany. They cite a large range of correlational studies that examine political outcomes but note that there are few causal studies in this area. Whilst OLS estimates are positive and significant, IV estimates suggest that education does not stimulate political participation or interest. In the context of Euroscepticism, Kunst et al. (2020) use compulsory schooling laws in the UK, the Netherlands, Denmark, and Sweden to find no statistically significant effects of education on sentiments towards Europe.[43]

The discussion around the extent to which schools can impart values is the key focus of this paper. If schools impart values, then increased years of schooling could lead individuals to be more likely to identify as British. The potential power of such a process is evidenced in Voigtländer & Voth (2015). Using a study of Germans interviewed in 1996 and 2006, they investigate the impact of Nazi indoctrination through education and advertising. Young Germans who lived through the period of indoctrination (around 1933-1945) were 2 to 3 times more likely to express anti-Jewish sentiments later in life than the population as a whole. Although an extreme example, this shows the pervasiveness of values that are learned when young. Even though the educational process in the UK would be starkly different to this case, we evidence the potential for mechanisms such as schooling to impart values. Similarly, Cantoni et al. (2017) show the effects of a school textbook reform in China, with more

---

[43] See the references in Kunst, Kuhn, and van der Werfhorst (2020) for a review of the literature around education and Euroscepticism.

favourable views of China's governance as well as scepticism of free markets, on the views of those subject to the new materials. Students' views are significantly impacted. [44]

Much of a young person's educational choices will be influenced, if not directed or dictated, by their parents. Giusta et al. (2017) present a model in which parents may choose lower levels of education for their children if they themselves identify with their original culture and see schools as favouring the majority culture. The model predicts that this situation will reverse if schools can suitably accommodate the values of the communities in which they reside. Mechanism aside, the benefit of the raising of the school leaving age is, as outlined below, that the choice of how long a child stays in school is removed from the parent (and the child for that matter).

Schüller (2015), using the GSOEP dataset, finds that parental identity is important for the educational attainment of their children. The author employs a sibling difference model using family fixed effects to attempt to control for unobserved family background to attempt to improve on a simple linear approach. There appears to be no detrimental impact of parents adhering to their original culture; in fact, identifying with the majority (German) culture rather than sticking to one's original culture each seem to have different potential benefits. Majority culture, acting mostly through the mother, has a positive impact on human capital accumulation, with the suggested mechanism being through language. Original culture is said to have a "stabilising effect" in terms of a child's feeling of place and self-worth (Ibid., p969). The method relies on age differences in siblings – as a result, much could have changed in the

---

[44] There is also a literature on assimilation of migrants. Gang & Zimmermann (2000), using the GSOEP dataset to compare educational attainment of migrants and natives of similar characteristics, suggest that the size of co-ethnic networks upon their arrival in Germany is important. Parental education level is said to be unimportant, but ethnicity appears to play a significant role. Most explicitly, Nekby et al. (2009) investigate the association between ethnic identity and education. Identifying with the 'home' culture of the country they are in results in a migrant being more likely to complete tertiary education.

intervening years between siblings; ultimately this method does not seem appropriate to identify effects.

3.3 Data

The primary dataset in this paper is the Quarterly Labour Force Survey (QLFS) (Office for National Statistics, 2021) one of the main datasets used in the UK for labour market analysis. The LFS has run since 1973. It ran biennially until 1983, annually until 1991, and quarterly from 1992 onwards (becoming the QLFS). The sample is large in any given quarter – up to 0.15% of the population. Though later waves of the data have sampled a smaller number of people, there is still a large pool of people to include in this analysis.

The QLFS contains a range of variables including month and year of birth, age the individual left education, their age at the time of the survey, nationality, country of birth, gender, ethnicity, and region of residence. Although the list could usefully include other questions – such as those covering non-cognitive skills – it represents a comprehensive list of covariates for analysis and is certainly sufficient for estimating a parsimonious instrumental variable specification. In the analysis that follows we focus on those born in England, Scotland, or Wales.

Since 2001 the QLFS has contained information on an individual's national identity.[45] In terms of national identity, the same six answers exist for each question: English, Welsh, Scottish, Northern Irish, British, or other.[46] The ordering of the options changes depending on which nation of the UK the question is being asked in, as shown in Table 3.1.

---

[45] The question wording changes slightly in 2011. From 2001 the wording is "How would you describe your national identity? Please choose all that apply." From 2011 the wording is "What do you consider your national identity to be? You may choose as many as apply. Is it...".

[46] The 'Northern Irish' option is just 'Irish' prior to 2011.

Table 3.1   Ordering of Question Responses by Nation of Residence (QLFS)

England	Scotland	Wales
(1) English	(1) Scottish	(1) Welsh
(2) Welsh	(2) English	(2) English
(3) Scottish	(3) Welsh	(3) Scottish
(4) Northern Irish	(4) Northern Irish	(4) Northern Irish
(5) British	(5) British	(5) British
(6) Other	(6) Other	(6) Other

Two national identity outcomes are constructed from this question. The first is British identity which is coded as 1 if the individual responds as British, even if they also respond as other identities, and 0 if they do not mention British.[47] This variable seeks to measure whether people are willing to say that they are British, even if they also have another identity.

The second identity outcome focuses on whether individuals identify with their home nation (i.e., England, Scotland, or Wales) or the parent nation (the UK). For this variable, referred to as unionist identity, individuals are coded as 1 if they identify as British, and 0 if they identify as English, Scottish, Welsh, or Northern Irish. Those with a dual identity are dropped as we are unable to tell if they prefer, or see themselves as more closely linked to, their British identity or their home nation identity. The smaller sample sizes in the regressions below reflect this choice and the fact that those who identify as "other" are dropped from this sample as well.

The dataset includes only those who are old enough to have been impacted by the RoSLA. I show a range of bandwidth estimates, but primarily (after Table 3.3) I use those born within 24

---

[47] Northern Irish identity is, in line with the footnote above, only counted as a British Identity after 2011.

months of the RoSLA on either side. This fairly small window of observations is kept to avoid confounding due to the number of years of education individuals were receiving rising slowly over time. It is not possible to know exactly where individuals were living at the time of the RoSLA, but, whilst prior papers examining compulsory schooling laws in the UK focused purely on England and Wales due to concerns around how the RoSLA was implemented in Scotland and Northern Ireland, recent work by Buscha & Dickson (2018) has shown that the RoSLA was implemented at the same time (in law) in Scotland and Northern Ireland as it was in England and Wales. They show similar positive first stage impacts in Scotland to England and Wales and state that many previous authors "incorrectly" believe that the RoSLA only took place in 1972 in England and Wales.

The QLFS is collected as a rolling panel (though here we have only the cross-sectional variant) which means that around a fifth of individuals in each wave are in their first interview, a fifth in their second interview, and so on. Following Dickson and Smith (2011), who also use the QLFS for their analysis of the RoSLA, we keep individuals in their first interview only.[48] Table 3.2 shows descriptive statistics for the pre- and post-RoSLA samples. As is evident, the two are similar in terms of characteristics like gender and race both sides of the RoSLA reform.

Further analysis is also conducted, in part as a robustness check, using the British Social Attitudes Survey (BSA). Since 1999 BSA has consistently asked about "Britishness", and about which national identity fits the individual best. This enables a comparison with the QLFS. This first outcome – British identity – is the same as for the QLFS; simply whether the individual identifies as British or not. The second outcome – unionist identity – is slightly different. It is the response to the question - which identity best represents the respondent. This addresses a flaw in the QLFS – where we do not know what the preferred identity of a dual

---

[48] This does not turn out to be consequential for results.

identity individual is. Table 3.3 shows descriptive statistics for the BSA data. These are in fact quite different to the LFS data, with higher average responses to the identity questions. The fraction leaving school at age 16 or older is also slightly higher.

Table 3.2   Descriptive Statistics Pre- and Post-RoSLA (QLFS)

Variable	Pre-RoSLA			Post-RoSLA		
	Mean	SD	N	Mean	SD	N
British Identity	0.415	0.493	110982	0.429	0.495	125596
Unionist Identity	0.346	0.476	97587	0.360	0.480	110008
RoSLA	0.000	0.000	111573	1.000	0.000	126311
Age Left FT Education	16.978	2.445	110582	17.235	2.224	125010
Left School Post-16	0.690	0.463	110582	0.924	0.265	125010
Female	0.517	0.500	111573	0.518	0.500	126311
Not White	0.022	0.146	111573	0.036	0.186	126311
Age	54.545	5.732	111573	48.354	5.730	126311
England	0.836	0.370	111573	0.843	0.364	126311
Year	2009.244	5.436	111573	2009.133	5.428	126311
Birth Year	1954.215	1.773	111573	1960.289	1.792	126311
Birth Month	6.353	3.411	111573	6.429	3.404	126311

Note: British Identity is a binary variable that takes value 1 if the individual identifies as British, regardless of dual identity, and 0 otherwise. Unionist Identity is different – a binary variable that takes value 1 if the individual identifies as British and 0 if they identify as English, Welsh, Scottish, or Northern Irish; dual identities are omitted. RoSLA takes value 1 if the individual was born in September 1957 or later; Age left FT Education is the age an individual left school; Left School Post-16 is a dummy variable that takes value 1 if an individual left school at age 16 or older. Female is 1 if the individual is a woman, and 0 if they are a man; Not White takes value 1 if the individual is not white, and 0 if they are white. Age is the age of the respondent at the time of the survey. England shows the proportion living in English regions. Year is the survey year of the survey, Birth Year and Birth Month provide the individuals date of birth.

3.4   The Raising of the School Leaving Age

The RoSLA reform has been widely used in empirical analysis, but it is worth outlining again in the context of identity. The 1972 reform raised the compulsory schooling age to 16 from 15, the age it had been since 1947 (before which it had been 14). This paper only makes use of the 1972 reform.

Table 3.3   Descriptive Statistics Pre- and Post-RoSLA (BSA)

Variable	Pre-RoSLA			Post-RoSLA		
	Mean	SD	N	Mean	SD	N
British Identity	0.686	0.464	6046	0.694	0.461	7064
Unionist Identity	0.515	0.500	4871	0.518	0.500	5678
RoSLA	0.000	0.000	6049	1.000	0.000	7067
Age Left FT Education	17.043	2.579	6049	17.466	2.434	7067
Left School Post-16	0.656	0.475	6049	0.912	0.283	7067
Female	0.539	0.499	6049	0.547	0.498	7067
Not White	0.054	0.226	6049	0.080	0.272	7067
Age	53.978	5.378	6049	46.509	5.385	7067
England	0.848	0.359	6049	0.860	0.347	7067
Year	2007.472	5.066	6049	2007.072	5.067	7067
Birth year	1953.494	1.703	6049	1960.563	1.692	7067

Note: British Identity is a binary variable that takes value 1 if the individual identifies as British, regardless or dual identity, and 0 otherwise. Unionist Identity is different – a binary variable that takes value 1 if the individual identifies as British and 0 if they identify as English, Welsh, Scottish, or Northern Irish; dual identities are omitted. RoSLA takes value 1 if the individual was born in September 1957 or later; Age left FT Education is the age an individual left school; Left School Post-16 is a dummy variable that takes value 1 if an individual left school at age 16 or older. Female is 1 if the individual is a woman, and 0 if they are a man; Not White takes value 1 if the individual is not white, and 0 if they are white. Age is the age of the respondent at the time of the survey. England shows the proportion living in English regions. Year is the survey year of the survey, Birth Year provide the individuals year of birth; month of birth is not available in BSA.

Ernest Bevin, a key proponent of the RoSLA reforms and a cabinet minister in the post-Second World War Attlee government, wanted the second RoSLA reform to occur within "no longer than 3 years" of the first or else he feared it would not happen for another 20 years (Woodin, McCulloch, and Cowan, 2013: p. 71). 25 years later, the leaving age was raised to 16. A host of reasons exist for why the implementation took so long. First was a lack of suitable buildings – the solution to which were the hastily erected prefabricated buildings, known as "RoSLA blocks". Economic factors were present too, which lead to concerns about the budgetary implications of such a large increase in education funding needed to accommodate the reform. Moreover, there were worries of a decline in the quality of provision in inner-city schools due to the increased numbers of students and their composition. Some teachers, it seems, were

concerned about, and may even have changed jobs as a result of, the racial composition of classes that would result from the RoSLA (Ibid; p.112).

Something that is generally ignored in papers that deal with the 1972 RoSLA reform is the curriculum changes that took place to coincide with the RoSLA. Curriculum content is shown to impact identity in Cantoni et al (2017), and so changes in the UK curriculum at the time of the reform are of interest. Woodin, McCulloch and Cowan (2013) report that "there was a widespread expectation that a new curriculum would address the problem of disaffection among young people" (p. 99). For the most part, this seems to have been related to making sure that a sufficient list of courses was available that would contribute towards future employment. Moreover, teachers would be expected to adapt their teaching to the interests of their pupils. Crucially though, "schooling was conceived as an agency for the promotion of culture, discrimination, and civilisation" (Ibid; p. 100). Essentially, the content became more explicitly linked to national identity. What this means is that the instrument, whilst still valid, could only be said to be a combined curriculum and schooling quantity change beginning with the September 1957 cohort. That said, this would be an issue for every paper that uses the RoSLA in the UK – to my knowledge this has not been acknowledged previously.

3.5 Empirical Strategy

The problem with identifying a causal impact of education on national identity is the potential endogeneity of the treatment. In essence, certain types of people stay on in school and those that do could be people already predisposed to respond to questions of identity in particular ways. Some individuals may be more inclined towards further education and identify in certain ways – their optimal amount of education would be higher but would not be the driver of how they identify.[49]

---

[49] For example, in the context of religious belief, Jehovah's Witnesses actively encourage against seeking out higher levels of education.

These challenges mean that an instrument is needed. We use the RoSLA in this paper - an instrument that has been used widely in economics.[50] Once enacted, the RoSLA removes choice from those 15 years olds who would have left in the absence of the reform and enables a comparison of those prior to the reform and those after. Specifically, the empirical specification is as follows:

$$Identity = \beta_0 + \beta_1 Stayed\ in\ School\ Post\ 16_i + \beta_2 X_i + \epsilon_i \qquad (6)$$

$$Stayed\ in\ School\ Post\ 16 = \gamma_0 + \gamma_1 RoSLA_i + \gamma_2 X_i + \mu_i \qquad (7)$$

Where RoSLA takes a value 1 if the individual is born after September 1957 and 0 if they were born before then. X_i is a vector of control characteristics including age the individual left education, nationality, country of birth, gender, ethnicity, region of residence and is common across the two stages. A strength of the data we have is the presence of month of birth, not available in the regular, "end user license" version of the data available from the UK Data Service website. This removes measurement error in the running variable and improves the precision of the second stage.

The endogenous variable is binary, rather than the more usual terminal education age. Figure 3.1 replicates, using the LFS data from this paper, a chart from Chevalier et al (2004) showing that the main effect of the RoSLA reform was to drastically reduce the share of individuals leaving school at 15 and to correspondingly increase the share who left at 16. There is little change in those leaving at older ages – essentially, there is virtually no ripple effect of the RoSLA. The RoSLA instrument is substantially stronger when used on the binary variable of leaving school at or below 15 versus 16 and older. Results are very close to those that are

---

[50] There is, of course, an extensive literature on the impact of compulsory schooling laws in relation to wages (see, for example, Brunello et al. (2016); Grenet (2013); Harmon & Walker (1995); Oreopoulos (2006)) as well as in relation to other outcomes, such as the impact of education on health (Clark & Royer, 2013); Lleras-Muney, 2005), likelihood of committing a crime or ending up in prison (Lochner & Moretti, 2004) and productivity (Chevalier et al., 2013).

obtained using the traditional terminal education age as the endogenous variable (not shown in the paper), but the increased strength of the instrument considering recent work by Lee et al (2020) makes the binary endogenous variable more appealing.

In practise we follow the implementation of a local-linear IV as recommended in Imbens and Lemieux (2008) and Gelman and Imbens (2014), running the analysis as a regression discontinuity design where linear regressions are run on data within some distance of the cut-off birthdate of September 1st, 1957. As in Buscha and Dickson (2018), who use the same data source as this paper to examine the RoSLA in Scotland and Northern Ireland, we show results for a range of bandwidths using a rectangular kernel.

Figure 3.1   Proportions Leaving School at Each Age Pre- and Post-RoSLA (QLFS)
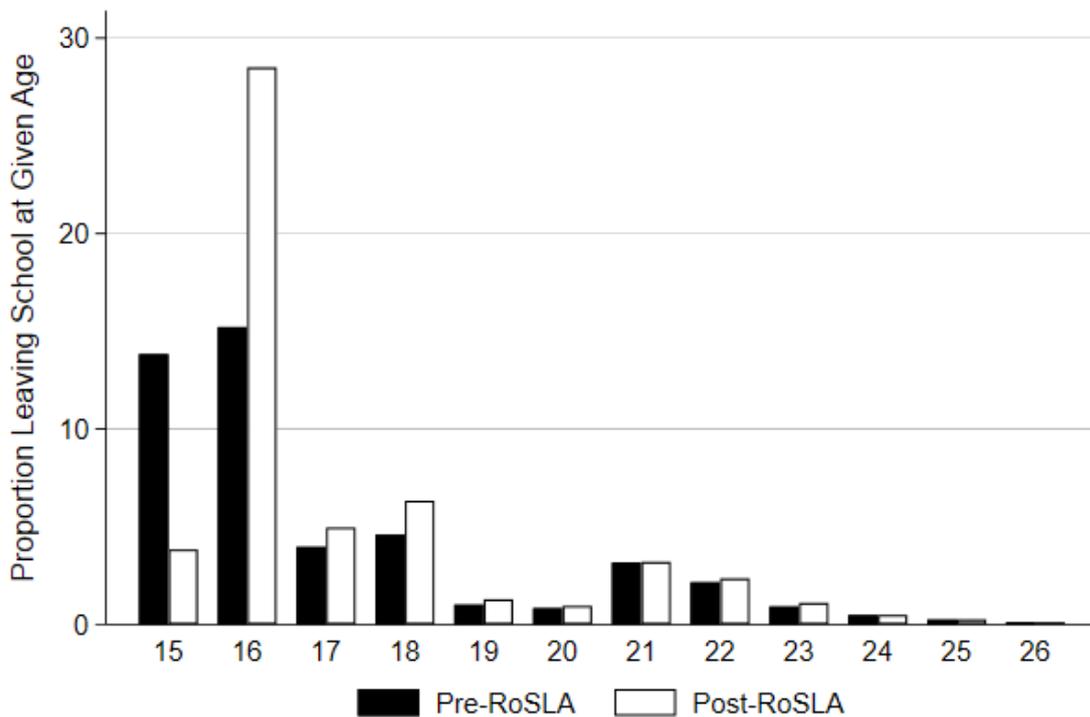
Figure 3.2 demonstrates the discontinuity at work. The x-axis shows "school cohorts" rather than years, so the line at 1957 represents the 1957-58 school year covering births from September of 1957 to August of 1958. As the reform takes effect in 1957 there is a jump in the

proportion staying in school until age 16 or later of around 20 percentage points – from just over 70 percent to just over 90 percent. This corresponds to an increase of roughly 28 percent.

Figure 3.2   Proportion Leaving School Post-15 (QLFS)
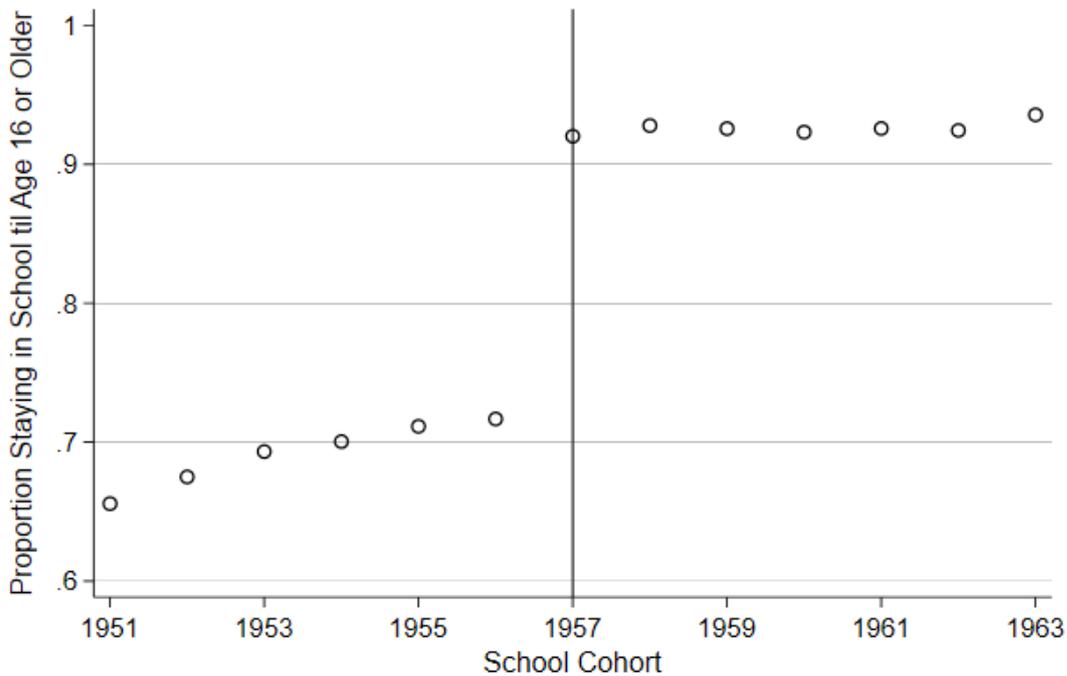
Figure Note: School Cohort refers to the year in which a given school year started. I.e., the 1957/58 school year in which the RoSLA was introduced appears as 1957 on the chart.

3.6   Identity in the Quarterly Labour Force Survey

3.6.1   Education's Impact on Identity

Table 3.4 shows the first stage and reduced form effects of the RoSLA reform at different bandwidths. These bandwidths increase by 12 months in each successive panel following Busha and Dickson (2018), who do this to avoid within school-year month of birth effects. Across all bandwidths the first stage is stable across columns and the magnitude of coefficients is virtually the same for all bandwidths from the "24 Month Bandwidth" onwards.

Table 3.4   **First Stage and Reduced Form Estimates (QLFS)**

	(1)	(2)	(3)	(4)	(5)
		First Stage		Reduced Form	
				British Identity	Unionist Identity
		Left School Post-Age 15			
Panel A – 12 Month Bandwidth					
RoSLA	0.253***	0.270***	0.271***	-0.016	-0.028**
	(0.009)	(0.011)	(0.011)	(0.013)	(0.013)
N	38,385	38,385	38,385	38,385	33,701
R-squared	0.075	0.082	0.082	0.035	0.041
Panel B – 24 Month Bandwidth					
RoSLA	0.220***	0.228***	0.228***	-0.009	-0.022**
	(0.006)	(0.007)	(0.007)	(0.009)	(0.009)
N	76,894	76,894	76,894	76,894	67,437
R-squared	0.078	0.083	0.084	0.033	0.038
Panel C – 36 Month Bandwidth					
RoSLA	0.216***	0.223***	0.224***	-0.007	-0.019**
	(0.005)	(0.006)	(0.006)	(0.008)	(0.009)
N	114,928	114,928	114,928	114,928	100,784
R-squared	0.081	0.086	0.087	0.032	0.038
Panel D – 48 Month Bandwidth					
RoSLA	0.216***	0.218***	0.218***	-0.009	-0.021**
	(0.004)	(0.006)	(0.006)	(0.008)	(0.008)
N	153,588	153,588	153,588	153,588	134,753
R-squared	0.083	0.088	0.089	0.032	0.038
Panel E – 60 Month Bandwidth					
RoSLA	0.212***	0.214***	0.214***	-0.009	-0.021**
	(0.004)	(0.006)	(0.006)	(0.008)	(0.008)
N	193,186	193,186	193,186	193,186	169,563
R-squared	0.086	0.092	0.092	0.032	0.038
Panel F – 72 Month Bandwidth					
RoSLA	0.208***	0.213***	0.213***	-0.009	-0.020**
	(0.003)	(0.006)	(0.006)	(0.008)	(0.008)
N	232,915	232,915	232,915	232,915	204,398
R-squared	0.091	0.098	0.098	0.033	0.039

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
Note: Column (1) includes the instrument and running variable without additional controls. Column (2) adds controls for year of birth, age and age squared, survey year interacted with survey quarter, and region of residence. Column (3) adds gender, ethnicity, and country of birth. Columns (4), (5), and (6) contain the column (3) controls. First stage sample corresponds to the sample present for the British Identity outcome; the other two outcomes vary in sample size as a result of how they are constructed.

In column (3) of Panel B, for example, the first stage coefficient of 0.228 is very close to the visual gap presented in Figure 2. The reduced form equations show little in the way of effects that are statistically different from zero for British identity, but there are statistically significant coefficients associated with unionist identity.

Table 3.5   OLS and Second Stage IV Results (QLFS)

	(1) OLS	(2) IV	(3) OLS	(4) IV
	British Identity		Unionist Identity	
Left School Post-15	0.068***	-0.039	0.075***	-0.105**
	(0.005)	(0.040)	(0.005)	(0.043)
Female	0.020***	0.021***	0.021***	0.022***
	(0.004)	(0.004)	(0.004)	(0.004)
Not White	0.202***	0.204***	0.287***	0.290***
	(0.017)	(0.017)	(0.018)	(0.019)
Born in Scotland	-0.109***	-0.111***	-0.139***	-0.134***
	(0.011)	(0.011)	(0.012)	(0.013)
Born in Wales	-0.152***	-0.151***	-0.196***	-0.194***
	(0.012)	(0.012)	(0.013)	(0.013)
First Stage F		1048.670		815.218
N	76,894	76,894	66,165	66,165
R-squared	0.036	0.029	0.041	0.022

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The first stage F is the Kleibergen-Paap F Statistic. Each column includes the instrument and running variable as well as controls for year of birth, age and age squared, survey year interacted with survey quarter, region of residence, gender, ethnicity, and country of birth.

An interesting result is presented in Table 3.5. In each of the OLS specifications, resulting from regressing each outcome on the treatment and the full range of controls *without* instrumenting whether the individual stayed in school until after age 15, there are significant effects associated with leaving school at 16 instead of 15 or younger. For British identity, the effect is around 7 percentage points, significant at the one percent level and for unionist identity around 8 percentage points, significant at the one percent level. Once education is instrumented the story

changes. Only the effect on unionist identity remains statistically significant – but has changed sign. The magnitude is greater at negative 11 percentage points. For British identity, the coefficient has also changed sign but is not statistically significant. The first stage F statistics are large in each case.

3.6.2   Heterogeneity

Tables 3.5 and 3.6 examine some dimensions of heterogeneity. The effect of education on unionist identity in Table 3.4 seems, in Table 3.6, to be primarily driven by women (in Panel A); the effect is large at negative 12 percent and is significant at the five percent level. That is the only statistically meaningful effect. For men (Panel B), there are no significant IV effects, although the OLS effects for columns (1) and (3) are basically indistinguishable.

Table 3.7 examines heterogeneity by country of birth. In Panel A, the results for England match those of the whole sample in Table 3.3 closely - unsurprising given that England makes up most of the sample. But the coefficient on unionist identity is no longer significant in column (4). Panel C, for Wales, does not display any significant effects across the six columns despite the instrument still being reasonably strong. In each of the IV cases the magnitudes are large but noisy. It appears that the results for unionist identity effect is driven by Scotland, in Panel B – where the effect is negative and large at around 20 percentage points. Strangely, these effects are not mirrored in Table C1, which looks at heterogeneity by nation of current residence rather than birth. There are no significant effects in the IV specifications in any of the three panels.

Table 3.6   Heterogeneity by Gender (QLFS)

	(1) OLS	(2) IV	(3) OLS	(4) IV
	British Identity		Unionist Identity	
Panel A - Female				
Left School	0.063***	-0.069	0.074***	-0.115**
Post-15	(0.007)	(0.054)	(0.007)	(0.055)
Not White	0.208***	0.210***	0.304***	0.304***
	(0.023)	(0.023)	(0.024)	(0.024)
Born in	-0.125***	-0.127***	-0.150***	-0.152***
Scotland	(0.015)	(0.015)	(0.016)	(0.016)
Born in	-0.172***	-0.170***	-0.215***	-0.211***
Wales	(0.016)	(0.016)	(0.017)	(0.018)
First Stage F		584.271		515.208
N	40,639	40,639	35,602	35,602
R-squared	0.040	0.031	0.047	0.026
Panel B - Male				
Left School	0.076***	-0.003	0.076***	-0.074
Post-15	(0.007)	(0.057)	(0.007)	(0.060)
Not White	0.195***	0.197***	0.275***	0.278***
	(0.025)	(0.025)	(0.026)	(0.026)
Born in	-0.090***	-0.092***	-0.130***	-0.135***
Scotland	(0.015)	(0.015)	(0.016)	(0.017)
Born in	-0.131***	-0.130***	-0.169***	-0.168***
Wales	(0.018)	(0.018)	(0.018)	(0.019)
First Stage F		484.997		416.462
N	37,680	37,680	33,096	33,096
R-squared	0.032	0.029	0.037	0.023

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The first stage F is the Kleibergen-Paap F Statistic. Each column includes the instrument and running variable as well as controls for year of birth, age and age squared, survey year interacted with survey quarter, region of residence, gender, ethnicity, and country of birth.

Table 3.7   Heterogeneity by Nation of Birth (QLFS)

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
	British Identity		Unionist Identity	
Panel A – England				
Left School Post-15	0.074***	-0.009	0.083***	-0.063
	(0.005)	(0.046)	(0.006)	(0.048)
First Stage F		801.375		677.378
N	64,147	64,147	56,139	56,139
R-squared	0.023	0.019	0.025	0.013
Panel B - Scotland				
Left School Post-15	0.064***	-0.092	0.048***	-0.205**
	(0.012)	(0.097)	(0.011)	(0.085)
First Stage F		125.208		119.241
N	8,872	8,872	7,835	7,835
R-squared	0.079	0.061	0.098	0.037
Panel C - Wales				
Left School Post-15	0.022	-0.231	0.028	-0.281
	(0.020)	(0.179)	(0.018)	(0.178)
First Stage F		46.983		35.712
N	3,927	3,927	3,437	3,437
R-squared	0.081	0.042	0.121	0.044

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The first stage F is the Kleibergen-Paap F Statistic. Each column includes the instrument and running variable as well as controls for year of birth, age and age squared, survey year interacted with survey quarter, region of residence, gender, ethnicity, and country of birth.

3.6.3   Further Robustness Tests

Robustness of the first stage is already established in Table 3.4. This also shows the stability of the reduced form estimates. But before proceeding to the second dataset which acts as an additional robustness check, it is worth conducting some further analysis on this QLFS dataset.

Firstly, region of current residence could be an important factor in shaping or reflecting characteristics of a particular individual. For example, those living in London may have a more liberal or metropolitan outlook. To examine whether one region is driving the results, regressions are run for each of the two national identity outcomes excluding one region at a time; the results are shown in Table 3.8.

Table 3.8   Removing Regions of Residence from the Sample (QLFS)

Omitted Region	British Identity			Unionist Identity		
	OLS	IV	N	OLS	IV	N
North East	0.068***	-0.045	72,665	0.075***	-0.116***	62,558
	(0.005)	(0.041)		(0.005)	(0.045)	
North West	0.069***	-0.035	69,107	0.076***	-0.099**	59,588
	(0.005)	(0.041)		(0.005)	(0.044)	
Mersey	0.068***	-0.045	74,884	0.075***	-0.114***	64,341
	(0.005)	(0.040)		(0.005)	(0.043)	
Yorkshire	0.068***	-0.054	69,545	0.074***	-0.125***	59,952
	(0.005)	(0.043)		(0.005)	(0.046)	
East Midlands	0.068***	-0.039	70,824	0.074***	-0.096**	60,810
	(0.005)	(0.041)		(0.005)	(0.045)	
West Midlands	0.068***	-0.034	70,108	0.074***	-0.111**	60,250
	(0.005)	(0.041)		(0.005)	(0.045)	
East England	0.067***	-0.046	69,237	0.075***	-0.112**	59,422
	(0.005)	(0.042)		(0.005)	(0.046)	
London	0.066***	-0.040	72,254	0.073***	-0.100**	62,160
	(0.005)	(0.040)		(0.005)	(0.043)	
South East	0.070***	-0.020	66,007	0.077***	-0.086*	56,896
	(0.005)	(0.041)		(0.005)	(0.045)	
South West	0.067***	-0.052	69,736	0.074***	-0.114**	60,267
	(0.005)	(0.042)		(0.005)	(0.045)	
Wales	0.069***	-0.025	72,666	0.077***	-0.092**	62,491
	(0.005)	(0.041)		(0.005)	(0.044)	
Scotland	0.065***	-0.029	68,801	0.075***	-0.092*	59,080
	(0.005)	(0.043)		(0.005)	(0.047)	
Scotland and Wales	0.066***	-0.012	64,573	0.076***	-0.076	55,406
	(0.005)	(0.044)		(0.005)	(0.049)	

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
Each column includes the instrument and running variable as well as controls for year of birth, age and age squared, survey year interacted with survey quarter, region of residence, gender, ethnicity, and country of birth.

The OLS results are very stable and barely change for any given region being removed from the sample. The IV estimates vary slightly more, although statistically significant differences are rarely present. The one case in which differences do occur is with the removal of Scotland and Wales. The final three rows of the table show each of these being removed separately and

together. The magnitudes dip, along with the significance, for the two removed separately, and then dips further for the two removed simultaneously. In this final case the result also becomes insignificant. This suggests that Wales and Scotland are driving the result, although a lack of precision remains an issue. Secondly, around 25,000 individuals (out of the full sample of around 230,000) report leaving school before 16 but possessing qualifications that they could only have acquired if they had undertaken schooling beyond that point. These observations could either be mistakes or people who did leave school at 15 but then later returned. Either of these could cloud results. Once these individuals are removed the result, in Table 3.9, is surprising – the coefficient is substantially larger for both British identity and unionist identity. The former is negative 16 percentage points, and the latter is negative 30 percentage points. Both are significant, at the 10% and 1% levels, respectively. This suggests that there is something different about the individuals that the RoSLA had an impact on.

The estimated specifications are, so far, quite parsimonious. This is by design – as treatment is in the past it makes sense to primarily focus on covariates that are time invariant. That said, a further check is to include some additional controls into the specifications. Including additional controls (not shown) for marital status, employment status, and religious affiliation leaves the estimated coefficients broadly the same. The effect on British identity is still insignificant and slightly smaller; the effect on unionist identity is also slightly smaller at around -0.071 but remains significant. These differences are not statistically significant. The range of available controls in the QLFS is not as rich as one might have hoped for – the additional controls represents the extent of what is reasonably possible in the data.[51]

---

[51] The same analysis was also run using the third quarter interview for each person (excluding people in their final interview to avoid double counting). This had initially been done in order to use the English as a first language variable. Sample sizes were small for this variable however and did not show anything worth including in this paper – but, comfortingly, results were consistent with use of the first interview only as occurs in this analysis. This is a further form of robustness check.

Table 3.9   Removing those with Qualifications from after their Leaving Age (QLFS)

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
	British Identity		Unionist Identity	
Left School Post-15	0.094***	-0.162*	0.097***	-0.306***
	(0.007)	(0.091)	(0.007)	(0.099)
Female	0.019***	0.013***	0.020***	0.013***
	(0.004)	(0.004)	(0.004)	(0.005)
Not White	0.188***	0.191***	0.272***	0.279***
	(0.018)	(0.018)	(0.019)	(0.020)
Born in Scotland	-0.116***	-0.120***	-0.144***	-0.145***
	(0.011)	(0.011)	(0.012)	(0.013)
Born in Wales	-0.153***	-0.152***	-0.193***	-0.192***
	(0.013)	(0.013)	(0.013)	(0.014)
First Stage F (Kleibergen-Paap)		254.893		249.828
N	69,093	69,093	59,476	59,476
R-squared	0.036		0.041	

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
Each column includes the instrument and running variable as well as controls for year of birth, age and age squared, survey year interacted with survey quarter, region of residence, gender, ethnicity, and country of birth.

3.7   Identity in the British Social Attitudes Survey

The British Social Attitudes data provides a form of robustness check on the QLFS results.

Figures 3.3 and 3.4 replicate for the BSA what Figures 3.1 and 3.2 show for the QLFS. Figure

3.3 shows much the same story as Figure 1, although there is slightly more evidence of a

RoSLA ripple effect that was evident in the QLFS. Figure 3.4 is strikingly like Figure 3.2, with

the jump at the discontinuity being virtually identical. In the BSA case the first treatment year

is 1958 as the data does not contain month of birth. In the regression analysis below, 1957 is

removed and the RoSLA variable takes value 0 if an individual was born in 1956 or earlier,

and value 1 if the individual was born in 1958 or later.

Figure 3.3   Proportions Leaving School at Each Age Pre- and Post-RoSLA (BSA)
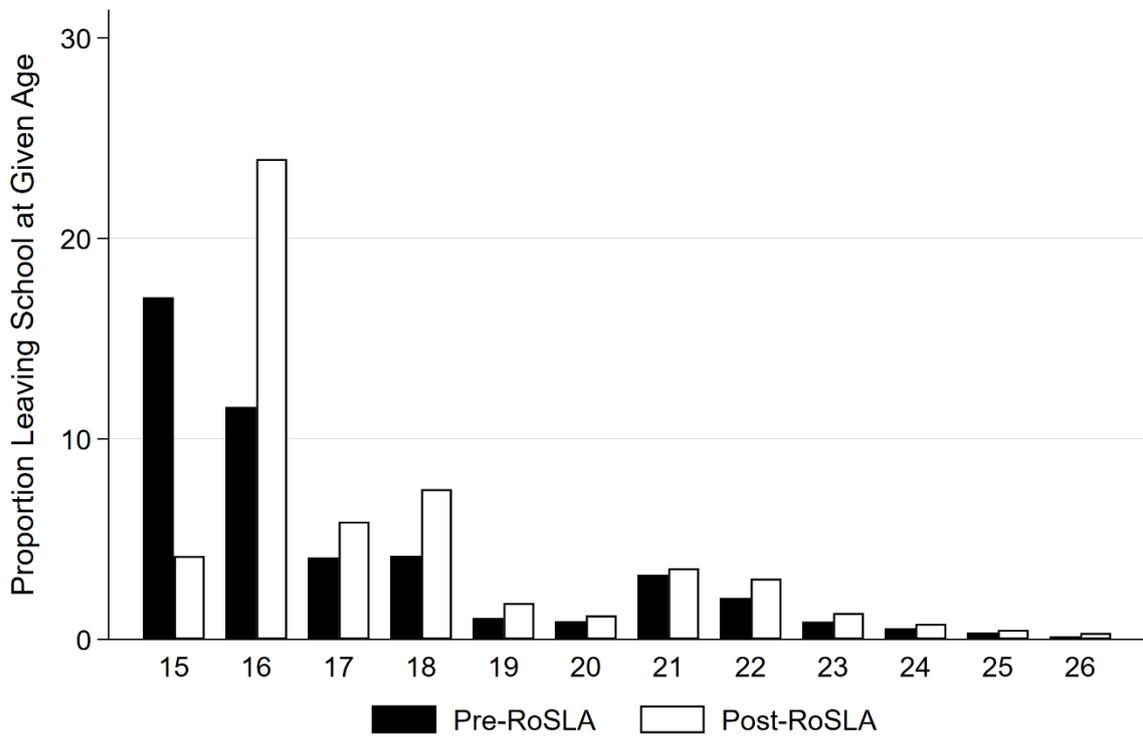
Figure 3.4   Proportions Leaving School Post-15 (BSA)

The similarities continue once the regression analysis is conducted. Table 3.10 shows the

complete array of regression specifications using the BSA data. Because of the smaller sample

size, and in the interests of efficiency, this uses all observations shown in Figure 3.4, running from 1951 (six full years before the reform) to 1963 (six full years after), excluding 1957. Table C2 in the appendix shows that this is not a consequential choice in terms of the magnitude of the first stage estimates – the first stage estimates are remarkably stable.

The first stage results in column (1) are close to the QLFS results above at around 20 percentage point increases in the likelihood of being in school post-15. The reduced form estimates are smaller in magnitude to those in Table 3.4 for British and unionist identities, around a third and a quarter of the size, respectively – though not statistically different due to the uncertainty in the BSA estimates. Unlike unionist identity in the QLFS, the BSA results are not statistically different from zero.

Table 3.10   First Stage IV, Reduced Form, OLS, and Second Stage IV Results (BSA)

	(1)	(2)	(3)	(4) Second
	First Stage	Reduced Form	OLS	Stage
Panel A - British Identity				
RoSLA	0.198***	-0.003		
	(0.017)	(0.020)		
Left School Post-15			0.054***	-0.017
			(0.012)	(0.103)
First Stage F				134.289
N	13,100	13,100	13,100	13,100
R-squared	0.109	0.026	0.028	0.024
Panel B - Unionist Identity				
RoSLA	0.211***	-0.005		
	(0.019)	(0.024)		
Left School Post-15			0.052***	-0.024
			(0.013)	(0.111)
First Stage F				122.079
N	10,545	10,545	10,545	10,545
R-squared	0.114	0.085	0.087	0.083

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The first stage F is the Kleibergen-Paap F Statistic. Each specification includes controls for age and age squared, survey year region of residence, gender and ethnicity.

The OLS estimates for the national identity outcomes in Table 3.10 are very similar to those in Table 3.4. Compared to the 0.068 and 0.075, for British and unionist identities respectively, the BSA estimates are 0.054 and 0.052, each significant at the five percent level. Column (4) the shows the second stage IV estimates. They are noisily estimated, though in each case the instrument is strong – as demonstrated in the Kleibergen-Paap F Statistic. In each case the effects are negative. For British identity, the estimate is around half the size of the QLFS result. The unionist identity coefficient is around a quarter of the QLFS coefficient, just as the reduced form result was. A reasonable reading of the results is that there is broad agreement between the two datasets in the causal IV specifications. But it remains the case that the QLFS results are not fully replicated in the BSA data and evidence of an impact on identity is not overwhelming.

3.8   Discussion and Conclusion

The results presented in this paper suggest that education can have an impact on national identity, though effects are not present for all groups. Using the change to compulsory schooling in the UK in 1972, a causal effect of education is identified on one of the two measures of national identity in the Quarterly Labour Force Survey. The result suggests that additional schooling makes one less likely to identify as British when compared to identifying with one of the UK's constituent nations. This is not immediately intuitive, based on the small amount of prior research that exists in this area. The explanation may be that there is some kind of alienation effect that occurs when individuals who may not otherwise have stayed in education are forced to. Heterogeneity analysis shows this effect is present for women and for those born in Scotland. The effects for these groups are large, negative twelve percentage points among women, and negative 20 percentage points among Scots, but for other groups are statistically indistinguishable from zero.

However, when turning to the same analysis conducted in the British Social Attitudes Survey, the result does not quite replicate, but shows a similar pattern. Magnitudes are generally smaller than in the QLFS and, though the direction of effects are generally the same and OLS estimates are statistically meaningful, statistical significance is not forthcoming for the IV specifications.

Previous work by other authors has suggested impacts of compulsory schooling changes on facets of identity, for example on Euroscepticism (Kunst, Kuhn, and van de Werfhorst, 2020), whilst others suggest compulsory schooling laws have been changed explicitly with the purpose of fostering a collective identity in the United States (Bandiera et al, 2019). Hungerman (2014) and Mocan and Pogorelova (2017) each examine religiosity, a similar form of individual identity to national identity, and find significant effects.

Future research might focus on the short-term effects of education on identity, as this paper can only look at long term effects where education's impact (if present) may have dissipated. Moreover, in terms of variables relating to national identity, we have only one indicator (transformed into two separate outcome variables). This enables analysis of the extensive margin of national identity. Other datasets do exist with more granular measures, such as the German Socio-Economic Panel, which asks how strongly German an individual feels giving an insight into the intensive margin of national identity. Analysis of national identity formation through peers could also be informative. This research was not able to make use of migrants as there were too few in the LFS; future research would do well to focus on this.

Appendix Table C.1   Heterogeneity by Nation of Current Residence (QLFS)

	(1)	(2)	(4)	(5)
	OLS	IV	OLS	IV
	British Identity		Unionist Identity	
Panel A - England				
Left School Post-15	0.071***	-0.011	0.080***	-0.076
	(0.005)	(0.046)	(0.006)	(0.049)
First Stage F		788.655		664.012
N	64,620	64,620	56,461	56,461
R-squared	0.020	0.017	0.020	0.006
Panel B - Scotland				
Left School Post-15	0.084***	-0.078	0.067***	-0.121
	(0.012)	(0.096)	(0.011)	(0.080)
First Stage F		128.105		125.756
N	8,104	8,104	7,223	7,223
R-squared	0.099	0.080	0.149	0.115
Panel C - Wales				
Left School Post-15	0.019	-0.211	0.017	-0.228
	(0.019)	(0.153)	(0.018)	(0.149)
First Stage F		61.931		50.449
N	4,222	4,222	3,727	3,727
R-squared	0.136	0.105	0.227	0.185

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The first stage F is the Kleibergen-Paap F Statistic. Each column includes the instrument and running variable as well as controls for year of birth, age and age squared, survey year interacted with survey quarter, region of residence, gender, ethnicity, and country of birth.

Appendix Table C.2   First Stage Estimates at Different Bandwidths (BSA)

	(1)	(2)	(3)	(4)	(5)	(6)
	Left School Post-15					
Bandwidth (Years)	1	2	3	4	5	6
RoSLA	0.207***	0.198***	0.192***	0.192***	0.190***	0.198***
	(0.019)	(0.041)	(0.028)	(0.022)	(0.019)	(0.017)
N	2,087	4,234	6,373	8,705	10,887	13,100
R-squared	0.085	0.091	0.092	0.099	0.106	0.109

Robust standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Each specification includes controls for age and age squared, survey year region of residence, gender and ethnicity.

Thesis Conclusion

The chapters presented in this thesis cover a range of topics in the economics of education. Each is policy relevant and provides a contribution to its part of the economics literature. In the first chapter I show, in work co-authored with my supervisor Ian Walker, that intrinsic religiosity at age 14 – measured by a variable we call faithfulness – is an important driver of some short- and long-term outcomes, namely GCSE attainment and future religious affiliation. More than that – it appears that it is important where faith schooling is not, and that interaction effects are all statistically zero. As faith schools congregate higher shares of religious students together, this finding is important. The perception that faith schools are better as institutions (at least in terms of the test scores they produce) appears misplaced, conditional on a wide range of school characteristics. In the second chapter I examine the Education Maintenance Allowance. Drawing on earlier work on EMA's pilot studies, I show that the benefits were of a similar magnitude at the end of EMA's life to the beginning in terms of increased attendance in higher education. As well as that, new findings are presented: EMA improved university attendance and reduced the likelihood of a person being in insecure work 8 years after they leave first receive the conditional cash transfer. In the context of an education system recovering from a global pandemic, EMA stands ready as a tool that policymakers should consider to help those from impacted cohorts who have now left or are about to leave compulsory education. Finally, my third chapter examines education's impact on identity. Schooling initially appears to have an impact in OLS specifications, but causal, IV estimates suggest otherwise. Effects do exist for some subgroups.

In the case of chapters 1 and 3, there is also a clear focus on identity which puts this thesis at the intersection of an established field and a relatively new area of economic inquiry. Beyond this, the methods employed are varied, ranging from simple linear regressions enhanced with

sensitivity analysis in chapter 1, to a conventional causal method in the form of an Instrumental Variables approach in chapter 3, to the use of machine learning for identifying heterogenous treatment effects in chapter 2. I have learnt and applied a wide range of skills during my studies. I hope to continue this in my future work.

The PhD process has not been uninterrupted, plain sailing; data access has been the major sticking point. The UK Data Service (UKDS) performs their job well and provide an essential service, without which this thesis would not have been possible, but are arguably understaffed and overworked. The *Next Steps* dataset took a little under a year from beginning to apply to the data access being granted. The LFS dataset application took closer to a year and a half. This led to chapter 3, originally conceived as chapter 2, which uses the LFS, being delayed. An idea was needed for chapter 2. It made sense to use the high-quality, secure access data that I already had access to – *Next Steps*. My desire to use new methods, not widely used in economics, led to the use of causal machine learning to examine EMA. This shows that I am able to innovate and respond to difficulties, but it might have been less stressful if such difficulties had not happened!

The work contained in this thesis has been presented at a range of events, demonstrating its relevance to academic audiences. These include the Royal Economic Society Conference 2021, the European Association of Labour Economists' Conference 2021, the International Workshop on Applied Education Economics 2021, and the Work, Pensions, and Labour Economics Study Group (WPEG) Conference 2021. I hope that this demonstrates the quality of my research; I intend to carry each chapter forward to publication.

References

Adamczyk, A. (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), 180–195. https://doi.org/10.1177/002214650905000205

Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, *115*(3), 715–753.

Akerlof, G. A., & Kranton, R. E. (2002). Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature*, *40*(4), 1167–1201.

Allen, R., & Vignoles, A. (2016). Can school competition improve standards? The case of faith schools in England. *Empirical Economics*, *50*(3), 959–973. https://doi.org/10.1007/s00181-015-0949-4

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005a). An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling. *The Journal of Human Resources*, *40*(4), 791–821.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005b). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, *113*(1), 151–184. https://doi.org/10.1086/426036

Anders, J. (2017). The influence of socioeconomic status on changes in young people's expectations of applying to university. *Oxford Review of Education*, *43*(4), 381–401. https://doi.org/10.1080/03054985.2017.1329722

Andrews, J., & Johnes, R. (2016). Faith Schools, Pupil Performance and Social Selection. *Education Policy Institute*, 1–46.

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science*, *485*(February), 483–485.

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7353–7360. https://doi.org/10.1073/pnas.1510489113

Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, *31*(2), 3–32. https://doi.org/10.1257/jep.31.2.3

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *Annals of Statistics*, *47*(2), 1179–1203. https://doi.org/10.1214/18-AOS1709

Athey, S., & Wager, S. (2019). Estimating Treatment Effects with Causal Forests: An Application. *Observational Studies*.

Attanasio, O. P., Meghir, C., & Santiago, A. (2012). Education choices in Mexico: Using a structural model and a randomized experiment to evaluate PROGRESA. *Review of Economic Studies*, *79*(1), 37–66. https://doi.org/10.1093/restud/rdr015

Bandiera, O., Mohnen, M., Rasul, I., & Viarengo, M. (2018). Nation-building through compulsory schooling during the age of mass migration. *The Economic Journal*, *129*(617), 62–109.

Battu, H., & Zenou, Y. (2010). Oppositional identities and employment for ethnic minorities: Evidence from England. *Economic Journal*, *120*(542), F52–F71. https://doi.org/10.1111/j.1468-0297.2009.02337.x

Becker, S. O., Nagler, M., & Woessmann, L. (2017). Education and religious participation: city-level evidence from Germany's secularization period 1890–1930. *Journal of Economic Growth*, *22*(3), 273–311. https://doi.org/10.1007/s10887-017-9142-2

Becker, S. O., & Woessmann, L. (2009). Was Weber wrong? A human capital theory of Protestant economic history. *The Quarterly Journal of Economics*, *124*(2), 531–596.

Benhassine, N., Devoto, F., Duflo, E., Dupas, P., & Pouliquen, V. (2015). Turning a shove into a nudge?: A "labeled cash transfer" for education. *American Economic Journal: Economic Policy*, *7*(3), 86–125. https://doi.org/10.1257/pol.20130225

Benjamini, Y., & Hochberg, Y. (1995). Controlling for the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.2307/2346101

Bertrand, M., Kamenica, E., & Pan, J. (2015). Gender identity and relative income within households. *Quarterly Journal of Economics*, *130*(2), 571–614. https://doi.org/10.1093/qje/qjv001

Bolton, P. (2011). Education Maintenance Allowance (EMA). In *Parliamentary Briefing Papers*. https://www.studentfinancewales.co.uk/fe/ema.aspx

Bommel, N., & Heineck, G. (2020). *Revisiting the Causal Effect of Education on Political Participation and Interest* (No. 13954; IZA Discussion Paper Series).

Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: Educational reform and the contradictions of economic life*. Basic Books.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1201/9780367816377-11

Britton, J., & Dearden, L. (2015). *The 16 to 19 bursary fund : impact evaluation* (Issue June).

Brown, S., & Taylor, K. (2007). Religion and education: evidence from the National Child Development Study. *Journal of Economic Behavior & Organization*, *63*(3), 439–460.

Buscha, F., & Dickson, M. (2018). A Note on the Wage Effects of the 1972 Raising of the School Leaving Age in Scotland and Northern Ireland. *Scottish Journal of Political Economy*, *65*(5), 572–582. https://doi.org/10.1111/sjpe.12187

Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2017). Curriculum and ideology. *Journal of Political Economy*, *125*(2), 338–392.

Centre for Longitudinal Studies. (2018). *Next Steps Sweep 8 - Age 25 Survey: User Guide* (L. Calderwood, Ed.; Second Edi).

Cesur, R., & Mocan, N. (2018). Education, religion, and voter preference in a Muslim country. *Journal of Population Economics*, *31*(1), 1–44.

Chowdry, H., Dearden, L., & Emmerson, C. (2008). *Education maintenance allowance evaluation with administrative data: the impact of the EMA pilots on participation and attainment in post-compulsory education. November*, 1–33. http://eprints.ucl.ac.uk/18324/

Clots-Figueras, I., & Masella, P. (2013). Education, language and identity. *The Economic Journal*, *123*(570), F332–F357.

Cohen-Zada, D., & Elder, T. (2009). Historical religious concentrations and the effects of Catholic schooling. *Journal of Urban Economics*, *66*(1), 65–74. https://doi.org/10.1016/j.jue.2009.04.002

Cohen-Zada, D., & Sander, W. (2008). Religion, religiosity and private school choice: Implications for estimating the effectiveness of private schools. *Journal of Urban Economics*, *64*(1), 85–100. https://doi.org/10.1016/j.jue.2007.08.005

Davis, J. M. V., & Heller, S. B. (2020). Rethinking the benefits of youth employment programs: The heterogeneous effects of summer jobs. *Review of Economics and Statistics*, *102*(4), 664–677. https://doi.org/10.1162/rest_a_00850

de Luca, G., Magnus, J. R., & Peracchi, F. (2019). Comments on "Unobservable Selection and Coefficient Stability: Theory and Evidence" and "Poorly Measured Confounders are More Useful on the Left Than on the Right." In *Journal of Business and Economic Statistics* (Vol. 37, Issue 2, pp. 217–222). American Statistical Association. https://doi.org/10.1080/07350015.2019.1575743

Dearden, L., Emmerson, C., Frayne, C., & Meghir, C. (2009). Conditional cash transfers and school dropout rates. *Journal of Human Resources*, *44*(4), 827–857. https://doi.org/10.3368/jhr.44.4.827

Dearden, L., & Heath, A. (1996). Income Support and Staying in School: What Can We Learn from Australia AUSTUDY Experiment? *Fiscal Studies*, *17*(4), 1–30.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics*, *132*(4), 1593–1640. https://doi.org/10.1093/qje/qjx022

Dynarski, S. M. (2003). American Economic Association Does Aid Matter ? Measuring the Effect of Student Aid on College Attendance and Completion. *The American Economic Review*, *93*(1), 279–288.

Evans, W. N., & Schwab, R. M. (1995). Finishing High School and Starting College: Do Catholic Schools Make a Difference? *The Quarterly Journal of Economics*, *110*(4), 941–974. https://doi.org/10.2307/2946645

Feinstein, L., & Sabatés, R. (2005). Education and Youth Crime: Effects of Introducing the Education Maintenance Allowance Programme. In *Wider Benefits of Learning Report No.14*. https://doi.org/10.2139/ssrn.901421

Fiszbein, A., & Schady, N. R. (2009). Conditional Cash Transfers: Reducing Present and Future Poverty. In *World Bank Report*. https://doi.org/10.1596/978-0-8213-7352-1

Fletcher, M. (2000). *Education Maintenance Allowances: The Impact on Further Education*. http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED441174&site=ehost-live

Fouka, V. (2019). How do immigrants respond to discrimination? The case of Germans in the US during World War I. *American Political Science Review*, *113*(2), 405–422.

Fruehwirth, J. C., Iyer, S., & Zhang, A. (2019). Religion and Depression in Adolescence. *Journal of Political Economy*, *127*(3), 1178–1209.

Gihleb, R., & Giuntella, O. (2017). Nuns and the effects of catholic schools. Evidence from Vatican II. *Journal of Economic Behavior and Organization*, *137*, 191–213. https://doi.org/10.1016/j.jebo.2017.03.007

Gitter, S. R., & Barham, B. L. (2009). Conditional cash transfers, shocks, and school enrolment in Nicaragua. *Journal of Development Studies*, *45*(10), 1747–1767. https://doi.org/10.1080/00220380902935857

Giusta, M. della, Hashimzade, N., & Myles, G. D. (2017). Schooling and the intergenerational transmission of values. *Journal of Public Economic Theory*, *19*(1), 1–17.

Glaeser, E. L., & Sacerdote, B. I. (2008). Education and Religion. *Journal of Human Capital*, *2*(2), 188–215.

Glewwe, P., & Kassouf, A. L. (2012). The impact of the Bolsa Escola/Familia conditional cash transfer program on enrollment, dropout rates and grade promotion in Brazil. *Journal of Development Economics*, *97*(2), 505–517. https://doi.org/10.1016/j.jdeveco.2011.05.008

Greenberg, D., Dechausay, N., & Fraker, C. (2011). *Learning Together: How Families Responded to Education Incentives in New York City's Conditional Cash Transfer Program*.

Gruber, J. 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.

Holford, A. (2015). The labour supply effect of Education Maintenance Allowance and its implications for parental altruism. In *Review of Economics of the Household* (Vol. 13, Issue 3). Springer US. https://doi.org/10.1007/s11150-015-9288-7

Hoxby, C. M. (1994). *Do Private Schools Provide Competition for Public Schools?* (Working Paper No.4978). National Bureau of Economic Research.

Humlum, M. K., & Vejlin, R. M. (2013). The Responses of Youth to a Cath Transfer Conditonal on Schooling: A Quasi-Experimental Study. *Journal of Applied Econometrics*, *28*, 628–649. https://doi.org/10.1002/jae

Hungerman, D. M. (2014a). Do religious proscriptions matter? Evidence from a theory-based test. *Journal of Human Resources*, *49*(4), 1053–1093. https://doi.org/10.3368/jhr.49.4.1053

Hungerman, D. M. (2014b). The effect of education on religion: Evidence from compulsory schooling laws. *Journal of Economic Behavior & Organization*, *104*, 52–63. https://doi.org/10.1080/07350015.2016.1227711

Iannaccone, L. R. (1998). Introduction to the Economics of Religion. *Journal of Economic Literature*, *36*(3), 1465–1495.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), 5–86.

Iyer, S. (2016). The new economics of religion. *Journal of Economic Literature*, *54*(2), 395–441.

Kremer, M., & Sarychev, A. (2000). Why do governments operate schools? *Harvard University: Cambridge, Mass. Available on Line at Http://Post. Economics. Harvard. Edu/Faculty/Kremer/Papers. Html. Processed*.

Kunst, S., Kuhn, T., & van de Werfhorst, H. G. (2020). Does education decrease Euroscepticism? A regression discontinuity design using compulsory schooling reforms in four European countries. *European Union Politics*, *21*(1), 24–42. https://doi.org/10.1177/1465116519877972

Lehrer, E. (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), 203–219. https://doi.org/10.1023/b:reho.0000031614.84035.8e

Long, R., & Danechi, S. (2019). Faith Schools in England: FAQs. *House of Commons Briefing Paper*, *06972*, 1–22.

McAndrew, S., & Voas, D. (2011). Measuring Religiosity Using Surveys. In *SURVEY QUESTION BANK: Topic Overview* (Vol. 4, Issue February).

McCullough, M. E., & Willoughby, B. L. B. (2009). Religion, self-regulation, and self-control: Associations, explanations, and implications. *Psychological Bulletin*, *135*(1), 69.

McFarland, M. J., Wright, B. R. E., & Weakliem, D. L. (2011). Educational attainment and religiosity: Exploring variations by religious tradition. *Sociology of Religion: A Quarterly Review*, *72*(2), 166–188. https://doi.org/10.1093/socrel/srq065

Mendolia, S., Paloyo, A. R., & Walker, I. (2018). Heterogeneous effects of high school peers on educational outcomes. *Oxford Economic Papers*, *70*(3), 613–634.

Mendolia, S., Paloyo, A., & Walker, I. (2019). Intrinsic Religiosity, Personality Traits, and Adolescent Risky Behaviors. *B.E. Journal of Economic Analysis and Policy*, *19*(3), 1–16. https://doi.org/10.1515/bejeap-2018-0311

Middleton, S., Maguire, S., Ashworth, K., Legge, K., Allen, T., Perrin, K., Battistin, E., Dearden, L., Emmerson, C., Fitzsimons, E., & Megir, C. (2004). *The evaluation of*

Education Maintenance Allowance Pilots: three years' evidence: a quantitative evaluation. http://eprints.ucl.ac.uk/18468/

Millán, T. M., Barham, T., Macours, K., Maluccio, J. A., & Stampini, M. (2019a). Long-term impacts of conditional cash transfers: Review of the evidence. *World Bank Research Observer*, *34*(1), 119–159. https://doi.org/10.1093/wbro/lky005

Millán, T. M., Barham, T., Macours, K., Maluccio, J. A., & Stampini, M. (2019b). Long-term impacts of conditional cash transfers: Review of the evidence. In *World Bank Research Observer* (Vol. 34, Issue 1, pp. 119–159). Oxford University Press. https://doi.org/10.1093/wbro/lky005

Milligan, K., Moretti, E., & Oreopoulos, P. (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics*, *88*(9–10), 1667–1695.

Morais de Sa e Silva, M. (2015). Conditional cash transfers and improved education quality: A political search for the policy link. *International Journal of Educational Development*, *45*, 169–181. https://doi.org/10.1016/j.ijedudev.2015.09.003

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106. https://doi.org/10.1257/jep.31.2.87

Natcen Social Research. (2021). British Social Attitudes Survey. In *[data collection]. UK Data Service. SN: 8772*.

Neal, D. (1997). The Effects of Catholic Secondary Schooling on Educational Achievement. *Journal of Labor Economics*, *15*(1, Part 1), 98–123. https://doi.org/10.1086/209848

Office for National Statistics, Social Survey Division, Northern Ireland Statistics and Research Agency, Central Survey Unit, . (2021). Quarterly Labour Force Survey, 1992-2020: Secure Access. *[Data Collection]. 20th Edition. UK Data Service. SN: 6727*. https://doi.org/http://doi.org/10.5255/UKDA-SN-6727-21

Peruffo, M., & Ferreira, P. C. (2017). the Long-Term Effects of Conditional Cash Transfers on Child Labor and School Enrollment. *Economic Inquiry*, *55*(4), 2008–2030. https://doi.org/10.1111/ecin.12457

Porter, C., & Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, *12*(3), 226–254. https://doi.org/10.1257/app.20180426

Riccio, J., Dechausay, N., Greenberg, D., Miller, C., Rucks, Z., & Verma, N. (2010). *Toward Reduced Poverty Across Generations: Early Findings from New York City's Conditional Cash Transfer Program*. http://ssrn.com/abstract=1786981

Riccio, J., & Miller, C. (2016). *NEW YORK CITY'S FIRST CONDITIONAL CASH TRANSFER PROGRAM What Worked, What Didn't*. http://ssrn.com/abstract=2821765

Sander, W. (2001). *The Effects of Catholic Schools on Religiosity, Education, and Competition* (No. 32). www.ncspe.org

Schaltegger, C. A., & Torgler, B. (2010). Work ethic, Protestantism, and human capital. *Economics Letters*, *107*(2), 99–101.

Schüller, S. (2015). Parental ethnic identity and educational attainment of second-generation immigrants. *Journal of Population Economics*, *28*(4), 965–1004.

Schultz, T. P. (2004). School subsidies for the poor: Evaluating the Mexican Progresa poverty program. *Journal of Development Economics*, *74*(1), 199–250. https://doi.org/10.1016/j.jdeveco.2003.12.009

SCOTUSblog. (2020). *Espinoza v. Montana Department of Revenue*. https://www.scotusblog.com/case-files/cases/espinoza-v-montana-department-of-revenue/

Silveus, N., & Stoddard, C. (2020). Identifying the causal effect of income on religiosity using the Earned Income Tax Credit. *Journal of Economic Behavior and Organization*, *178*, 903–924. https://doi.org/10.1016/j.jebo.2020.08.022

Spenkuch, J. L. (2017). Religion and work: Micro evidence from contemporary Germany. *Journal of Economic Behavior & Organization*, *135*, 193–214.

Spenkuch, J. L., & Tillmann, P. (2018). Elite Influence? Religion and the Electoral Success of the Nazis. *American Journal of Political Science*, *62*(1), 19–36. https://doi.org/10.1111/ajps.12328

Squicciarini, M. P. (2020). *Devotion and Development : Religiosity, Education, and Economic Progress in 19th-Century France*. *110*(11), 3454–3491.

Stone, J. (2017, August 7). Brexit caused by low levels of education, study finds. *Independent*.

Sullivan, A., Parsons, S., Green, F., Wiggins, R. D., Ploubidis, G., & Huynh, T. (2018). Educational attainment in the short and long term: was there an advantage to attending faith, private, and selective schools for pupils in the 1980s? *Oxford Review of Education*, *44*(6), 806–822. https://doi.org/10.1080/03054985.2018.1481378

Tibshirani, J., Athey, S., Friedberg, R., Hadad, V., Hirshberg, D., Miner, L., Sverdrup, E., Wager, S., & Wright, M. (2020). *grf: Generalised Random Forests. R Package* (1.2.0).

Torgler, B. (2006). The importance of faith: Tax morale and religiosity. *Journal of Economic Behavior and Organization*, *61*(1), 81–109. https://doi.org/10.1016/j.jebo.2004.10.007

University College London, UCL Institute of Education, & Centre for Longitudinal Studies. (2021a). Next Steps: Linked Education Administrative Datasets (National Pupil Database), England, 2005-2009: Secure Access. In *[data collection]. 6th Edition. UK Data Service. SN: 7104*.

University College London, UCL Institute of Education, & Centre for Longitudinal Studies. (2021b). Next Steps: Sweeps 1-8, 2004-2016. . In *[data collection]. UK Data Service. SN: 5545*.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28. https://doi.org/10.1257/jep.28.2.3

Voigtländer, N., & Voth, H.-J. (2015). Nazi indoctrination and anti-Semitic beliefs in Germany. *Proceedings of the National Academy of Sciences*, *112*(26), 7931–7936.

Wadsworth, A. A., & Walker, J. K. (2017). Religiosity and the Impact of Religious Secondary Schooling. *Journal of School Choice*, *11*(1), 131–147. https://doi.org/10.1080/15582159.2016.1257900

Wager, S., & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, *113*(523), 1228–1242. https://doi.org/10.1080/01621459.2017.1319839

Weber, M. (2001). *The Protestant Ethic and the Spirit of Capitalism*.

West, M., & Woessmann, L. (2010). 'Every Catholic Child in a Catholic School': Historical Resistance to State Schooling, Contemporary Private Competition and Student Achievement across Countries. *The Economic Journal*, *120*(546), F229–F255.

Zhang, A. (2018). New Findings on Key Factors Influencing the UK's Referendum on Leaving the EU. *World Development*, *102*, 304–314. https://doi.org/10.1016/j.worlddev.2017.07.017