Labour Market Returns to Vocational Education in the Presence of Multiple Alternatives

Sönke Hendrik Matthewes* Guglielmo Ventura†

January 23, 2022

DRAFT, NOT FOR CIRCULATION.

Abstract

Across many countries there is a growing interest in the expansion of vocational programmes to boost skills and reduce youth unemployment. However, disagreement remains about whether vocational routes are beneficial for students' labour market outcomes. Critics argue that the specific human capital acquired in vocational tracks might quickly become obsolete in the face of rapidly changing labour markets. Proponents argue that vocational education increases the employability of less academically inclined students who are at risk of dropping out of academic tracks or who fail to complement their academic secondary education with higher education. We aim to reconcile these opposing views by empirically decomposing the net effect of vocational education on earnings into the effect for students at the margin between vocational and academic education and the effect for students at the margin between vocational education and no further education. To estimate margin-specific complier treatment effects we leverage an identification approach based on multiple instrumental variables together with rich linked administrative education and earnings data from England, where students have to choose between no further education, a vocational and an academic upper secondary track at the age of 16. We find that vocational education boosts earnings at age 30 for students who choose it as an alternative to no post-16 education, mainly through extensive margin effects, but decreases earnings for a majority of (male) students who are diverted from the academic track, through wage and intensive margin effects. The negative results at the vocational vs. academic education margin cannot be explained by differential access to higher education. Our results caution against an across-the-board expansion of vocational upper secondary education.

*University of Potsdam; WZB Berlin Social Science Center; Centre for Vocational Education Research, London School of Economics; . *Email*: soenke.matthewes@wzb.eu.

†Centre for Economic Performance and Centre for Vocational Education Research, London School of Economics; University College London. *Email*: g.ventura@lse.ac.uk.

1 Introduction

As demand for skills in the labour market is rising, it is increasingly difficult for young workers without tertiary education to secure stable and well-paying jobs (e.g. Autor, 2019). Youth unemployment and underemployment levels are high, while firms lament the lack of skilled workers in technical occupations (OECD, 2017). Many herald improving and expanding vocational courses in secondary education as a means to boost skills and, consequently, employability and earnings of the non-tertiary educated. Recently, policies to improve and expand vocational education have gained momentum also in countries with weak traditions of providing high quality vocational education and training to its secondary students, like the US and UK. However, considerable disagreement remains on whether vocational routes really benefit young people economically.

This debate is often framed in terms of a dichotomy between vocational (also known as technical) and academic (also known as general) education and revolves around the types of skills these provide and their value on the labour market. The main benefit of an academic curriculum lies in equipping students with general knowledge and analytical skills that are transferable across occupations (Goldin, 2001). Compared to the more occupation-specific curricula taught in vocational tracks, this might help students on the labour market through increased flexibility, especially in the long run (Hanushek *et al.*, 2017). This argument is particularly cogent in the face of rapidly changing labour demand due to rising automation and globalisation.

To these points, advocates of vocational education wield two main counter-arguments: first, the general skills taught in academic tracks might, in fact, be too generic to be readily deployable on labour markets unless complemented with tertiary education, which far from all students pursue (Bertrand *et al.*, 2019). Second, the abstract nature of learning in academic tracks might disengage less academically inclined students, leaving them at risk of dropping out of secondary education altogether (Hall, 2016). Across OECD countries 15% of 25–34 year-olds have not completed secondary education, indicating that drop-out is indeed an important problem.

Given the plausibility of both of these mechanisms, the labour market effects of vocational education are ultimately an empirical question. However, conclusive empirical evidence is lacking because selection problems plague comparisons of vocationally educated and other students and natural experiments are rare (Shavit *et al.*, 1998). While there is a consensus that vocational education, through establishing links to specific occupations, facilitates school-to-work transitions, the evidence on longer run outcomes is more mixed and less positive (Ryan, 2001). In an influential study, Hanushek *et al.* (2017) compare employment/earning-age profiles of vocationally and academically educated students across multiple countries and find that initial advantages for the vocationally educated reverse over time, lending support to the skill arguments of vocational education critics. However, two recent studies with highly credible quasi-experimental research designs find that vocational education can boost medium and long run labour market outcomes for particular types of students (Bertrand *et al.*, 2019; Silliman and Virtanen, 2019).

In this paper we aim to reconcile the seemingly conflicting theoretical predictions and empirical findings by looking at the problem through the lens of effect heterogeneity in an unordered choice framework. Returns to different educational choices are likely to vary substantially, not only based on students' abilities and inclinations but also on their access to different alternatives. For example, a student might benefit greatly from vocational education if her alternative is to drop out of school (e.g. through disinterest in academic subjects or inability to cope with the scholastic demands of academic education). At the same time, a student whose alternative to vocational education is academic upper secondary school and subsequently the completion of a university degree might lose out from choosing a vocational track, especially in the longer run. These examples suggest, first, that advocates and critics of vocational education could both be right but are talking about different students and, second, that returns might crucially depend on students' counterfactual education choice. Identifying this effect heterogeneity is crucial for policy design as it can inform the optimal targeting of vocational programmes beyond simple calls for expansion or contraction.

We study returns to vocational education in England, where at the age of 16, after the completion of comprehensive compulsory education, students need to choose between three alternatives: (i) conclude their education and enter the labour market directly, (ii) pursue upper secondary education in the vocational track or (iii) pursue upper secondary education in the academic track.[1] This setting allows us to separately consider the effects of vocational education *vs.* academic education and *vs.* no post-16 education. Thereby, we uncover patterns of effect heterogeneity that are hidden in most other contexts but that are crucial for coming to a comprehensive judgement of vocational education as they relate directly to the above-mentioned theoretical arguments.

To identify the two alternative-specific effects of interest, net of self-selection into educational tracks, we rely on an identification approach based on multiple instrumental variables (IVs) proposed by Mountjoy (2019). To construct the IVs, we exploit the fact that upper secondary educational tracks in England are linked to specific institutions: the vocational track is offered by Further Education (FE) colleges and the academic track is offered by integrated secondary schools and designated Sixth Form (SF) colleges. Focusing on students from non-integrated schools who need to switch institution regardless of which track they choose, we can construct two alternative-specific IVs by computing students' geographical distance to the nearest vocational (FE) college and the nearest academic (SF) college. Thus equipped, Mountjoy's (2019) method allows us to separately identify causal returns to vocational education for students at the margin between academic and vocational education and for students at the margin between vocational and no post-16 education. Our estimated margin-specific complier treatment effects speak directly

[1] Since 2013, students need to stay on in some form of (part-time) education until 17. In 2015, this was raised to age 18. However, as post-16 students can get educated in many different institutions (including in work-based forms of education), it is unclear how enforceable this law is (Hupkau *et al.*, 2017). Regardless, for the cohorts studied here leaving the education system at age 16 was still possible.

to the effects of expanding access to vocational colleges, but offer a more nuanced view than would be possible with conventional methods that are only able to identify the net effect of such a policy.

For estimation we leverage unique education administrative data linked to tax records that allow us to follow two full cohorts of state-school educated pupils in England through their school careers, post-16 upper secondary and tertiary education and into the labour market. We record our two main outcomes of interest, employment and annual earnings, when students are 29–30 years of age. To construct the two required distance instruments we use geospatial information on all education institutions in England and students' home addresses. To account for the non-random location of post-16 education providers, next to student- and school-level controls, we directly control for distance to local economic centre and detailed measures of neighbourhood quality, as well as region fixed effects to ensure that we compare students from similar neighbourhoods who face similar labour market conditions when they make their post-16 education choices. Identification stems from conditional variation in the distance to vocational college, holding constant distance to academic college, and *vice-versa*, which is much more plausibly exogenous than distance to any kind of education provider (Mountjoy, 2019). In balance tests we show that our instruments are empirically balanced across a range of student characteristics, including nationally administered achievement tests at age 11.

Descriptively, we find an initial earnings advantage for students that enter the labour market directly which quickly reverses over time. Students who choose vocational education catch up with their less educated peers by age 19 and are consistently better off thereafter. Academically educated students, who have the lowest earnings initially, catch up with both groups by the age of 23, after which their advantage grows steadily. By the ages 29–30, annual earnings of vocational-track students are about £4,000 higher than those of students without any upper secondary education but £6,400 lower than those of academic-track students. Flexibly controlling for detailed student demographics, previous achievement and school and neighbourhood characteristics roughly halves the raw differences; still, they remain substantial.

Instrumenting vocational track attendance with students' geographical distance to their closest vocational college in a classical IV framework, we find a positive net effect of vocational-track attendance on earnings at ages 29–30 of about £2,500 for compliers, i.e. students at the margin of choosing vocational education over one of the other two options. Using the multiple-instrument identification approach, we show that this net effect is composed of opposite effects along the two complier margins: students who otherwise would have chosen the academic track (80 percent of compliers) suffer from earnings losses of roughly £1,000, whereas students who otherwise would have quit education (20 percent of compliers) experience very large gains of about £17,000, which are quite imprecisely estimated but explain the positive effect on net. The effect of vocational *vs.* no post-16 education is, for a large part, driven by extensive margin (i.e. employment) effects, whereas the effect *vs.* academic education is driven by wage (and intensive margin) effects.

Splitting the results by gender, we find that both margins are primarily driven by male students. They make up the vast majority of compliers at the no post-16 education margin and only for them the large positive effect on earnings is statistically significant, though estimates for female students are qualitatively similar. In contrast, at the academic education margin, where the complier population is split roughly evenly between males and females, results differ markedly by gender: we find a substantial negative effect of vocational education on the earnings of male students of about £3,000, or 15%, but a zero effect for females. To understand the mechanisms underlying this result, we study the effects of vocational track attendance on a number of intermediary education outcomes: we find no significant effect on upper secondary degree completion (i.e. a Level 3 qualification) or 3-year university degree completion for either gender. Yet, only for female students, who make more use of the built-in flexibilities of the vocational track and study a high share of academic courses regardless of track choice, this seems to suffice to guard against negative earnings effects of vocational education.

Therefore, our results suggest that vocational education is a double-edged sword: highly beneficial for students who are discouraged by the academic track and otherwise would leave school at 16 but detrimental for (male) students who otherwise would pursue an academic education. As the first study to uncover this effect heterogeneity empirically, our findings can reconcile ambiguous previous empirical evidence and the opposing theoretical arguments by vocational education advocates and critics. On the one hand, our results confirm advocates' claim that for less academically inclined students vocational education confers large labour market benefits, mainly by channelling students into employment. On the other hand, they reinforce critics' concerns that for more academically inclined students vocational education has detrimental career effects, as students are channelled into lower-wage jobs.

Contrary to widely held beliefs, this is not due to differential access to higher education. We show, first, that that the majority of vocational-academic compliers does not pursue higher education *regardless of their track choice* and, second, that the English vocational track does not close the door to university for students who aspire to pursue a degree. Accordingly, the negative earnings effect must stem from problems specific to the skills acquired in – or signals conveyed by – the upper secondary vocational track. Well known problems of English vocational education include an overwhelmingly wide course offer with (too) many narrow qualifications, hard to navigate for students and without clear recognition among employers, the absence of good career guidance and a lack of clear progression routes into higher *vocational* education and work-based learning opportunities, such as apprenticeships (Musset and Field, 2013; Wolf, 2011). All these factors have contributed towards deep-rooted negative perceptions about vocational education in Britain, so that, graduation from a vocational college might also be read as a negative signal by employers (Wolf, 2011).

Our findings have important implications for policy – especially so, since our instrumental variables, which represent students' distance to different post-16 education providers, map directly into expansion (or contraction) policies. They suggest that an across-the-board expansion of vocational education opportunities, which is likely to switch a population of students into choosing the vocational track that resembles our group of compliers, would reduce earnings for the majority of affected students who are diverted from the academic track and the associated higher earnings. In contrast, designing policies targeted at students at risk of dropping out seems to be a promising avenue for education policy as it can boost earnings for a group of students that is particularly disadvantaged.

The paper is structured as follows: section 2 lays out the English institutional context. Section 3 describes our data sources and presents descriptive findings. Section 4 discusses identification challenges and presents our research design. Section 5 assesses the validity of our identification assumptions. Section 6 present the estimation results. Finally, section 7 discusses implications and concludes.

2 Institutional Context

In England, students start school after turning 5 and remain in compulsory education until the age of 16. This phase is divided into six years of primary school and five years of comprehensive lower secondary school, during which students follow a common nationally defined curriculum. Compulsory education concludes with the General Certificate of Secondary Education (GCSE) examinations, which students typically take in eight to ten subjects. These are high-stakes exams that influence which courses they can enter in upper secondary education and are sometimes even taken into account for university admission and job applications. After their GCSEs, students have to choose among three alternatives: (i) conclude their education and enter the labour market directly, (ii) pursue upper secondary education in the vocational track or (iii) pursue upper secondary education in the academic track.[2]

Entry to the academic track typically requires that a student obtains at least five GCSEs at grade C or higher. It lasts two years, during which students study towards academic qualifications known as A-Levels, which are the traditional prerequisite for university entrance. These two years of academic upper secondary education are referred to as 'sixth form'. They are offered by integrated secondary schools that have their own sixth form (which thus integrate lower and upper secondary schooling) and by designated Sixth Form Colleges. Virtually all students from an integrated secondary school who pursue the academic track attend their own school's sixth form. Students from non-integrated secondary schools who pursue the academic track generally enrol in Sixth Form Colleges, of which there are 94 across England (they cater to more than 90% of this group).

[2]Since 2013, students need to stay on in some form of (part-time) education until 17. In 2015, this was raised to age 18. However, as post-16 students can get educated in many different institutions, it is unclear how enforceable this law is (Hupkau *et al.*, 2017). In any case, for the cohorts studied here leaving the education system at age 16 was still possible.

The fact that secondary schools in England can be divided into integrated and non-integrated schools has important implications for our study. First, integrated schools attract more students who intend to pursue the academic track from an early age. Accordingly, fewer of them consider vocational routes: while 58% of students from non-integrated schools enrol in the vocational track, only 38% from integrated schools do. Second, and most importantly for our purposes, students from integrated schools do not have to switch institutions at the age of 16 to attend the academic track. This would be a problem for our instrumental variables strategy because it relies on differences in distance to different education providers playing a (marginal but non-negligible) role for students' educational choices. For these reasons, in our analysis we focus on the population of students who attend a non-integrated lower secondary school, for whom the choice between academic and vocational education is more salient and more symmetric in terms of the role that distance might play.[3]

In contrast to the academic track and to other vocational education systems, which usually comprise a limited and well-defined number of programmes, the vocational track in England is much less harmonised. Students who follow the vocational track can choose from a plethora of (often hyper-specialised) vocational courses at three different levels: Level 3, 2 and 1. Level 3 qualifications are equivalent to A-Levels, have similar entry requirements in terms of GCSEs, are taught full-time, mostly as two-year courses and also count towards university admission.[4] For students that do not meet the entry requirements for Level 3, there are less demanding vocational qualifications at Level 2 (which count as equivalent to GCSEs) and Level 1 (which count as equivalent to primary school education). At each level the number of of courses to choose from is very large: at Level 3 alone, there are more than 3,700 different qualifications (Hupkau *et al.*, 2017). The vast majority of vocational courses for 16–18-year-old learners are classroom-based; apprenticeships that include workplace training typically only start after completion of a classroom-based qualification.

About 80% of vocational-track students attend at Further Education (FE) Colleges, of which there are 247 across England (Hupkau and Ventura, 2017).[5] Like community colleges in the US, these vocational institutions were historically established to offer training to adult learners. While this function remains important, over the last decades FE Colleges have increasingly shifted their focus to vocational-track students coming straight out of secondary school. Next to vocational courses, FE colleges also offer courses in basic and soft skills (such as employability or

[3]An added benefit from this sample restriction is that we compare vocational-track students on Further Education (FE) colleges (see below) with academic-track students on Sixth Form (SF) colleges but (save for a few exceptions) not with those on schools. FE and SF colleges both only cater to students above 16 and have similar governance and funding structures, whereas schools operate very differently. Hence, this way we minimise the influence of institutional differences on our treatment definitions so that they capture only the type of education students follow: academic *vs.* vocational.

[4]Though students with Level 3 vocational qualifications usually face more restrictions regarding the degrees and universities they can enter than with A-Levels.

[5]The remaining share of vocational learners are trained by private training providers, like large firms, and other publicly funded providers, like local authorities, both of which are numerous but small.

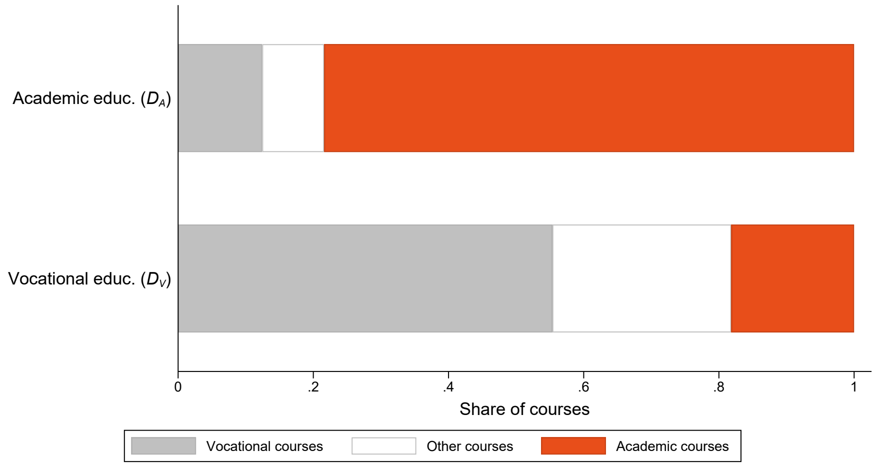


Figure 1. Course contents by educational track.

Notes: Shares are constructed considering all courses and modules (of more than one month of length) studied within 24 months of the relevant enrolment. We weight courses based on their officially recommended hours of studying ('guided learning hours'). We adopt the following classification: A-Levels and GCSE qualifications are classified as Academic; non-A-Levels qualifications in vocational subjects are classified as Vocational; qualifications whose recorded subject is 'Preparation for Work and Life' (such as qualifications known as Key Skills or Functional Skills) are classified as Other.

communication skills) and some academic courses (remedial courses in English and maths but also A-Levels). Therefore, students enrolled in the vocational track typically study a more mixed curriculum than students in the academic track. Figure 1 shows the average shares of different types of courses for the two tracks: while vocational students spend more than half of their time studying vocational subjects they also spend more time studying modules in basic and soft skills than academic students who predominantly study academic subjects.

3 Data and Descriptive Results

3.1 Data Sources and Variables

We use a unique ensemble of administrative datasets from England, known as Longitudinal Education Outcomes (LEO), to follow two full cohorts of state-school educated pupils in England through their school careers, post-16 upper secondary and tertiary education and into the labour market up until the age of 30. The two cohorts we focus on took their GCSEs in the academic years 2001/02 and 2002/03. Even though data for more recent cohorts is available, we chose to follow the earliest two cohorts for which we observe all the required information for our analyses in order to maximise the time we can follow students into the labour market.

Our initial sample definition is based on the pupil census of the National Pupil Database (NPD), which reports information on the universe of students enrolled in state-funded schools in England.[6] For the two abovementioned academic years, we retain all students in the final year before their GCSE exams (year group 11) to define our cohorts of interest. Student-level demographic characteristics reported in the pupil census include binary variables for gender, ethnicity (White British or other), special educational need (SEN), language spoken at home (English or other) and free school meal (FSM) eligibility, as a proxy for socio-economic status. We link the pupil census to the NPD's exam data to record students' scores in standardised national end of primary school tests in English, Maths and Science (Key Stage 2 exams) and the overall score they obtained in their GCSE exams at the end of lower secondary school. All test scores are standardised to mean zero and standard deviation one within cohorts. In order to measure secondary school quality, we calculate school-level averages of the three Key Stage 2 (KS2) test scores and the share of FSM eligible, White British and English as a second language students.

Our use of distance instruments makes it paramount to control for residential sorting (e.g. Spiess and Wrohlich, 2010). Accordingly, on top of the student- and school-level covariates, we compile a third set of neighbourhood-level controls. The pupil census contains a fairly precise measure of students' residential location, namely an identifier for the Lower Layer Super Output Area (LSOA). LSOAs divide the whole surface of England into about 33,000 small geospatial units of 1,000–1,500 inhabitants each. Using the LSOA identifier, we can merge in eight so-called Indices of Deprivation (IoD), constructed by British Ministry of Housing, Communities and Local Government, to measure neighbourhood quality at the LSOA-level in the following domains: income deprivation; employment deprivation; education, skills and training deprivation; health deprivation and disability; crime; barriers to housing and service; living environment deprivation; and income deprivation affecting children.

To construct the distance instruments required by our identification strategy, we proxy a student's home address with the population-weighted centroid of her LSOA. In combination with a registry of all state-funded educational institutions in England, we then calculate the ellipsoidal distance in kilometres between a student's (proxied) residential address and the closest Further Education College and Sixth Form College, respectively. Henceforth, we will refer to these two instruments as distance to vocational college, Z_V, and distance to academic college, Z_A. Even though we observe detailed measures of neighourhood quality, concerns about educational institutions concentrating in local centres and this confounding our instruments might remain. Therefore, we also calculate the distance between students' home address and the nearest local economic centre and add this variable to our control set.

The three exhaustive and mutually exclusive treatments of interest are starting at an academic institution (i.e. Sixth Form College or a school's sixth form), starting at a vocational institution (i.e. Further Education College or other vocational education provider) and not enrolling in any upper secondary education after completing compulsory education. In order to observe all post-16 educational choices, we link the NPD data with the Individualised Learner Records (ILR),

[6] State-funded schools comprise 93% of the total English student population.

a dataset which covers all publicly-funded education and training activities. We thus observe any enrolment at a school's sixth form in the NPD and any enrolment at a Sixth Form College, Further Education College or other private or public vocational education provider in the ILR. Equipped with this information, we define treatment by the institution type of students' first observed enrolment, if any, within a two-year window after finishing their GCSEs.[7]

Our main outcomes relating to students' labour market performance come from Her Majesty's Revenue and Customs (HMRC) tax records, which we can link to our student data for the tax years 2004 to 2017. This dataset reports earnings spells for all employed individuals in England. From 2014 onwards, we additionally observe earnings from self-employment. We sum the earnings accruing from all employment spells and from self-employment in a given year and deflate by the annual UK consumer price index (base year 2010) to construct a measure of real annual earnings for each student. As this dataset covers all earnings (except for earnings from abroad and from the informal sector), we define employment as having positive earnings in a given year. We average the annual earnings and employment variables within persons over ages 29–30 and, in the case of earnings, winsorise at the 99.5^{th} percentile, to arrive at our two main outcomes of interest. We include observations with zero earnings, whether unemployed or inactive, throughout our analysis.[8] To study dynamics, we also construct an annual panel of employment and (winsorised) earnings at each observed age, for which we are forced to exclude self-employment to ensure comparability across the whole age range.

We construct a number of intermediate educational outcomes to get a better understanding of the mechanisms that might underlie potential labour market effects of track choice. First, using the NPD and ILR, we calculate the share of vocational *vs.* academic content students study across all their courses during their first two years of post-compulsory education (see Figure 1). Second, we construct an indicator for whether students achieve a Level 3 qualification, i.e. A-Levels or an equivalent vocational qualification, which is the stated goal of both tracks and widely considered an important stepping-stone. Third, using the ILR, we construct an indicator for whether students ever start an apprenticeship, as many vocational courses aim at channelling students into apprenticeships and they are generally considered to yield returns on the labour market (Cavaglia *et al.*, 2020). Fourth, we link our sample to data from the Higher Education Statistics Agency (HESA) containing the universe of university enrolments to construct indicators for completing a 3-year university degree.

[7]In order to avoid misclassification from short summer courses or initial enrolments that are subsequently not actually taken up, we ignore learning spells shorter than two months in the treatment assignment.

[8]Hence, our effect estimates for earnings combine extensive and intensive margin and wage effects. Restricting the sample to individuals with non-missing earnings would introduce sample selection bias that renders effect estimates uninterpretable. Using the employment outcome we can discern extensive margin effects.

Table 1. Summary statistics.

	All students	Integrated schools	Non-integrated schools	Distances less than 33km	Mean-weighted sample
	(1)	(2)	(3)	(4)	(5)
A. Treatment choices					
No post-16 education (D_N)	0.13	0.11	0.15	0.15	0.15
Vocational education (D_V)	0.47	0.39	0.58	0.56	0.57
Academic education (D_A)	0.40	0.50	0.27	0.29	0.27
B. Demographic characteristics					
Female	0.49	0.50	0.48	0.48	0.48
White British	0.80	0.80	0.80	0.79	0.79
English as second language	0.09	0.08	0.11	0.12	0.11
Special educational need	0.18	0.15	0.22	0.21	0.22
Free school meal	0.14	0.11	0.19	0.20	0.20
C. Previous achievement					
KS2 score English	0.00 (1.00)	0.09 (0.99)	-0.13 (1.00)	-0.13 (1.00)	-0.14 (1.01)
KS2 score Maths	0.00 (1.00)	0.08 (1.00)	-0.12 (0.99)	-0.12 (0.99)	-0.12 (0.99)
KS2 score Science	0.00 (1.00)	0.08 (0.98)	-0.12 (1.01)	-0.12 (1.02)	-0.12 (1.02)
GCSE points	0.00 (1.00)	0.12 (0.98)	-0.18 (1.00)	-0.17 (1.00)	-0.19 (1.00)
D. Neighbourhood characteristics					
Income deprivation for children	0.21 (0.18)	0.19 (0.16)	0.25 (0.19)	0.26 (0.19)	0.26 (0.19)
Environment deprivation	-0.02 (1.00)	-0.11 (0.97)	0.12 (1.03)	0.16 (1.04)	0.16 (1.04)
Crime	0.03 (1.00)	-0.09 (0.98)	0.21 (1.00)	0.26 (1.00)	0.27 (0.98)
Barriers to housing	-0.06 (0.98)	-0.03 (0.98)	-0.10 (0.98)	-0.13 (0.97)	-0.13 (0.95)
Education deprivation	0.15 (1.07)	-0.02 (0.98)	0.39 (1.16)	0.41 (1.17)	0.44 (1.19)
Health deprivation	0.08 (1.01)	-0.10 (0.96)	0.36 (1.01)	0.39 (1.02)	0.42 (1.02)
Employment deprivation	0.09 (1.04)	-0.08 (0.94)	0.35 (1.14)	0.38 (1.15)	0.41 (1.16)
Income deprivation	0.11 (1.08)	-0.05 (0.98)	0.35 (1.17)	0.38 (1.19)	0.38 (1.18)
Distance to local centre in km	9.0 (13.7)	9.0 (13.9)	9.0 (13.5)	7.2 (9.7)	6.6 (8.8)
E. Distance instruments					
Dist. vocational college in km (Z_V)	6.3 (5.8)	7.1 (6.2)	5.1 (4.9)	4.7 (4.2)	4.4 (3.0)
Dist. academic college in km (Z_A)	18.0 (22.0)	21.1 (21.5)	13.3 (22.0)	7.8 (7.6)	7.6 (5.6)
F. Outcomes					
Annual earnings in £ (ages 29–30)	14,778 (13,343)	15,810 (13,790)	13,239 (12,488)	13,281 (12,549)	13,230 (12,508)
Employment (ages 29–30)	0.80 (0.38)	0.81 (0.37)	0.79 (0.39)	0.79 (0.39)	0.79 (0.39)
Observations	1,131,424	677,266	454,158	408,459	388,927

Notes: The table presents variable means and standard deviations in parentheses for different samples.

3.2 Sample Construction and Summary Statistics

For the estimation of our IV models we focus on students from non-integrated secondary schools who need to switch institution for both the vocational and the academic track, so that all students in our sample face the same symmetric educational choice problem, where both distance to vocational college and distance to academic college enter as cost shifters influencing choices. To gauge the extent to which this might limit the external validity of our effect estimates, in this and the following subsections, we will pay close attention to any differences in treatment selection and outcomes between students from non-integrated and integrated secondary schools.

Table 1 reports summary statistics for the whole student population in column 1 and splits the sample by integrated and non-integrated secondary schools in columns 2 and 3. These groups correspond to 60% and 40% of all students, respectively. Columns 4 and 5 refer to sub-samples of the non-integrated schools sample that are used for different estimations and are discussed in further detail below.

Panel A shows that post-compulsory education choices differ substantially between students from integrated and non-integrated schools: the latter are 19 percentage points (pp) more likely to enrol in vocational education, 4 pp more likely to pursue no post-16 education and, conversely, 23 pp less likely to enrol in academic education. This might be due to better academic preparation on integrated schools, peer effects, lower costs of enrolling in the academic track and differences in self-selection into these schools, apparent from the remainder of the table. Students from non-integrated secondary schools are slightly more likely to be male and to speak English as a second language and substantially more likely to have a special educational need or to be economically disadvantaged (panel B). Their scores in the end of primary school KS2 tests are about 0.2 standard deviations lower in all three subjects (panel C). They live in neighbourhoods that are more deprived along all eight dimensions measured by our indices (panel D). Similarly, their annual earnings at the ages of 29–30 are lower by about £2,500 (panel F). Note, however, that they do not represent a more rural student population as evidenced by parity in the distance to the closest local economic centre measure (panel D). As seen in Panel E, they do live closer to both Sixth Form Colleges and Further Education Colleges, indicating some sorting of either these institutions towards their constituencies or *vice-versa*.

In sum, students from non-integrated secondary school represent a moderately negatively selected group in terms of achievement and socio-economic background for whom vocational education plays a particularly important role. Accordingly, they are the group most affected by changes in the size of the vocational sector and most likely to have substantial shares of students at the margin between vocational and no post-16 education and at the margin between vocational and academic education. These facts make our estimation sample relevant from a policy perspective, even though it does not cover the whole population. Further, we show below that education premiums for employment and earnings in the non-integrated school sample are very similar to those in the full sample, suggesting that our effect estimates might well extend to the whole population.

To reduce functional form dependence we estimate our IV models using flexible locally linear specifications, stratifying along the two-dimensional distance grid defined by our instruments. This means that all our estimates are "local", as in belonging to certain values of the instruments, and the main results will be evaluated at their means in the non-integrated school sample. Column 5 summarises this mean-weighted sample, where observations are weighted by their proximity to the instrument means using a two-dimensional Epanechnikov kernel with a 15km bandwidth. We assess the representativeness of the mean results by estimating our model over a wide range of points across the distance grid. As the data becomes sparse and the first stages break down

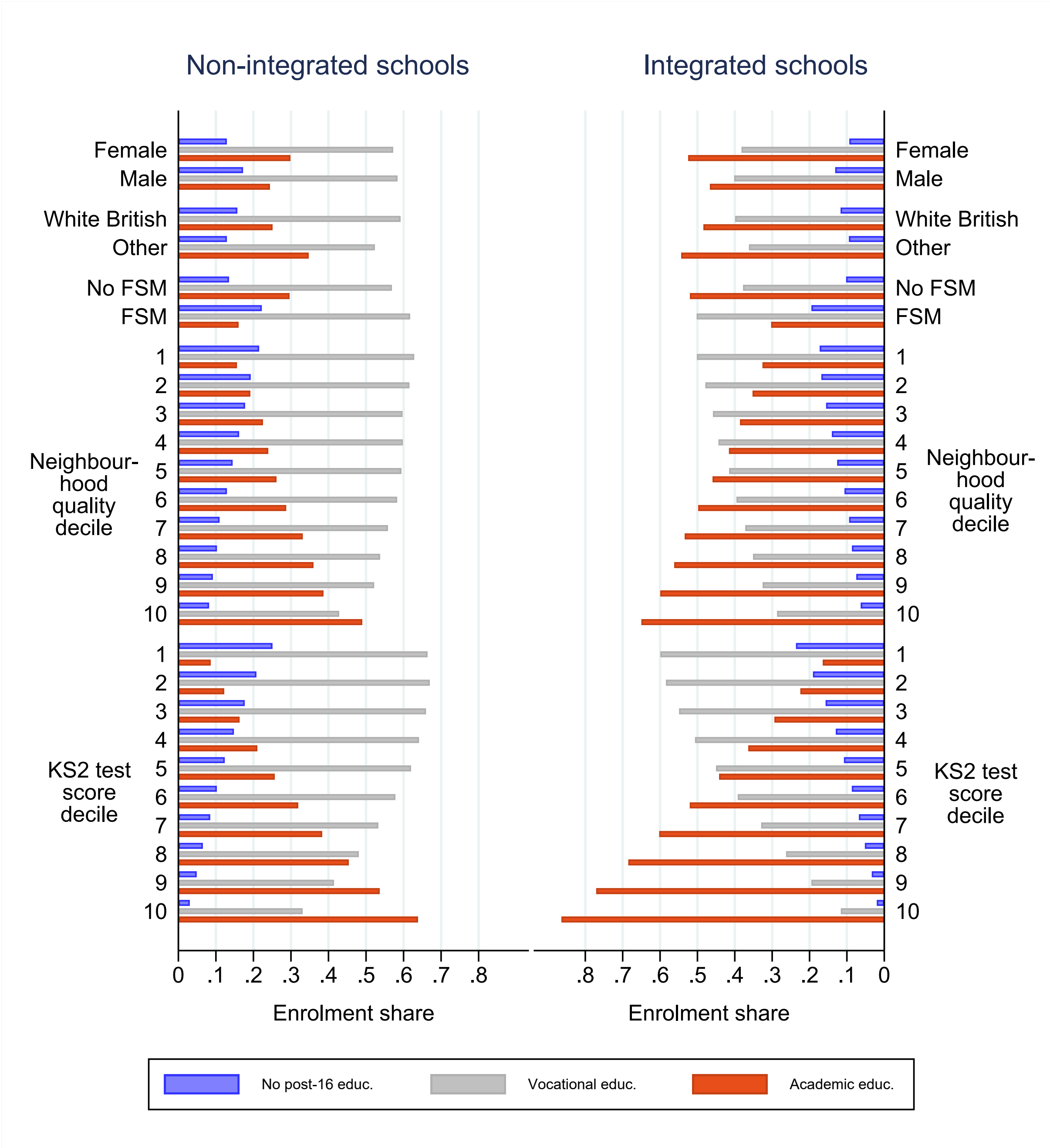


Figure 2. Education choices by observable characteristics.

Notes: This figure shows the share of students in each treatment by observable characteristics, separately for students from non-integrated and integrated schools. Neighbourhood quality deciles are deciles of the (inverse) first principal component of all eight deprivation indices. KS2 test score deciles are deciles of the first principal component of all three separate KS2 scores. The principal components are extracted (and their deciles calculated) in the whole sample, so that the deciles refer to the same categories for integrated and non-integrated schools.

at large distances, we do not use the most remote observations with distances greater than 33km to either college in the estimation of our IV models (about 10% of non-integrated school students). Column 4 summarises all observations that are used in the estimation of our IV models. Comparison of columns 3–5 suggests that focusing on certain distances does not affect the sample's representativeness much. In particular, the mean-weighted sample seems almost perfectly representative of the full sample of non-integrated school students.

3.3 Selection into Treatment

Figure 2 describes how initial post-compulsory education choices depend on observable student characteristics, separately for students from non-integrated and integrated schools. Across all categories students from integrated school are more likely to enrol in the academic track and less likely to enrol in the vocational track or to choose no post-16 education. However, the selection patterns appear to be roughly identical in both groups.

In both groups, the no post-16 education and academic education shares vary stronger with demographic and neighbourhood characteristics than the vocational share does. Vocational enrolment is very similar for men and women, but women are more likely to enrol in the academic track and less likely to pursue no upper secondary education. Non-white British students are less likely to be in the vocational track or in no education but much more likely to enrol in the academic track. Disadvantaged (i.e. FSM) students are far less likely to enrol in the academic track, only somewhat more likely to enrol in the vocational track and much more likely to not enrol in either. Also with respect to neighbourhood quality the enrolment gradient is steepest for academic education, with students from the highest decile more than (close to) twice as likely to enrol in the academic track than those from the lowest decile for non-integrated school students (for integrated school students). The no post-16 education and vocational education shares decrease roughly equally with neighbourhood quality.

Education choices appear to be close to monotonic in test scores. In both groups, students from the highest test score decile are roughly six times as likely to enrol in the academic track than those from the lowest decile, which translates to an increase from below 10% to about 65% for non-integrated school students and an increase from 17% to about 88% for integrated school students. The share of students who pursue no post-16 education decreases from about 25% in the bottom decile to about 2% in both groups. Also the vocational share is decreasing in test scores, though here the gradient is slightly steeper for students from integrated schools and on non-integrated schools it exhibits a slight hump-shape, increasing from the first to the second decile before monotonically falling thereafter.

3.4 Differences in Education Outcomes

Figure 3 compares the three education groups in terms of three key education outcomes, using our estimation sample of non-integrated school students. The differences between groups are stark: about 90% of academic-track students successfully complete a upper secondary Level 3 qualification, compared to only about 60% of vocational-track students, despite this being the

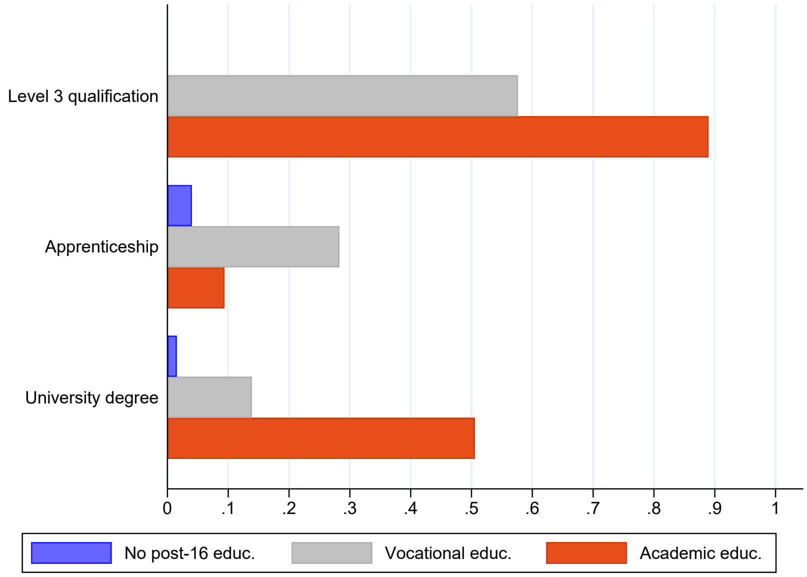


Figure 3. Intermediate education outcomes by initial education choice.

Notes: This figure show the share of students (in the non-integrated schools sample) achieving the following education outcomes by treatment status: achieving a Level 3 qualification within 24 months from the relevant enrolment; starting an apprenticeship at any time after sitting GCSEs; completing a 3-year university degree at any time after sitting GCSEs.

stated goal of both tracks. Half of the academic-track students complete a 3-year university degree, compared to only 13% of vocational-track students. Though this difference is large, the numbers indicate that the vocational track in England does not rule out higher education later on per se. Just below 30% of vocational-track students start an apprenticeship, compared to less than 10% of academic-track students. Unsurprisingly, students that initially choose no post-16 education obtain almost no further educational credentials, apart from about 4% who start an apprenticeship.

3.5 Differences in Labour Market Performance

Figure 4 plots raw employment and earnings dynamics by post-16 education choice over the full observed age range, separately for women and men. Note that unlike in the main regressions (for outcomes at ages 29–30) these figures exclude employment and earnings from self-employment to ensure comparability across all years.

Panel A shows a clear ranking for the probability of being employed, constant across gender and age: academic track students are more likely to be employed than vocational track students who in turn are more likely to be employed than students without post-16 education.[9] For both genders, the raw differences stabilise rather quickly, remaining roughly constant from the age

[9]Note that employment is operationalised as having (any) positive earnings. Accordingly, between-group differences might be somewhat muffled at younger ages, as the measure picks up any income as employment, even it comes from a short-term summer or student job.

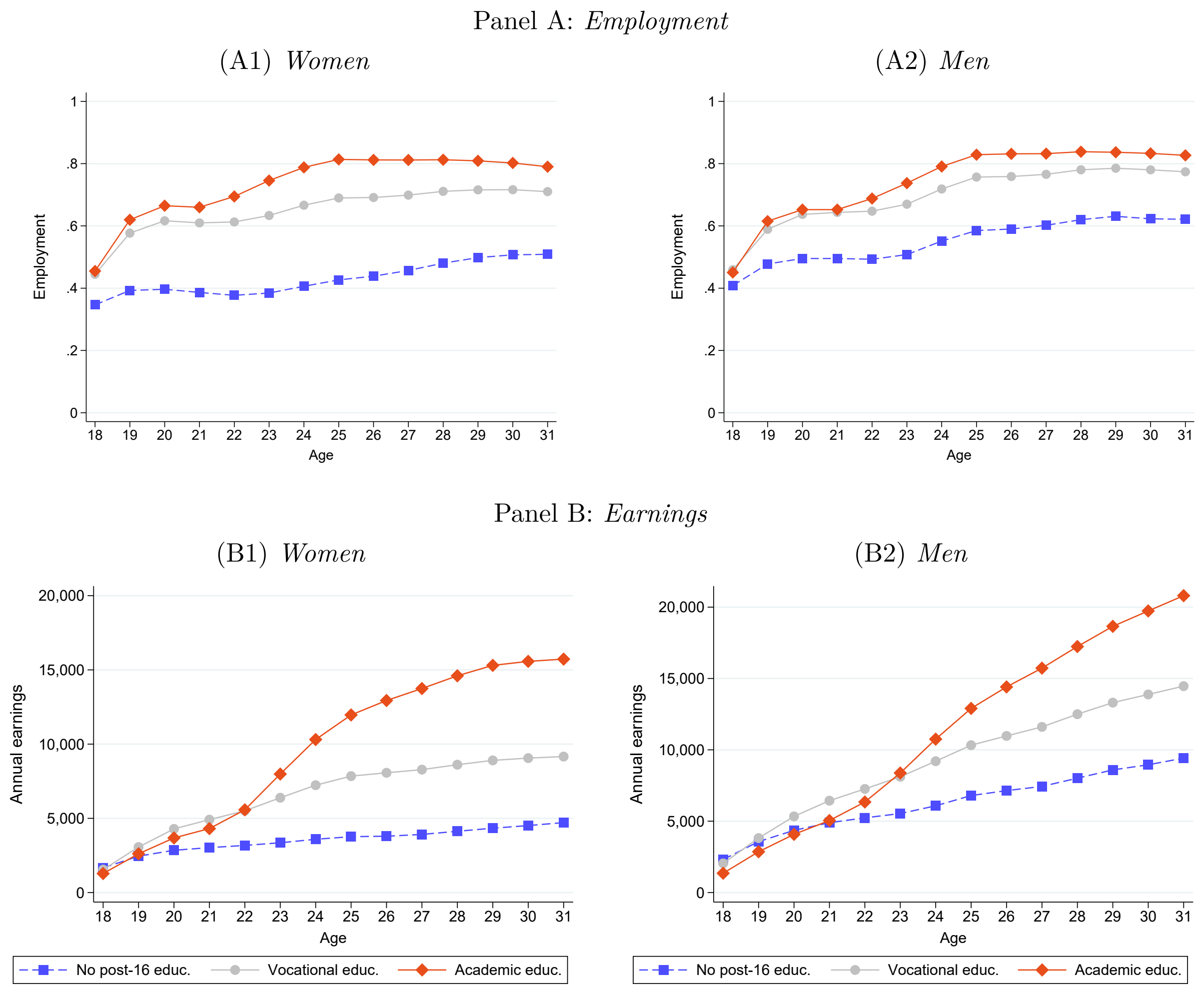


Figure 4. Labour market trajectories by initial enrolment.

Notes: The figure is based on students from non-integrated secondary schools. Annual earnings are measured in real 2010 British pounds. For comparability across the whole age range only earnings (and employment) from employed, but not from self-employed, work are included. Outcomes at age 31 are only available for the 2002 cohort.

of 23 onwards, and are much larger between vocational and no-post 16 education than between academic and vocational education. Two gender differences emerge: first, within each education group men are more likely to be employed than women. This gender difference is largest for the no post-16 education group, smaller but substantial for the vocational education group and only barely visible for the academic education group.[10] Second, the raw education premiums for employment are slightly larger for women.

The earnings trajectories in panel B show the expected pattern: those without post-16 education have the highest earnings at the age of 18 but are immediately overtaken by those with vocational education, who in turn are overtaken by those with academic education a couple of

[10]Note that unobserved differences at the intensive margin are most probably larger than the observed extensive margin differences.

years later. Academically educated women overtake their vocationally educated peers at the age of 22 – a year prior to men. Whereas differences between the no post-16 and vocational education groups stabilise rather quickly, earnings differences between academically educated students and the rest continue to grow throughout students' mid-twenties – but at a decreasing rate. For women, differences seem to have stabilised by the age of 29. For men, they might continue to grow slightly beyond the age of 31 but, if so, at a slow pace. Strikingly, the differences between academically and vocationally educated students are much more pronounced for earnings than they are for employment.

Gender differences mirror those from before: first, men earn more than women within each education group. Again, this difference is strongest for the no post-16 and vocational education groups but now it is substantial also for the academically educated. Second, the raw education premiums for earnings are larger for women than for men, though this difference shrinks over the age range and has almost vanished by the age of 30 (in absolute terms). Therefore, we will perform much of our analyses pooling over men and women for the sake of brevity. Of course, the main models will also be presented separately by gender. When we investigate how treatment effects vary across the age range, analyses are performed separately by gender due to the observed differences in earnings dynamics.

3.6 OLS Results

The raw labour market outcome differences across education groups represent a mixture of causal effects and selection. As a first step to approximate the causal returns to upper secondary education, we use and OLS regressions and our rich data on students' demographics, previous performance, school- and neighbourhood characteristics to estimate the *controlled* education premiums for employment and earnings. In particular, we estimate models of the following form:

$$Y = \alpha + \beta_{N \leftarrow V}\, D_N + \beta_{A \leftarrow V}\, D_A + \gamma\, \mathbf{X} + \varepsilon, \tag{1}$$

where the dependent variable Y is either average employment or annual earnings at ages 29–30 and D_N and D_A are indicator variables for no post-16 education and academic education, respectively, making vocational education the reference category. The control set $\mathbf{X}$ consists of cohort fixed effects; region fixed effects; all student demographics listed in Table 1, including all their two-way interactions; all school-level controls and cubic polynomials in all three KS2 test scores, all eight neighbourhood quality indices and distance to local economic centre.

Table 2 presents estimation results for both outcomes, with and without controls and for three different samples: all students, only non-integrated school students and the mean-weighted IV sample. Note that, in contrast to Figure 4, both outcomes now include earnings from self-employment, though columns 1 and 3 reveal that this has little effect on the raw gaps. Comparing panels A and B shows that for both employment and earnings the raw vocational vs. no post-16

Table 2. Raw and controlled OLS regressions for labour market outcomes at ages 29–30.

Dependent variable:	Employment		Annual earnings	
	Raw	Controlled	Raw	Controlled
	(1)	(2)	(3)	(4)
A. All students ($N = 1,131,424$)				
No post-16 education (D_N)	-0.161***	-0.152***	-3860***	-3125***
	(0.001)	(0.001)	(34)	(31)
Academic education (D_A)	0.053***	0.031***	6431***	3246***
	(0.001)	(0.001)	(32)	(29)
B. Non-integrated schools ($N = 454,158$)				
No post-16 education (D_N)	-0.171***	-0.156***	-4332***	-3170***
	(0.002)	(0.002)	(48)	(43)
Academic education (D_A)	0.050***	0.028***	5632***	2838***
	(0.001)	(0.001)	(54)	(49)
C. Mean-weighted sample ($N = 388,927$)				
No post-16 education (D_N)	-0.171***	-0.157***	-4327***	-3187***
	(0.002)	(0.002)	(55)	(49)
Academic education (D_A)	0.054***	0.030***	5895***	2933***
	(0.001)	(0.002)	(59)	(55)

Notes: The table presents estimation results from OLS regressions of average employment and annual earnings at ages 29–30 on two treatment indicators for no post-16 education and academic education (making vocational education is the reference category) for different samples. In the raw regressions the control set includes cohort fixed effects only. In the controlled regressions the control set includes cohort fixed effect, region fixed effects and the full set of (interacted) student-, school- and neighbourhood-level controls. Panel A uses all students, panel B only those from non-integrated school and panel C uses the mean-weighted sample, which weights observations by their proximity to the instrument means using a two-dimensional Epanechikov Kernel and a bandwidth of 15km. Standard errors, reported in parentheses, are clustered at the LSOA level. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

education premium is slightly larger in the non-integrated schools sample, whereas the academic vs. vocational education premium is slightly larger in the full sample, especially for earnings. However, once we condition on the full set of covariates these differences shrink considerably, suggesting that, after conditioning, education premiums estimated from on the non-integrated schools sample are informative of the full population. Like in Table 1, weighting observations by their proximity to the instrument means yields results that are practically identical to those from unweighted OLS.

Even though the inclusion of covariates has a considerable impact on the estimated gaps – halving it in the case of academic vs. vocational education – the differences between education groups remain large even after conditioning. Taken at face value, the OLS results suggest that the earnings premium from academic vs. vocational education is about as large as the that from vocational vs. no-post 16 education, with both being equal to about £3,000 in annual earnings at ages 29–30. While the advantage of vocationally educated students vs. their peers without post-16 education seems, for a large part, be driven by extensive margin effects – the effect on the probability of employment at this margin is about 15 percentage points (pp) – their disadvantage vs. academically educated students seems to be driven mainly by intensive margin and wage effects, given that the effect on employment at that margin equals only about 3 pp. Though indicative, there are several problems with the interpretation of the OLS estimates, as we will detail in the next section.

4 Research Design

A key complication in our setting is that students select into one of three, instead of two, unordered treatments. This limits the interpretability of common estimators, such as OLS and two-stage least squares (2SLS), and thus warrants an alternative identification approach to obtain meaningful estimates of the returns to vocational education. In this section, we briefly explain the limitations of OLS and 2SLS and subsequently introduce our alternative empirical strategy.

4.1 Limitations of OLS

There are three problems with interpreting the OLS estimates as policy-relevant causal returns to education. To see this, define the three discrete and mutually exclusive treatment conditions as $D = N$ (no post-16 education), $D = V$ (vocational education) and $D = A$ (academic education), with corresponding binary treatment indicators D_N, D_V and D_A, and denote the associated potential outcomes as Y_N, Y_V and Y_A. Observed outcomes are given by $Y = \sum_{d\in\{N,V,A\}} D_d Y_d$. Now, consider the OLS estimate of the effect of academic vs. vocational education, though the same arguments apply to the other margin. The OLS estimand from (1) (without controls) can we written in terms of potential outcomes as follows:

$$\beta_{A\leftarrow V} = \mathbb{E}[Y|D=A] - \mathbb{E}[Y|D=V] = \underbrace{\mathbb{E}[Y_A - Y_V|D=A]}_{\text{Effect of } A \text{ vs. } V} + \underbrace{\mathbb{E}[Y_V|D=A] - \mathbb{E}[Y_V|D=V]}_{\text{Selection bias}} \quad (2)$$

The first, standard, problem relates to the presence of the selection bias term, which might remain even after conditioning on $\mathbf{X}$. The second problem arises because students who choose the academic track may differ in their next-best choice. Define the track a student chooses if alternative d is removed from her choice set as $D_{/d}$. Using this notation and expanding the first term on the right-hand side of (2), we get:

$$\begin{aligned}\mathbb{E}[Y_A - Y_V|D=A] = {} & \mathbb{E}[Y_A - Y_V|D=A,\, D_{/A}=V]\ \Pr(D_{/A}=V|D=A) \\ & + \mathbb{E}[Y_A - Y_V|D=A,\, D_{/A}=N]\ \Pr(D_{/A}=N|D=A),\end{aligned}$$

which is a weighted average of the effect of academic vs. vocational education for students with different next-best alternatives, some of whom would not actually choose the vocational track could they not choose the academic track. This illustrates that, even in the absence of selection bias, the OLS estimate is difficult to interpret as a return. The third problem is that

OLS estimates treatment-on-the-treated effects that might well be unrepresentative of effects for students whose education choices are responsive to policy changes, like an expansion or contraction of the vocational sector. From the perspective of such a policy, the effect for a student who under no circumstances would consider switching track is of little relevance. And even without such effect heterogeneity, OLS gives no indication of the relative responsiveness of the two margins, i.e. of how many students are likely to switch from no post-16 education to vocational education (or vice-versa) and from academic education to vocational education (or vice-versa), making it hard to gauge a given policy's overall impact.

We do not wish to argue that the OLS estimates are meaningless. Regarding the first problem, our comparatively elaborate control set might make any remaining bias small. The second problem might be small at the academic vs. vocational margin due to the small size of the no post-16 education group and at the other margin one might argue that the next-best alternative for most students in the no post-16 group is vocational education. Yet, all of these are strong assumptions and hard to assess *ex ante*. Furthermore, we expect the third problem to be substantial given that the literature generally finds large heterogeneity in returns to education (see e.g. Carneiro *et al.*, 2011). Therefore, we deem a solely OLS-based analysis unsatisfactory. In the next subsection, we consider 2SLS as a potential remedy.

4.2 Insufficiency of 2SLS

Instrumental variables (IV) offer a potential solution to the problem of selection bias. Depending on the instrument, the local average treatment effect (LATE) identified by IV can be directly policy-relevant, as it applies to a sub-population at the margin of treatment (Imbens & Angrist, 1994). However, multiple margins of treatment present a challenge to the interpretability of LATE. Kline and Walters (2016) show that in a setting with three discrete treatments and a instrumental variable for one of the alternatives available, 2SLS identifies a pooled LATE that fuses the two margin-specific effects into one weighted average. To exemplify this point, if we would use binary variable Z to instrument the vocational education indicator, D_V, in the simple outcome equation $Y = \alpha + \beta_V D_V + \varepsilon$, under the usual assumptions, the 2SLS estimand equals:

$$\beta_V^{IV} = \underbrace{LATE_V}_{\text{Net effect of } V} = \lambda \underbrace{LATE_{V \leftarrow N}}_{\text{Effect of } V \text{ vs. } N} + (1-\lambda) \underbrace{LATE_{V \leftarrow A}}_{\text{Effect of } V \text{ vs. } A} \tag{3}$$

The weight λ equals the share of Z-compliers who are at the vocational-no post-16 education margin and is identified by the reduction in $Pr(D = N)$ induced by Z as a share of the increase in $Pr(D = V)$. The two margin-specific complier treatment effects, however, are not identified. Heckman and Urzúa (2010) show the equivalent result for the marginal treatment effect (MTE), the limit version of LATE, which can be estimated with a continuous instrument by the method of local instrumental variables (LIV).[11] Hence, with a single instrument we can only retrieve the net effect of vocational education.

Using distance to vocational college as an instrument, we could thus estimate the net effect of expanding access to vocational colleges in England, which, arguably, is an important parameter. However, identification of the margin-specific effects would allow for a much more comprehensive evaluation of such a policy. Equation (3) reveals that the same net effect can be composed of many different combinations of margin-specific effects, with potentially very different policy implications. For example, a moderately positive net effect composed of a substantial effect at the no post-16 education margin and a zero effect at the academic margin might warrant large-scale expansion of vocational colleges, whereas the same net effect composed of large effects at the no post-16 education margin but substantial negative effects at the academic margin would suggest more targeted policies instead.

Given that we observe distance to vocational and to academic college, one might think that multivariate 2SLS applied to (1), instrumenting the two treatments D_N and D_A with the two distance instruments Z_V and Z_A, secures identification of the margin-specific effects. However, as shown by Kirkeboen *et al.* (2016), even with as many instruments as treatments, 2SLS does not identify well-defined treatment effects along policy-relevant margins. Instead, the 2SLS estimands amalgamate all three effect margins, resulting in fundamentally uninterpretable quantities. These shortcomings of conventional multivariate IV motivate our use of an alternative IV-based identification approach.

4.3 Identification Framework

To separately identify the two margin-specific treatment effects of vocational education we follow an identification procedure proposed by Mountjoy (2019). The core of this method lies in "cross-instrumenting" educational choices using exogenous variation in the attractiveness of alternative choices, holding fixed their own attractiveness. Our two distance instrument allow us to do exactly this: for example, conditional on distance to academic college, variation in distance to vocational college only changes the attractiveness of vocational education, but not that of no post-16 or academic education. Under a standard monotonicity assumption, this restricts the possible complier flows and thereby allows for margin-specific identification of some potential outcomes. Under an additional complier comparability assumption, which arises naturally in a setting where both instrumental variables represent distance to education providers, all the relevant potential outcomes – and, hence, also treatment effects – are identified.

[11]The MTE is the continuous-instrument analogue to LATE in that it is defined without any parametric assumptions or restrictions on effect heterogeneity (Kennedy *et al.*, 2019). In the binary instrument case (without covariates), 2SLS equals the Wald ratio and thus non-parametrically identifies LATE. This is no longer true in the continuous instrument case, where 2SLS imposes parametric assumptions on the first stage relationship. LIV, in contrast, non-parametrically identifies MTE.

Mountjoy's (2019) procedure relies on three identification assumptions. To formalise these, denote potential treatment choice as $D(z_V, z_A) \in \{N, V, A\}$. This represents the education choice a student would make if exogenously assigned to instrument values $(Z_V, Z_A) = (z_V, z_A)$. Corresponding binary indicators are defined analogously. To simplify notation, we suppress the individual index i and implicitly condition on our control set $\mathbf{X}$ in everything that follows.

The first assumption is the canonical IV assumption of independence and exclusion, adapted to the multiple treatments and two instruments setting:

Assumption A1. Independence and Exclusion:

$$(Z_V, Z_A) \perp\!\!\!\perp \left(Y_N, Y_V, Y_A, \{D(z_V, z_A)\}_{\forall(z_V, z_A)}\right)$$

This assumption requires the two distance instruments to be as good as randomly assigned with respect to students' potential outcome and treatment choices, conditional on the implicit control set $\mathbf{X}$. The primary challenge to A1 is residential sorting, which is why we include detailed geographical controls in $\mathbf{X}$ next to student- and school-level variables. We assess the plausibility of A1 by means of balance tests on observed pre-determined covariates.

The second assumption adapts the usual monotonicity condition:

Assumption A2. Partial Unordered Monotonicity:

For all triples $(z_V,\, z'_V,\, z_A)$ with $z'_V < z_V$ we have: $D_V\left(z'_V, z_A\right) \geq D_V\left(z_V, z_A\right)$ but $D_A\left(z'_V, z_A\right) \leq D_A\left(z_V, z_A\right)$ and $D_N\left(z'_V, z_A\right) \leq D_N\left(z_V, z_A\right)$ for all individuals.

For all triples $(z_A,\, z'_A,\, z_V)$ with $z'_A < z_A$ we have: $D_A\left(z_V, z'_A\right) \geq D_A\left(z_V, z_A\right)$ but $D_V\left(z_V, z'_A\right) \leq D_V\left(z_V, z_A\right)$ and $D_N\left(z_V, z'_A\right) \leq D_N\left(z_V, z_A\right)$ for all individuals.

A2 extends the intuition of "no defiers" from the binary to the multi-valued treatment case, considering only conditional instrumental variation.[12] It requires that a decrease (increase) in the distance to either type of college, holding constant distance to the other, renders the associated education choice weakly more (less) attractive for all students. It does not restrict the complier flows to a certain margin, however. For example, as the distance to vocational college decreases ($z'_V < z_V$), but distance to academic college is held fixed, some people may switch into but no one out of vocational education ($D_V\left(z'_V, z_A\right) \geq D_V\left(z_V, z_A\right)$); whether these compliers come from academic education ($D_A\left(z'_V, z_A\right) \leq D_A\left(z_V, z_A\right)$) or no post-16 education ($D_N\left(z'_V, z_A\right) \leq D_N\left(z_V, z_A\right)$) if left unrestricted. However, nobody may switch between no post-16 and academic education upon such a decrease in Z_V.

Given the exogeneity of the instruments, partial unordered monotonicity is natural assumption in the case of our distance instruments. It would be violated if there are complementarities

[12] Heckman and Pinto (2018) develop the general 'unordered monotonicity' condition for the unordered multi-valued treatment case. It requires that treatment responses are uniform across *all* possible shifts in the instruments. Mountjoy's (2019) 'partial unordered monotonicity' relaxes this assumption by looking only at *conditional* variation in the instruments, i.e. focusing on the subset of shifts where one of the two instruments stays constant. This means that we make no assumptions about the behaviour of students in cases where distance to both colleges decreases simultaneously.

between the two college types, so that the one's attractiveness is tied to that of the other. This is unlikely as academic and vocational colleges are substitutes and enrolling in one is not a preparatory step for enrolling in the other. We assess the plausibility of A2 by graphically inspecting conditional associations between our treatment variables and distance instruments.

The third assumption is specific to Mountjoy's (2019) framework and draws a connection between the two sets of vocational-academic compliers induced by Z_V and Z_A:

Assumption A3. Complier Comparability:

For all pairs (z_V, z_A):

$\lim_{z'_V \uparrow z_V} \mathbb{E}\left[Y_V \mid D\left(z'_V, z_A\right) = V, D\left(z_V, z_A\right) = A\right] = \lim_{z'_A \downarrow z_A} \mathbb{E}\left[Y_V \mid D\left(z_V, z'_A\right) = V, D\left(z_V, z_A\right) = A\right]$

This assumption states that compliers shifted from academic to vocational college by a marginal *decrease* in distance to vocational college (left-hand side) must be comparable, in terms of potential outcomes, with compliers shifted from vocational to academic college by a marginal *increase* in distance to academic college (right-hand side). Given that both sets must be students at a margin of indifference between vocational and academic education and both instruments represent simply the distance to the closest respective provider, it is hard to imagine how these two complier types could systematically differ.[13] Even though A3 cannot be tested directly, we will assess its plausibility by comparing the two complier groups in terms of their pre-determined characteristics, which are separately identified in the framework developed next.

Under assumptions A1–A3, mean potential outcomes for compliers along the vocational-no post-16 education margin and the vocational-academic education margin can be identified from the data as ratios of partial derivatives.[14] This section states and briefly explains the intuition behind the identification results, closely following Mountjoy (2019). For a formal proof the interested reader is referred to Appendix D of his paper.

Start by decomposing the reduced form's partial derivative with respect to Z_V, using the fact that $D_N + D_V + D_A = 1$:

$$\begin{aligned}\frac{\partial \mathbb{E}\left[Y \mid Z_V, Z_A\right]}{\partial Z_V} &= \frac{\partial \mathbb{E}\left[Y D_N + Y D_V + Y D_A \mid Z_V, Z_A\right]}{\partial Z_V} \\ &= \frac{\partial \mathbb{E}\left[Y D_N \mid Z_V, Z_A\right]}{\partial Z_V} + \frac{\partial \mathbb{E}\left[Y D_V \mid Z_V, Z_A\right]}{\partial Z_V} + \frac{\partial \mathbb{E}\left[Y D_A \mid Z_V, Z_A\right]}{\partial Z_V}.\end{aligned}$$

We consider each of these partial derivatives separately. To begin, note that $Y D_N = Y_N$

[13] Mountjoy (2019) shows that this condition is implied by a standard Roy-style selection model: both Z_V and Z_A act as costs shifting a single index that governs the relative attractiveness of vocational vs. academic education. Hence, students who switch their treatment choice in response to a marginal change in the index are the same regardless of whether this change is is induced by a marginal decrease in Z_V or a marginal increase in Z_A (or *vice-versa*).

[14] Theory for identifying complier potential outcome distributions using instruments was developed by Imbens and Rubin (1997), Abadie (2002) and Carneiro and Lee (2009). Mountjoy (2019) extends this logic to multiple treatments using conditional variation in alternative-specific instruments.

if $D_N = 1$ and $YD_N = 0$ otherwise, so that instrument-induced changes in $\mathbb{E}[YD_N \mid Z_V, Z_A]$ contain information about Y_N for compliers who are switching in or out of D_N in response to changes in the instrument. Under partial unordered monotonicity (A2), Z_V-induced variation in D_N must come from compliers at the $V \leftarrow N$ margin, because there are no movements between N and A when only the attractiveness of V changes (remember that Z_A is held fixed by conditioning). Therefore, under A1 and A2, we have that:

$$
\begin{aligned}
\frac{\partial \mathbb{E}[YD_N \mid Z_V, Z_A]}{\partial Z_V} &= \lim_{z'_V \uparrow z_V} \frac{\mathbb{E}[YD_N \mid z_V, z_A] - \mathbb{E}[YD_N \mid z'_V, z_A]}{z_V - z'_V} \\
= \lim_{z'_V \uparrow z_V} \mathbb{E}\left[Y_N \mid D(z'_V, z_A) = V, D(z_V, z_A) = N\right] & \frac{\Pr(D(z_V, z_A) = N) - \Pr(D(z'_\nu, z_A) = V)}{z_V - z'_V} \\
= \mathbb{E}[Y_N \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)] \frac{\partial \mathbb{E}[D_N \mid Z_V, Z_A]}{\partial Z_V},
\end{aligned}
$$

where we use $\mathbb{E}[Y_N \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)]$ as shorthand notation to refer to $\lim_{z'_V \uparrow z_V} \mathbb{E}[Y_N \mid D(z'_V, z_A) = V, D(z_V, z_A) = N]$ evaluated at a given instrument point $(Z_V, Z_A) = (z_V, z_A)$. As the partial derivatives $\frac{\partial \mathbb{E}[YD_N|Z_V,Z_A]}{\partial Z_V}$ and $\frac{\partial \mathbb{E}[D_N|Z_V,Z_A]}{\partial Z_V}$ are directly identified from the data, their ratio identifies the first potential outcome of interest $\mathbb{E}[Y_N \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)]$ over all points in the instrument support where the first stage $\frac{\partial \mathbb{E}[D_N|Z_V,Z_A]}{\partial Z_V}$ is non-zero.

Analogously, changes in D_A with respect to Z_V are driven by $V \leftarrow A$ compliers so that for the derivative of $\mathbb{E}[YD_A \mid Z_V, Z_A]$ with respect to Z_V we get that:

$$
\frac{\partial \mathbb{E}[YD_A \mid Z_V, Z_A]}{\partial Z_V} = \mathbb{E}[Y_A \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] \frac{\partial \mathbb{E}[D_A \mid Z_V, Z_A]}{\partial Z_V},
$$

which identifies the second potential outcome of interest.

Thanks to A2, the first two derivatives only captured changes along one margin of treatment. The partial derivative of $\mathbb{E}[YD_V \mid Z_V, Z_A]$ with respect to Z_V, however, reflects changes in D_V coming from both margins:

$$
\begin{aligned}
\frac{\partial \mathbb{E}[YD_V \mid Z_V, Z_A]}{\partial ZV} = \mathbb{E}[Y_V \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)] \frac{\partial \mathbb{E}[D_N \mid Z_V, Z_A]}{\partial Z_V} \\
+ \mathbb{E}[Y_V \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] \frac{\partial \mathbb{E}[D_A \mid Z_V, Z_A]}{\partial Z_V}.
\end{aligned}
\tag{4}
$$

To disentangle these margins we require the second instrument: changes in D_V with respect to distance to academic college, Z_A, must be driven by $V \leftarrow A$ compliers, so that:

$$
\frac{\partial \mathbb{E}[YD_V \mid Z_V, Z_A]}{\partial Z_A} = \mathbb{E}[Y_V \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] \frac{\partial \mathbb{E}[D_V \mid Z_V, Z_A]}{\partial Z_A}.
$$

This expression identifies the third potential outcome of interest,

$\mathbb{E}[Y_V \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)]$, which subsequently can be plugged into (4) to identify the final potential outcome of interest $\mathbb{E}[Y_V \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)]$, given that all other quantities in the equation are directly identified from the data. These two steps require the complier comparability assumption (A3) to ensure that potential outcomes of compliers induced by Z_A are identical to those induced by Z_V.

We have thus secured all four mean potential outcomes necessary for forming the margin-specific treatment effects of interest:

$$\begin{aligned}
\mathbb{E}[Y_N \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)] &= \frac{\frac{\partial \mathbb{E}[YD_N \mid Z_V, Z_A]}{\partial Z_V}}{\frac{\partial \mathbb{E}[D_N \mid Z_V, Z_A]}{\partial Z_V}} \\
\mathbb{E}[Y_A \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] &= \frac{\frac{\partial \mathbb{E}[YD_A \mid Z_V, Z_A]}{\partial Z_V}}{\frac{\partial \mathbb{E}[D_A \mid Z_V, Z_A]}{\partial Z_V}} \\
\mathbb{E}[Y_V \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] &= \frac{\frac{\partial \mathbb{E}[YD_V \mid Z_V, Z_A]}{\partial Z_A}}{\frac{\partial \mathbb{E}[D_V \mid Z_V, Z_A]}{\partial Z_A}} \\
\mathbb{E}[Y_V \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)] &= \frac{\frac{\partial \mathbb{E}[YD_V \mid Z_V, Z_A]}{\partial Z_V}}{-\frac{\partial \mathbb{E}[D_N \mid Z_V, Z_A]}{\partial Z_V}} - \frac{\frac{\partial \mathbb{E}[YD_V \mid Z_V, Z_A]}{\partial Z_A}}{\frac{\partial \mathbb{E}[D_V \mid Z_V, Z_A]}{\partial Z_A}} \frac{\frac{\partial \mathbb{E}[D_A \mid Z_V, Z_A]}{\partial Z_V}}{\frac{\partial \mathbb{E}[D_N \mid Z_V, Z_A]}{\partial Z_V}}
\end{aligned} \tag{5}$$

Then, treatment effects are given by simple subtraction:

$$\begin{aligned}
&\mathbb{E}[Y_V \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)] - \mathbb{E}[Y_N \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)] \\
&= \mathbb{E}[Y_V - Y_N \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)] \\
&= MTE_{V \leftarrow N}(Z_V, Z_A)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[Y_V \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] - \mathbb{E}[Y_A \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] \\
&= \mathbb{E}[Y_V - Y_A \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)] \\
&= MTE_{V \leftarrow A}(Z_V, Z_A)
\end{aligned}$$

These marginal treatment effects (MTEs) are simply the continuous instrument analogues to discrete LATEs. They represent the treatment effect for marginal students at specific values of the instruments (Z_V, Z_A). Accordingly, they can be computed across the empirical support of the instruments, provided that the first stage partial derivatives are non-zero.

Finally, recall from the previous subsection that the net marginal treatment effect of vocational education is identified by the local instrumental variables estimand for the binary treatment indicator, D_V, instrumented with Z_V,

$$MTE_V(Z_V, Z_A) = \frac{\frac{\partial \mathbb{E}[Y \mid Z_V, Z_A]}{\partial Z_V}}{\frac{\partial \mathbb{E}[D_V \mid Z_V, Z_A]}{\partial Z_V}},$$

and that the share of $V \leftarrow N$ compliers is given by the Z_V-induced reduction in D_N as a share of the increase in D_V:

$$\lambda\left(Z_V, Z_A\right) = \frac{-\frac{\partial \mathbb{E}[D_N|Z_V,Z_A]}{\partial Z_V}}{\frac{\partial \mathbb{E}[D_V|Z_V,Z_A]}{\partial Z_V}}.$$

Combining these parts we arrive at the decomposition of interest:

$$\underbrace{MTE_V}_{\text{Net effect of } V} = \lambda \underbrace{MTE_{V\leftarrow N}}_{\text{Effect of } V \text{ vs. } N} + (1-\lambda) \underbrace{MTE_{V\leftarrow A}}_{\text{Effect of } V \text{ vs. } A}. \tag{6}$$

4.4 Estimation

All parameters of interest in (6) are identified as ratios of partial derivatives of the conditional expectations of $\{D_N, D_V, D_A, Y, YD_N, YD_V, YD_A\}$ with respect to the instruments (Z_V, Z_A). In principle, these could consistently be estimated from local linear regressions across the empirical support of the instruments. However, as independence and exclusion of the instruments (A1) is only plausible after conditioning on our control vector $\mathbf{X}$, the curse of dimensionality prohibits a fully non-parametric estimation procedure. To avoid imposing a restrictive constant and linear relationship between distance and choices, we follow Mountjoy's (2019) suggestion and use a version of Hastie and Tibshirani's (1993) varying coefficient model: we estimate locally linear regressions across the bi-variate distribution of the two distance instruments (thus reducing the dimensionality to two), where all variables enter additively but with coefficients that are allowed to vary arbitrarily across different (z_V, z_A) evaluation points. Formally, for a given variable $W \in \{D_N, D_V, D_A, Y, YD_N, YD_V, YD_A\}$, the coefficients are estimated by solving a kernel-weighted least squares problem at each (z_V, z_A) evaluation point:

$$\begin{pmatrix} \hat{\alpha}^W\left(z_V, z_A\right) \\ \hat{\beta}_V^W\left(z_V, z_A\right) \\ \hat{\beta}_A^W\left(z_V, z_A\right) \\ \hat{\beta}_x^W\left(z_V, z_A\right) \end{pmatrix} = \underset{\alpha,\beta_V,\beta_A,\beta_x}{\operatorname{argmin}} \sum_{i=1}^{N} K\left(\frac{Z_{Vi}-z_V}{h}, \frac{Z_{Ai}-z_A}{h}\right)\left(W_i - \alpha - \beta_V Z_{Vi} - \beta_A Z_{Ai} - \mathbf{X}_i'\beta_x\right)^2,$$

where $K()$ is a two-dimensional Epanechnikov kernel function and the bandwidth, h, is set to 15km. As mentioned above, we report our main results evaluated at the mean values of the instruments. Subsequently, we evaluate the model across 7×7 grid points (along both distance to college dimensions we move in increments of three from 0km to 18km) to assess the representativeness of the mean results and to inspect selection patterns.

To purge the distance instruments of correlations with factors that might directly influence education choices, we construct our control set, $\mathbf{X}$, as flexibly as possible. It contains the following variables: cohort fixed effects; fixed effects for the nine regions of England; a cubic polynomial in distance to economic centre, the secondary school share of FSM eligible students and the secondary school share of White British students all interacted with the nine region fixed effects;

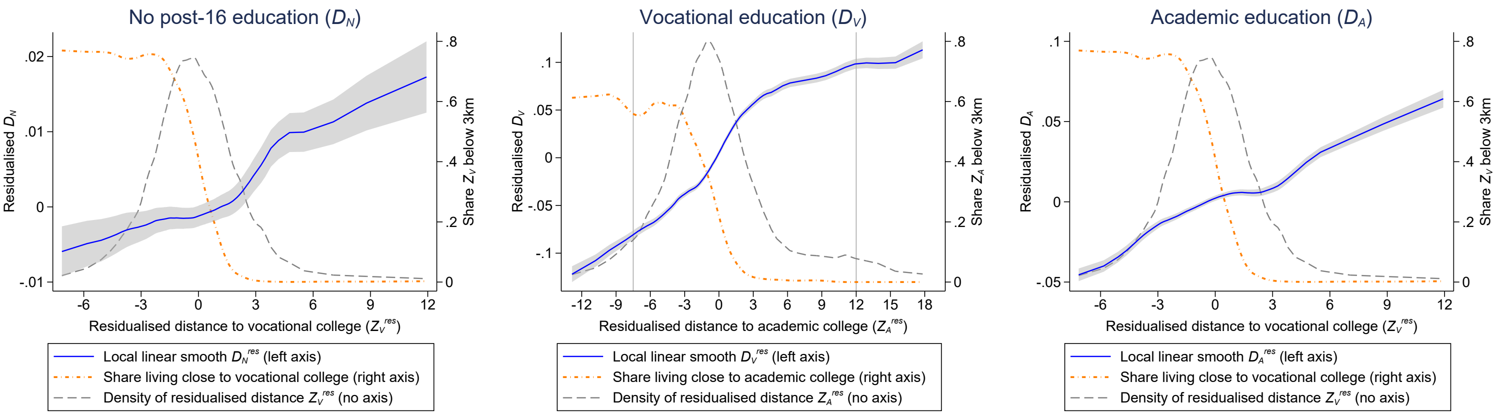


Figure 5. Education choices by distance to college.

Notes: This figure displays the three first stage relationships of interest in the IV sample. For no post-16 education (left panel) and academic education (right panel) it depicts the (conditional) relationship between the treatment indicator and distance to vocational college, both residualised with respect to distance to academic college and the full control set, plotting fitted values from local linear regression at each percentile of the residualised distance, including 95%-confidence sets. The middle panel proceeds analogously to depict the relationship between vocational education and distance to academic college. On top of this, the figure shows fitted values for an indicator for students living close to the respective college type and kernel density estimates for the respective residualised distance variable. All three kernel estimators use adaptive bandwidths, inversely proportional to the density of the residualised distance (Abramson, 1982).

secondary school share of English as a second language students; secondary school averages in students' Key Stage 2 (KS2) test scores in English, maths and science; cubic polynomials in all eight neighbourhood quality indices; cubic polynomials in students' KS2 English, maths and science test scores interacted with students' gender and, finally, all two-way interactions between indicators for students' gender, ethnicity, FSM eligibility, language status and special educational need status.[15]

5 Validity of the Research Design

5.1 First Stages

This section presents our first stage estimates as a first steps towards assessing the validity of our research design. As can be seen in (5), our identification approach involves three first stage relationships of interest: the (conditional) effect of distance to vocational college, Z_V, on no post-16 education, D_N, and academic education, D_A, and the (conditional) effect of distance to academic college, Z_A, on vocational education, D_V. In Figure 5 we visualise these empirical relationships by plotting local linear smooths of the residualised treatment indicators across the distributions of the respective residualised distance instruments. We residualise through

[15]We do not include students' GCSE scores in the control set because these are arguably 'bad controls': students have different incentives to perform in their GCSE exams depending on which track they choose so that performance in these exams might be influenced by treatment choice.

Table 3. First stages with and without test score controls.

Dependent variable:	No post-16 educ. (D_N)		Vocational educ. (D_V)		Academic educ. (D_A)	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance vocational college (Z_V)	0.0015***	0.0013***	-0.0067***	-0.0068***	0.0053***	0.0055***
	(0.0003)	(0.0003)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
		[24.6]				[160.9]
Distance academic college (Z_A)	0.0010***	0.0009***	0.0111***	0.0112***	-0.0121***	-0.0121***
	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
				[1898.6]		
Key Stage 2 test scores		✓		✓		✓
Remaining controls	✓	✓	✓	✓	✓	✓
R^2	0.06	0.08	0.06	0.09	0.18	0.25
N clusters	21,732	21,732	21,732	21,732	21,732	21,732
N students	388,927	388,927	388,927	388,927	388,927	388,927

Notes: The table displays results from locally weighted OLS regressions, where observations are weighted by their proximity to the instrument means using a two-dimensional Epanechikov Kernel and a bandwidth of 15km. The three KS2 test scores enter as gender-specific cubic polynomials. Standard errors, reported in parentheses, are clustered at the LSOA level. 'Effective' F-statistics (Olea and Pflueger, 2013) for the first stage coefficients of interest are are reported in squared brackets. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

regressing on the other distance instrument and the full control set.

The first observation from Figure 5 is that the instruments predict choices as expected: conditional on distance to academic college (and the other controls), larger distances to vocational college are associated with a higher probability of both no post-16 (left panel) and academic education (right panel), while, conditional on distance to vocational college, larger distances to academic college are associated with a higher probability of vocational education (middle panel). Second, these relationships are weakly monotonic, lending credence to A2. However, third, the strength of these relationships is not constant. As residualised distances do not directly translate into absolute ones, we also plot the share of students living very close ($<$ 3km) to the respective college across the residualised distance distribution. The strongest effects of distance on choices seem to be found where the share of students living very close is rather low. It makes intuitive sense that distance mainly plays a role at medium distances, whereas marginal changes in distance matter less when students live close or far to colleges. In particular, the treatment of no post-16 education, which appears least responsive to distance (compare the scales of the vertical axes), seems to have a rather weak first stage at distances too small or too far. The other two first stages, however, are strong across the whole range. The first stage for vocational education is clearly strongest, though this is partly due to the fact that distance to academic college exhibits more conditional variation. Restricting the comparison to a range that is shared between Z_V and Z_A (indicated by two vertical bars in the middle panel) the "cross-effects" of distance to vocational and academic college on choices are quite similar, as one would expect.

We now turn to the first stage regressions, estimated using the local linear specification introduced in the previous section and evaluated at the mean values of the instruments. Table

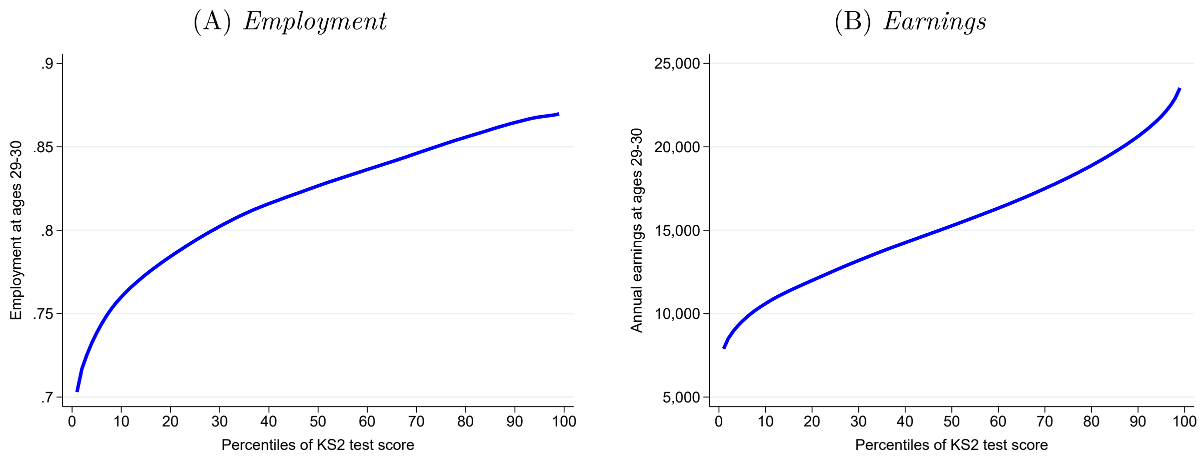


Figure 6. Labour market outcomes by KS2 test score percentile.

Notes: This figure shows fitted values for average employment and annual earnings at age 29–30 (in 2010 pounds) across percentiles of the first principal component of the KS2 English, Maths and Science scores, estimated via local linear regressions.

3 presents estimates for the three first stages, comparing each across ex- and inclusion of the test score controls, which enter as gender-specific cubic polynomials. The estimates are remarkably unaffected by this change in the control set, indicating that our demographic, school and neighbourhood controls suffice for purging the distance-choice relationships of residential sorting. Of course, this test is only meaningful if test scores predict not only choices, as shown in section 3, but also outcomes. Indeed, Figure 6 shows a tight relationship between test scores and labour market performance: for example, average annual earnings at ages 29–30 more than double across the test score distribution.

All coefficients in Table 3 show their intuitive signs: distance to vocational college decreases the probability of vocational education but increases that of no post-16 and academic education, while distance to academic college decreases the probability of academic education but increases that of vocational education. For the three coefficients of interest, in squared brackets we also report so-called 'effective' F-statistics, proposed by Olea and Pflueger (2013) to test first-stage strength under heteroskedasticity and clustering. All of them exceed their rule-of-thumb critical value of 23.1, indicating that, at the mean values of the instruments, the strength of all three first stages suffices to draw inference.[16] Nevertheless, the case of D_N is close and warrants special care when interpreting the results.

[16]Similarly, empirical p-values for the null hypothesis of a zero effect from a block bootstrap at the LSOA-level using 999 repetitions are zero for all three coefficients of interest, so that our first stages pass both tests recommended by Young (2020).

Table 4. Instrument balance tests.

Dependent variable:	White (1)	FSM (2)	KS2 English (3)	KS2 Maths (4)	KS2 Science (5)
Distance vocational college (Z_V)	-0.0005	-0.0001	-0.0002	-0.0004	0.0003
	(0.0003)	(0.0002)	(0.0006)	(0.0005)	(0.0006)
Distance academic college (Z_A)	-0.0009***	0.0006***	-0.0002	-0.0005*	0.0011***
	(0.0002)	(0.0001)	(0.0003)	(0.0003)	(0.0003)
White British		✓	✓	✓	✓
Free school meal (FSM)	✓		✓	✓	✓
KS2 score English	✓	✓		✓	✓
KS2 score Maths	✓	✓	✓		✓
KS2 score Science	✓	✓	✓	✓	
Remaining demographics	✓	✓	✓	✓	✓
School characteristics	✓	✓	✓	✓	✓
Neigbourhood characteristics	✓	✓	✓	✓	✓
Region FEs	✓	✓	✓	✓	✓
N clusters	21,732	21,732	19,791	19,913	20,088
N students	388,927	388,927	333,395	337,931	339,900

Notes: The table displays results from locally weighted OLS regressions, where observations are weighted by their proximity to the instrument means using a two-dimensional Epanechikov Kernel and a bandwidth of 15km. Standard errors, reported in parentheses, are clustered at the LSOA level. Missing values in the KS2 scores for some students explain the lower number of observations in columns 3–5. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2 Balance Tests

The first stage estimates' invariance to the in- and exclusion of the test score controls provided strong initial evidence in favour of the independence and exclusion assumption (A1). Here, we present instrument balance tests as a more direct test of this assumption: one at a time, we exclude the White British indicator, free school meal (FSM) eligibility indicator and the three test scores from the control set to regress the excluded variable on the two distance instruments and the remaining controls. This way we assess covariate balance with respect to changes in distance.

Table 4 presents the results. For distance to vocational college not a single significant coefficient is observed, indicating perfect balance. For distance to academic college we observe some statistically significant associations, though economically these are extremely small. For example, an additional kilometre in distance to academic college is associated with only a 0.06 percentage point increase in the probability of FSM eligibility and a 0.1 percent of a standard deviation increase in KS2 science scores. Note that the former is a sign of disadvantage and the latter a sign of advantage so this does not point at a clear directional pattern of sorting. Their small size, inconsistent pattern and the fact that this exercise gives an upper bound of the remaining selectivity because each time one control variable (and its interactions) is excluded from the control set lead us to the the conclusion that we can safely ignore any potential remaining selectivity.

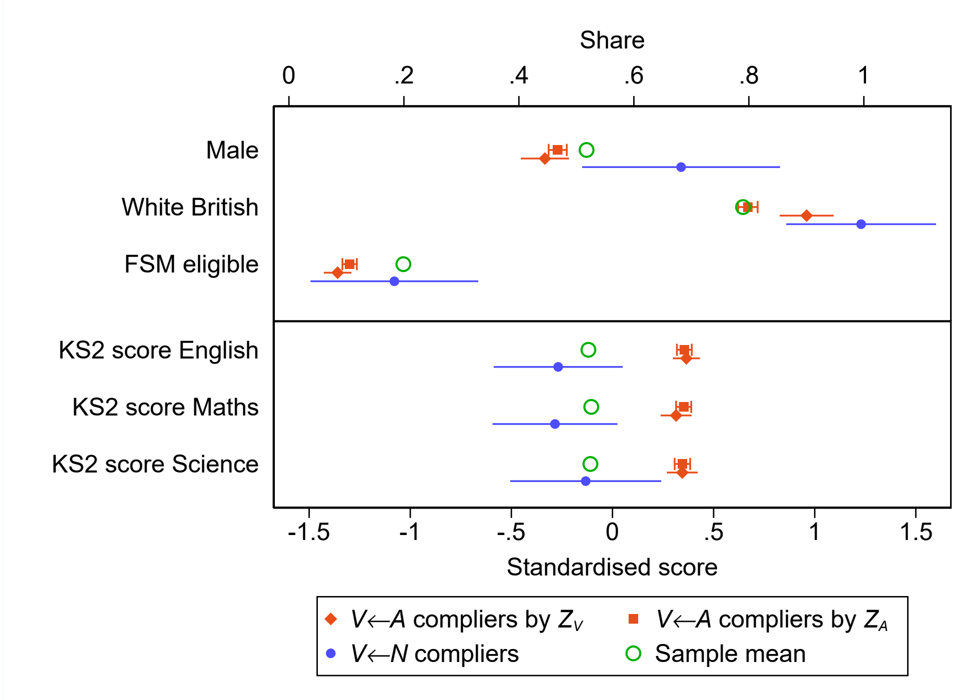


Figure 7. Complier characteristics.

Notes: The figure shows estimated mean characteristics, evaluated at the mean values of the instruments, for three different complier groups: vocational-academic compliers induced by Z_A, vocational-academic compliers induced by Z_V and vocational-no post-16 compliers induced by Z_V. Estimation for characteristic X is based on kernel-weighted 2SLS regressions of, respectively, XD_V on D_V instrumenting with Z_A; XD_A on D_A instrumenting with Z_V; and XD_N on D_N instrumenting with Z_V, each time controlling for the other instrument and full control set $\mathbf{X}$ and clustering at the LSOA level. For vocational-academic compliers induced by Z_A we report 95% confidence intervals. For the other two groups, confidence intervals are adjusted as described in Wright *et al.* (2019) so that non-overlap with the former indicates a statistically significant mean difference.

5.3 Complier Characteristics

So far, we have presented evidence in favour of the plausibility of assumptions A1 and A2. For the internal validity of our estimates, we further require complier comparability (A3), namely that compliers at the margin between vocational and academic education induced by Z_V and Z_A are comparable in terms of their mean potential outcomes. To this end we compare the two complier sets in terms of a number of pre-determined observable characteristics. We can do so because, replacing Y with any covariate X, the second and third equations of (5) separately identify the mean of X for Z_V- and Z_A-compliers at the vocational-academic margin (Mountjoy, 2019). The results from this exercise, presented in Figure 7, are reassuring: at the mean values of the instruments, the two complier types are indistinguishable in all but one of six variables and, in particular, their KS2 scores, which are by far the most predictive for earnings, match precisely.

Beyond testing complier comparability, this exercise allows us to describe the complier subpopulation for whom we can estimate causal effects. To this end Figure 7 also reports the characteristics of compliers at the vocational-no post-16 education margin, which are identified by the first equation of (5). Note that their characteristics are estimated with less precision because of the weaker first stage for D_V with respect to Z_V. Nonetheless, the figure reveals stark differences between the two complier populations. Compared to the average, students at the margin between vocational and no post-16 education are more likely to be male and White British and perform slightly worse academically. Students at the margin between vocational and academic education, in contrast, are less likely to be male, less disadvantaged and perform above average academically.

We have contrasted compliers at the two margins in terms of observable characteristics and below we will estimate their respective shares. Before we do so, however, it is interesting to assess the overall size of the complier population. Dahl *et al.* (2014) show that it can be estimated by comparing treatment take-up at the extreme values of the instrument: the share of students who choose D_V at maximum distance to vocational college equals the share of always-takers, while the share of students who do not choose D_V at minimum distance to vocational college equals the share of never-takers. The rest are compliers would have chosen a different post-16 education at at least some point of the distance distribution. We estimate the share of always-takers as the mean of D_V at the 99^{th} percentile of the residualised Z_V distribution and the share of never-takers as the mean of $(1 - D_V)$ at the 1^{st} percentile, which equal 0.45 and 0.38, respectively. Accordingly, compliers make up about 17% of non-integrated school students.

6 Results

6.1 Main Results

In this section we present our main findings for the margin-specific returns to vocational education. We report estimates for marginal treatment effects (MTEs), evaluated at the mean values of the instruments. Accordingly, strictly speaking, these results pertain to a specific point in the two-dimensional distance to colleges distribution. In section 6.4 below, we assess the representativeness of these estimates and show that they are, in fact, close to average effects.

Table 5 presents the estimation results. The first three columns present estimates for the full (mean-weighted) sample. The remaining columns show results separately for women (columns 4–6) and men (columns 7–9). Within each column triple, the first column reports the estimated share of compliers at each margin and the second and third column report results for employment and earnings at ages 29–30, respectively.

The first row in bold shows the net complier treatment effect of vocational education, amalgamating both margins (MTE_V in (6)). For the average complier induced by closer access to vocational college, vocational upper secondary education increases annual earnings by about £2,500. This corresponds to about 15% of the pooled counterfactual complier potential outcome mean of $£15,761$. This effect is partly driven by extensive margin effects, as indicated by a

Table 5. Margin-specific effects of vocational education on employment and earnings at ages 29–30.

	All students			Females			Males		
	Complier share (1)	Employ-ment (2)	Annual earnings (3)	Complier share (4)	Employ-ment (5)	Annual earnings (6)	Complier share (7)	Employ-ment (8)	Annual earnings (9)
Net effect vocational educ.		**0.056**	**2,516****		**0.025**	**1,871**		**0.076**	**2,653**
		(0.038)	**(1,215)**		**(0.055)**	**(1,392)**		**(0.053)**	**(2,017)**
No post-16 education margin	0.196			0.113			0.280		
	(0.033)			(0.043)			(0.047)		
$\mathbb{E}[Y_V \mid V \leftarrow N \text{ complier}]$		1.009	30,085		0.649	22,421		1.155	32,939
		(0.158)	(6,126)		(1.083)	(43,064)		(0.162)	(6,744)
$\mathbb{E}[Y_N \mid V \leftarrow N \text{ complier}]$		0.844	13,006		0.601	8,546		0.958	15,363
		(0.106)	(2,482)		(0.568)	(17,436)		(0.114)	(3,070)
$MTE_{V \leftarrow N}$		**0.165**	**17,079****		**0.047**	**13,874**		**0.197**	**17,576****
		(0.181)	**(6,078)**		**(1.473)**	**(38,317)**		**(0.185)**	**(6,973)**
Academic education margin	0.804			0.887			0.720		
	(0.033)			(0.043)			(0.047)		
$\mathbb{E}[Y_V \mid V \leftarrow A \text{ complier}]$		0.853	15,404		0.811	13,523		0.903	17,634
		(0.009)	(279)		(0.014)	(323)		(0.012)	(454)
$\mathbb{E}[Y_A \mid V \leftarrow A \text{ complier}]$		0.824	16,433		0.789	13,176		0.874	20,779
		(0.022)	(985)		(0.030)	(1,125)		(0.031)	(1,623)
$MTE_{V \leftarrow A}$		**0.029**	**-1,029**		**0.023**	**347**		**0.029**	**-3,145***
		(0.024)	**(1,015)**		**(0.033)**	**(1,165)**		**(0.033)**	**(1,672)**

Notes: This table presents complier share, potential outcome and, in bold, marginal treatment effect estimates, evaluated at the mean values of the instruments. The number of locally weighted observations is 388,927 in the full sample, 188,078 in the female sample and 200,849 in the male sample. Standard errors are block bootstrapped at the LSOA-level using 999 iterations. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

positive effect on employment of 5.6 percentage points (pp) that is close to significant. Estimates are not significant in the smaller gender-specific samples, but the point estimates indicate that, on net, men profit more from vocational education than women.

The remaining rows decompose these net complier treatment effects into margin-specific effects. The first block of rows pertains to the vocational-no post-16 education margin and the second block pertains to the vocational-academic education margin. Within each block the first row shows the estimated share of compliers, the next two rows show potential outcome estimates and the final row in bold shows the corresponding margin-specific treatment effect.

For about 20% of compliers the alternative to vocational is no post-16 education. For these students, vocational education appears hugely beneficial, as indicated by a very large (and significant) effect of about £$17,000$ on earnings, which, at least in part, is driven by a similarly large (but insignificant) 17pp effect on employment. The estimates at this margin are quite imprecise as the expression for the potential outcome under vocational education ($\mathbb{E}[Y_V|V \leftarrow N$ complier]) involves three separately estimated partial derivative ratios, making it very demanding to estimate. This is amplified by the fact that they rely on the weaker first stage of D_N with respect to Z_V, which additionally might inflate the estimates. Accordingly, the size of these estimates should be taken with a grain of salt. Nevertheless, they indicate that for marginal students gains from vocational compared to no upper secondary education are large, clearly exceeding the OLS estimates from section 3.6.

The results by gender indicate that this effect is mainly driven by men, who represent the vast majority of compliers at the no post-16 margin (28% of male compliers are at this margin compared to only 11% for females). For them, the estimated effect for earnings corresponds to that of the full sample and the one for employment even exceeds it. For the few women, inflated standard errors (due to an even weaker first stage for D_N) prohibit precise statements. At least qualitatively, the point estimate suggests a similarly large effect on earnings.

For the other 80% of compliers, for whom the alternative to vocational is academic education, results look very different, however. In the full sample, vocational education appears to be a neutral choice at best, as indicated by a small negative (but insignificant) effect of about £$1,000$ on earnings without substantial differences in employment. Splitting the sample by gender, we find that the effect on earnings is composed a zero effect for women and a substantial negative effect of about £$3,000$ for men. Compared to the counterfactual, where they would have chosen academic upper secondary education, these students earn 15% less annually. Importantly, this difference seems to be due to higher wages (and potentially higher working hours) as there are no differences in the probability of being employed (if anything vocationally educated students are slightly more likely to be employed).

In sum, our results suggest that vocational education is a double-edged sword: highly beneficial for students that otherwise would not have pursued upper secondary education, but detrimental for male students that otherwise would have pursued academic upper secondary education. The

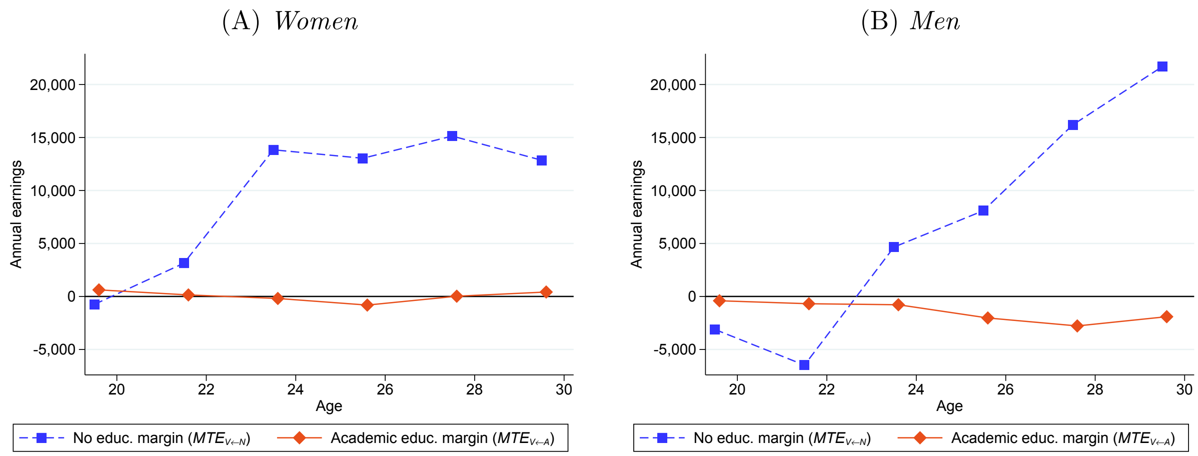


Figure 8. Margin-specific returns by age.

Notes: This figure display point estimates for the margin-specific effects of vocational education on average annual earnings at ages 19–20, 21–22, 23–24, 25–26, 27–28 and 29–30, separately by gender. The same sample sizes as in Table 5 apply. In contrast to the main results, earnings from self-employment are excluded to ensure comparability across the age range.

estimated complier shares reveal that the vast majority of compliers fall into the latter category. This means that an untargeted expansion of vocational colleges would harm the majority of students even though the net effect of vocational education is positive. Accordingly, identification of the margin-specific effects proves crucial in this context.[17]

6.2 Effects by Age

To get an idea of the dynamics of the returns of vocational education, Figure 8 plots the estimated margin-specific effects on average annual earnings at six age bands between 19–20 and 29–30, separately by gender. Note that for this exercise earnings from self-employment have to be excluded to ensure comparability across the age range, which has little effect on the estimates for women but leads to an underestimate (overestimate) of the effect of vocational *vs.* academic education (vocational *vs.* no post-16 education) for men. The figure reveals that the large positive effect of vocational education at the no post-16 education margin take about six years to materialise. For women, it remains roughly constant from ages 23–24 onward. For men, it continues to grow until the age of 30. At the academic education margin, the zero effect of vocational education for female compliers is stable over the entire age range. The negative effect for male compliers materialises in students' mid-twenties and remains roughly constant thereafter.

[17]Not only do our analysis reveal that the positive net effect is based on a small minority of students, they also show that the first stage relationship at that margin is weaker so that the positive net effect could actually be due to inflated point estimates among the group diverted from no post-16 education.

Table 6. Effects on education outcomes for vocational-academic compliers.

Parameter:	Females			Males		
	Vocational PO	Academic PO	Difference	Vocational PO	Academic PO	Difference
	$\mathbb{E}[Y_V \mid V \leftarrow A]$	$\mathbb{E}[Y_A \mid V \leftarrow A]$	$MTE_{V \leftarrow A}$	$\mathbb{E}[Y_V \mid V \leftarrow A]$	$\mathbb{E}[Y_A \mid V \leftarrow A]$	$MTE_{V \leftarrow A}$
	(1)	(2)	(3)	(4)	(5)	(6)
A. Course content						
Share vocational	0.400	0.255	**0.145*****	0.493	0.227	**0.266*****
	(0.017)	(0.028)	**(0.032)**	(0.023)	(0.035)	**(0.043)**
Share academic	0.688	0.800	**-0.112****	0.651	0.966	**-0.315*****
	(0.018)	(0.038)	**(0.042)**	(0.021)	(0.051)	**(0.055)**
B. Attainment						
Level 3 qualification	0.818	0.784	**0.034**	0.812	0.847	**-0.034**
	(0.017)	(0.025)	**(0.030)**	(0.020)	(0.032)	**(0.038)**
Apprenticeship	0.168	0.172	**-0.004**	0.129	0.087	**0.042**
	(0.015)	(0.024)	**(0.028)**	(0.018)	(0.030)	**(0.035)**
University degree	0.367	0.296	**0.071**	0.370	0.365	**0.005**
	(0.013)	(0.045)	**(0.046)**	(0.014)	(0.054)	**(0.055)**

Notes: The number of locally weighted observations is 188,078 in the female sample and 200,849 in the male sample. All estimates are evaluated at the mean values of the instruments. Standard errors are block bootstrapped at the LSOA-level using 999 iterations. Stars indicate significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.3 Effects on Educational Outcomes

Our estimates show that vocational education is an improvement over no upper secondary education – a (perhaps unsurprising) result in line with the theoretical predictions of canonical human capital and signalling models. As highlighted in the introduction, the relative merits of vocational education as compared to academic upper secondary education are much less clear in theory. Theoretical predictions crucially depend on the types of skills that students acquire in the respective educational tracks and on the opportunities for higher education they face afterwards. At this margin, our results suggest that the earnings of women are largely unaffected by their track choice, whereas those of men suffer from vocational education. To get a better understanding of the mechanisms behind this finding, in this section we investigate how enrolling in the vocational instead of the academic track affects the types of courses students attend during their upper secondary education and their final educational attainment.

Panel A of Table 6 shows the results for the type of courses students attend in upper secondary education. Comparing the estimated potential outcomes with the sample averages for the shares of vocational and academic courses (see Figure 1) reveals that vocational-academic compliers are students that make full use of the built-in flexibilities of the English vocational education system: while the average vocational-track student attends 57% vocational, 18% academic and 25% basic and soft skill courses, compliers at this margin attend slightly more than half academic and slightly less than half vocational courses even when they attend a vocational college. When they enrol in the academic track, they behave like typical academic-track students and attend

80-90% academic and only very few vocational courses. Though visible for both genders, for men these differences in courses attended between treatments are far more pronounced than for women. This might explain the earnings penalty male students experience from choosing vocational education: unlike their female peers, they fail to make full use of the possibility offered by vocational colleges to study many academic subjects.

Panel B shows that the effects on several measures of educational attainment are insignificant. About 80% of vocational-academic compliers achieve at least one upper secondary Level 3 qualification, regardless of which track they choose This is far above the average for vocational-track students and slightly below the average for academic-track students (see Figure 3). Furthermore, under vocational education, these compliers are half as likely to start an apprenticeship but far more likely to complete a university degree than the average vocational-track student. For both of these outcomes, treatment effects are insignificant, but, if anything, vocational education boosts university degree completion for women and apprenticeship take-up for men, reflecting the gender differences in the effects on course content and potentially contributing to the differential effects on earnings. More generally, the fact that educational attainment of marginal students is not significantly affected by the track they attend implies that the negative earnings effects (for men) must stem from problems specific to upper secondary vocational education instead of subsequent educational pathways.

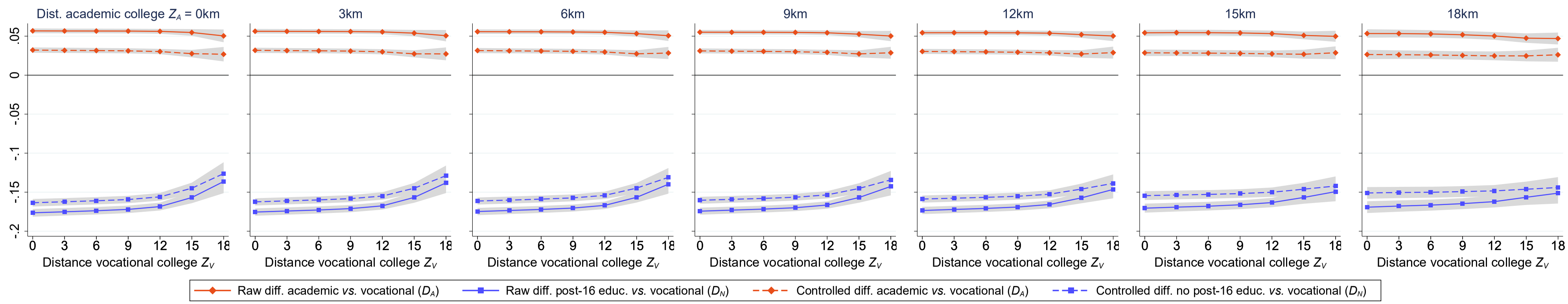
6.4 Effects across the Distance Grid

So far, we evaluated all our models at the mean values of the instruments. In this section, we stratify the results across different points in the two-dimensional distance to academic and vocational colleges distribution. The motivation for this is two-fold: first, it allows us to assess the external validity of the local estimates at the mean and paint a more comprehensive picture of the (margin-specific) returns to vocational education. Second, it allows us to explore potential patterns of selection into educational choices.

To establish a benchmark for the IV results, Figure 9 presents results from estimating the simple OLS model of (1) with and without controls, evaluated at different distances to vocational and academic colleges. Along either dimension (i.e. distance to vocational college, Z_V, and distance to academic college, Z_A) we move in increments of 3km from 0km to 18km, so that the distance grid along which we estimate consists of 49 points in total. Panel A for employment and panel B for earnings show that raw and controlled outcome differences between the treatment groups are remarkably stable across the distance grid. This confirms our conclusions from Table 2 that average treatment effects do not systematically vary with students' residential location.

Note that the fact that average effects do not vary much with distance does not imply that the same holds for complier treatment effects. This is because the population of compliers, i.e. students at the margin between two tracks, is likely to be different at different distances: for

(A) *Employment*



(B) *Earnings*

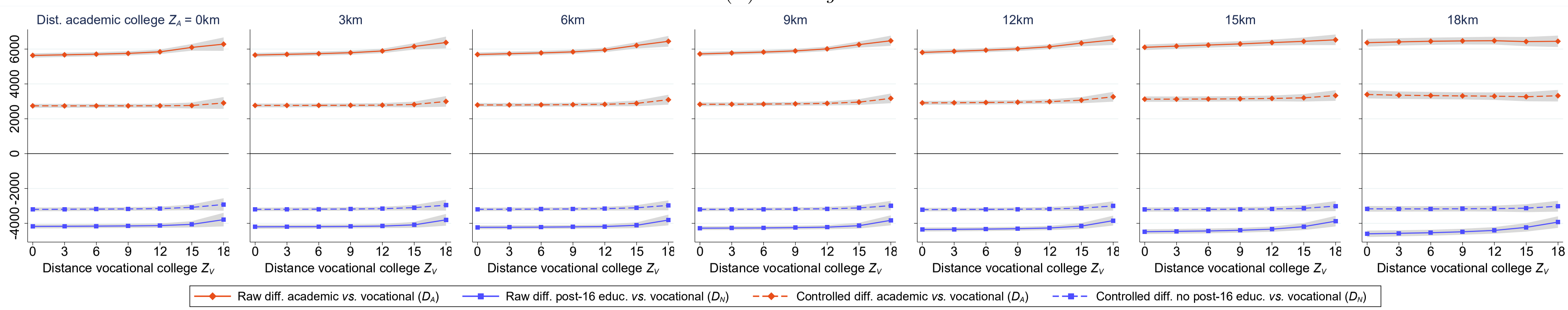


Figure 9. Local OLS regression results for employment and earnings at ages 29–30 across the distance grid.

Notes: This figure shows raw and controlled differences in employment (panel A) and earnings (panel B) between (i) academic and vocational education and (ii) no post-16 and vocational education across the distance grid. Estimation is based on locally weighted OLS regressions of equation (1) with and without controls, **X**. At each grid point, observations are weighted by their proximity to the evaluation point using a two-dimensional Epanechnikov kernel with 12.5km bandwidth. 95% confidence intervals are based on standard errors that are clustered at the LSOA level.

example, all else equal, students living close to a vocational college face lower costs for entering the vocational track than students living far (they need to travel less). Accordingly, the closer the distance to a particular college type, the higher should be the share of students self-selecting into the corresponding track. This, in turn, implies that marginal (i.e. compliant) students at closer distances must have lower unobserved preference for the education choice in question than marginal students living far. If students' unobserved preferences reflect perceived gains from the different education choices, as in a standard Roy (1951)-style selection model, we would thus expect lower treatment effects at closer distances. Therefore, inspecting the effect heterogeneity in our IV models across the distance grid, allows us to explore how potential outcomes and treatment effects relate to student self-selection.

Before we estimate the IV models across the distance grid, we need to assess the validity of the identifying assumptions. For this, we proceed analogously to before for the mean results (section 4). Appendix Figure A2 plots the three first stage coefficients of interest across the distance grid, confirming that these have the expected signs and are strong at all points – potentially except for D_N, which remains significant everywhere but is small at distances too close or far. Remember that Figure 5 showed that the empirical relationship between education choices and the distance instruments is weakly monotonic over the entire distance distribution. Together this suggests that partial unordered monotonicity (A2) is likely to be satisfied across all points in the grid.

To assess the plausibility of independence and exclusion of the instruments (A1), we again present instrument balance tests. For this, we focus on the most predictive covariate for selection into treatments and outcomes – KS2 test scores. For brevity, we simultaneously exclude all three test scores from the control set and regress their first principal component on the remaining controls and our two distance instruments. Panel A of Appendix Figure A3 plots the coefficients for the two distance measures, showing that these are close to zero across the entire grid, even in this particularly conservative test where all three ability measures are excluded at once. Accordingly, we see no reason to suspect that A1 does not hold across the entire grid.

Finally, to assess the plausibility of complier comparability (A3), as before, we compare $V \leftarrow A$ compliers induced by Z_V with those induced by Z_A in terms of pre-determined observable characteristics. We again focus on the first principal component (PC) of KS2 test scores as the most predictive covariate, for ease of interpretation converted into percentiles. Panel B of Appendix Figure A3 plots the estimated KS2 PC means for the two complier types across the distance grid. Problematically, the results show that at some points of the distance distribution the two complier types differ in terms of their test scores. While at medium and far distances to either college type compliers are close to perfectly comparable, at points close to either college type compliers induced by distance to vocational college have lower test scores on average than those induced by distance to academic college. Though these differences are modest, at most 7 percentile points, this suggests that, at those grid points, potential outcomes identified from conditional variation in Z_A slightly overestimate those of Z_V-induced compliers that we are interested in.

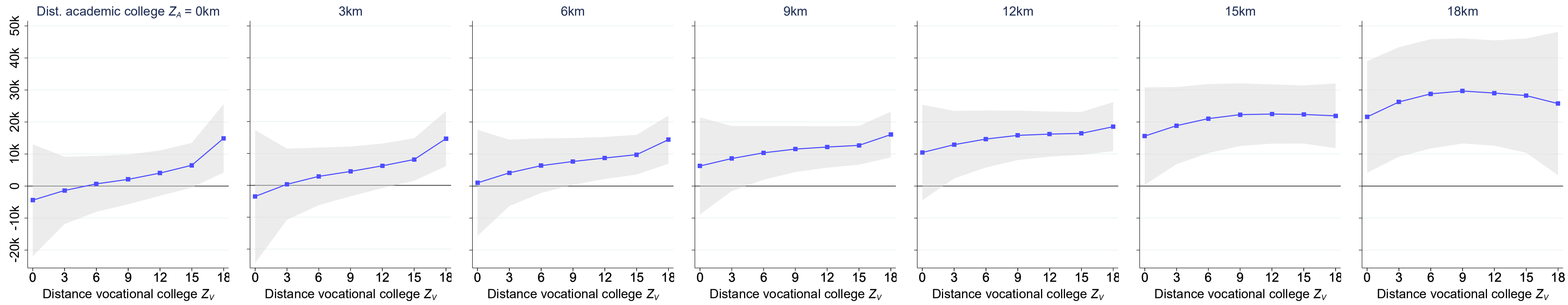
To solve this problem we exploit the tight relationship between KS2 test scores and earnings. In particular, at each point of the distance grid, we estimate the relationship between earnings and KS2 test scores and the gap in KS2 test scores between the two complier types. Equipped with these two pieces of information, we then correct the estimated potential outcome for Z_A-induced compliers for the difference in KS2 scores compared to Z_V-induced compliers, imposing mild parametric assumptions.[18] Appendix Figure A4 plots the original and corrected potential outcome estimates across the distance grid. Reassuringly, and expectedly given the modest differences in KS2 scores, the effects from this correction are small and do not alter any of our conclusions.

Figure 10 plots estimates for the two margin-specific marginal treatment effects of vocational education on earnings, constructed from the corrected potential outcomes, across the distance grid. The results show that effects estimated at the instrument means ($Z_V = 5.1, Z_V = 13.3$) are indeed close to average effects at both margins, but, at the same time, they point at non-negligible effect heterogeneity across the distances distribution. Panel A shows that the effect of vocational *vs.* no post-16 education grows with distance to either college type and is large and positive at all distances, except for students living close to both college types for whom effects seem to be null. This pattern is consistent with the notion that students who are prepared to sustain a higher cost (travel longer distances) to enrol into vocational education do so because they expect a higher pay-off. Note that the share of compliers at either margin is roughly constant across the grid (see panel C of Appendix Figure A4).

The results for vocational *vs.* academic education in panel B show zero effects for students living close to both college types (for students living very close effects may even be slightly positive), negative but insignificant effects at medium distances and large negative and significant effects for students living farther away from either college type. The finding that the effect of vocational education becomes more negative with increased distance to academic college is unsurprising and suggests selection on gains into the academic track. The finding that conditional on distance to academic college, the effect becomes more negative with increased distance to vocational college is more surprising because it suggests reverse selection on gains into the vocational track. However, panel B of Appendix Figure A3 shows that compliers living farther

[18] More precisely, at each grid point (z_V, z_A), we proceed as follows: First, in the sample of vocational- and academic-track students (i.e. excluding no post-16 education students) we estimate a locally weighted regression of earnings on an indicator for vocational education and the KS2 principal component, $Y = \alpha(z_V, z_A) + \delta(z_V, z_A)\, D_V + \beta(z_V, z_A)\, KS2 + \varepsilon$, and record the local estimate $\beta(z_V, z_A)$. Second, we estimate the local gap in KS2 scores between $V \leftarrow A$ compliers induced by Z_A and those induced by Z_V, $\Delta KS2(z_V, z_A)$, as in Appendix Figure A3. Third, we correct the estimate for vocational-academic compliers' vocational education potential outcome, $\mathbb{E}\left[Y_V \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)\right] = \frac{\frac{\partial \mathbb{E}[Y D_V | Z_V, Z_A]}{\partial Z_A}}{\frac{\partial \mathbb{E}[D_V | Z_V, Z_A]}{\partial Z_A}}$, which pertains to Z_A-induced compliers, using the estimated gap and the estimated coefficient as follows: Let $\hat{Y}_V^{orig}(z_V, z_A)$ denote the original local estimate of this potential outcome. We replace it with $\hat{Y}_V^{new}(z_V, z_A) = Y_V^{\hat{orig}}(z_V, z_A) - \hat{\beta}(z_V, z_A)\, \Delta K\hat{S}2(z_V, z_A)$. Note that this leads to a correction at both margins because the the estimate for $\mathbb{E}\left[Y_V \mid V \leftarrow A \text{ complier at } (Z_V, Z_A)\right]$ is also used in the calculation of $\mathbb{E}\left[Y_V \mid V \leftarrow N \text{ complier at } (Z_V, Z_A)\right]$ (see (5)).

(A) *Vocational vs. no post-16 education margin*



(B) *Vocational vs. academic education margin*

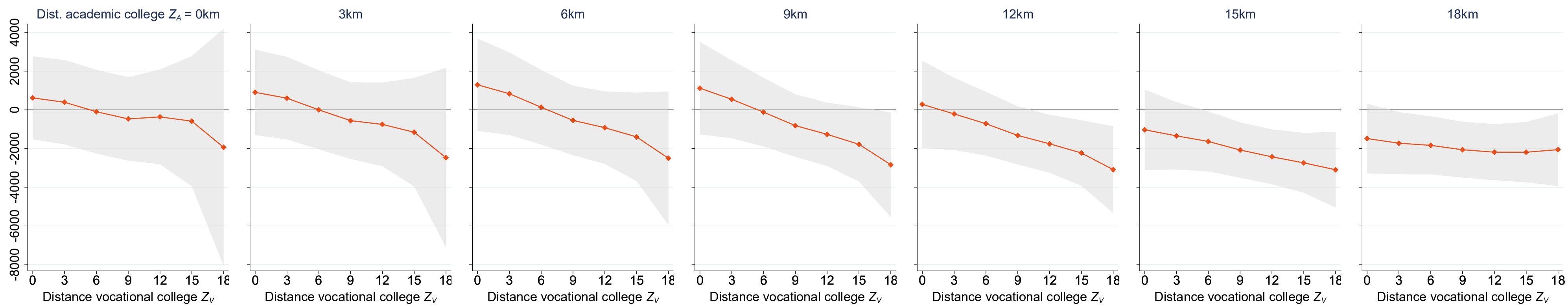


Figure 10. Marginal treatment effect estimates for earnings at ages 29–30 across the distance grid.

Notes: This figure shows point estimates for the marginal treatment effect (MTE) of vocational education *vs.* no post-16 education (panel A) and for the MTE of vocational education *vs.* academic education (panel B) evaluated across all 49 points of the two-dimensional distance grid. At each point, the effects are estimated using locally weighted regressions analogously to the results at the mean reported in Table 5. 90%-confidence intervals are constructed using a block bootstrap at the LSOA-level with 220 repetitions.

away from both vocational and academic colleges have higher test scores on average, so that it is unsurprising that these students benefit more from choosing academic rather than vocational education.

7 Discussion and Conclusions

Debates on the merits of vocational education are often framed in terms of a dichotomy between vocational and academic education. One the one hand, opponents of vocational education argue that academic curricula improve students' chances on the labour market by equipping them with general knowledge and analytical skills that are transferable across occupations. On the other hand, proponents argue that vocational education increases the employability of less academically inclined students who are at risk of dropping out of academic tracks or who fail to complement their academic secondary education with higher education. In this paper we estimate labour market returns to vocational upper secondary education in England but move beyond the dichotomous perspective in an effort to reconcile the two opposing views: using distance to different post-16 education providers as instrumental variables in an identification framework proposed by Mountjoy (2019), we empirically decompose the net effect of vocational upper secondary education into two alternative-specific effects *vs.* no post-16 education and *vs.* academic education for students at the respective margins.

We evaluate our main model locally at the average proximity to either college type. As the effects at the mean appear to be close to average effects – at distances farther than the means they become more pronounced, whereas at distances closer than the mean they tend towards zero – the following discussion will loosely treat them as averages.

For students who are at the margin between vocational and no post-16 education, we find significant positive, and potentially very large, returns to enrolling in the vocational track for earnings at ages 29-30. We caution against taking the estimated effect of £17,000 at face value given the large uncertainty around the point estimate. Nevertheless, our results lend support to the argument that vocational education is highly beneficial for students who have grown disengaged with academic education and tend to drop out of school at the earliest point possible. This is in line with predictions from canonical human capital and signalling models on returns to education. A large part of the effect on earnings seems to be driven by extensive margin effects, i.e. by channelling students into employment. The effect is more pronounced among male students, who represent the majority of compliers at this margin.

For male students at the margin between vocational and academic education, we find a large negative effect of enrolling in the vocational track on earnings of about £3,000, or 15%, while for females, who represent roughly half of compliers at this margin, the effect is close to zero. The negative effect for males can be attributed to lower wages (and working hours) since, if anything, the effect of vocational education on employment is moderately positive.

To understand the mechanisms behind this result, we consider the effect of vocational track enrolment on a number of education outcomes. For starters, we confirm that marginal students who enrol in the vocational track study more vocational courses than if they had enrolled in the academic track. This difference is much more pronounced for male students, suggesting that vocational track attendance, indeed, leads these students to obtain a more occupation-specific skill set. Next, we consider upper secondary achievement: we find that the vast majority of male and female compliers achieve at least one Level 3 qualification (A-Levels and equivalent) regardless of the track they choose. If anything, for female students upper secondary achievement is slightly higher under vocational education. This suggests that, at least for the positively selected group of marginal students, there are no big differences in teaching effectiveness between vocational (i.e. Further Education) and academic (i.e. Sixth Form) colleges. Finally, we consider students' progression into higher education. Interestingly, we find that enrolling in the vocational track does not decrease the likelihood of completing a 3-year university degree: for male compliers the effect is zero, with about 37% completing a degree, whereas for female compliers, if anything, degree completion is less likely under academic education.

This result starkly contrasts with a widely-held belief that vocational education in England impairs educational progression after upper secondary education. Indeed, on average, only 13% of students in the vocational track complete a university degree (see Figure 3), highlighting that vocational-academic compliers are positively selected (as reflected in their above-average prior attainment and below-average levels of economic disadvantage shown in Figure 7). Given the inherent flexibilities within this system, it is not entirely surprising to see Further Education colleges offer successful pathways to university for more academically-inclined students. Note that, despite this neutral result in university completion, it is still possible that vocational enrolment affects the type of university degrees and institutions students enrol in. These are two factors expected to influence earnings that we will consider in future versions of this work.[19]

While progression into higher education is an important outcome, one of the main arguments for vocational education relates to students who fail to complement their academic upper secondary education with higher education and might therefore suffer from a lack of readily deployable occupation-specific skills. This argument would suggest positive returns to vocational education among the 62% of marginal students in our sample who do not pursue higher education. Taking the effect estimate £-3,145 for males at face value, a back-of-the envelope calculation shows that, for the effect among non-university-bound male students to be non-negative (as postulated), the effect for students who complete university would need to be £-8,500. Pending further investigation, we deem such a large negative effect among students who completed university after the vocational track unlikely. Hence, at least for male students, vocational track attendance

[19] It may also be possible that, despite being as likely to complete a university degree as academic students, the fact that vocational students acquire fewer academic skills puts them at a disadvantage during their studies. Among other things, this may be directly reflected in longer graduation times and lower final grades: two aspects that would be read as negative signals by employers.

seems to reduce earnings regardless of students' subsequent education choices. The fact that the negative effect for males materialises quite soon after leaving school and for females it is zero throughout students' twenties indicates that these results do not stem from the rapid obsolescence of occupation-specific skills (this may, however, worsen the gap as students get older).

This implies that canonical arguments in favour of vocational as compared to academic education, related to school-to-work transitions and corresponding earnings advantages in the early career, do not apply to England, despite having been empirically verified elsewhere (Zimmermann *et al.*, 2013). For example, a closely-related recent study by Silliman and Virtanen (2019) for Finland finds large positive earnings effects for marginal applicants who are admitted to the vocational instead of the academic track that persist at least until age 35. Our findings call into question how generalisable this result is beyond the group of Nordic and central European countries with well-established and highly-regarded vocational education systems. In countries like the UK and US, vocational education has historically been regarded with more suspicion, so that enrolment in the vocational track may be read as a negative signal by employers, which might partly explain the negative earnings effects we find. Additionally, the vocational system in England suffers from numerous well-known issues: vocational students face a flexible but confusingly wide choice of courses, only few of which have currency with employers (Musset and Field, 2013). Additionally, there is a lack strong career guidance and no well-signposted progression routes into higher vocational education (which is nearly absent anyway) and other work-based learning opportunities, such as apprenticeships (Wolf, 2011).[20]

In terms of policy relevance, one of the main advantages of our approach is that changes in enrolment probabilities induced by the distance instruments map directly into expansion (or contraction) policies of vocational (Further Education) colleges. We show that 17% of students are indeed responsive to distance to vocational college in their education choice and that, on net, the effect of vocational education for these students is positive. However, our margin-specific estimates reveals that this positive net effect is driven by only 20% of students who would otherwise pursue no post-16 education, whereas the remaining 80% of students are diverted from the academic track and see their earnings reduced as a result of enrolling in vocational college. Accordingly, we caution against across-the-board expansion of vocational education opportunities (in their current form). Instead, policy-makers should try to target students who are at risk of dropping out of education, since for these students gains are very large. Moreover, to offset the negative effects of vocational education for male students at the academic education margin the vocational curriculum could be better integrated with post-secondary and work-based education opportunities.

[20]This may not always have been the case: Brunello and Rocco (2017) look at the 1958 cohort of students in Britain and find that vocational students educated at the same level as academic students enjoyed a relatively large earnings premium at age 23 which then becomes negative by age 33. However, it is not clear how comparable these estimates of returns to vocational education can be: since then Britain has morphed into a post-industrial more service-oriented economy and countless changes have affected the education system.

References

Abadie, A. (2002). 'Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models', *Journal of the American Statistical Association*, vol. 97(457), pp. 284–292.

Abramson, I.S. (1982). 'On Bandwidth Variation in Kernel Estimates-A Square Root Law', *Annals of Statistics*, vol. 10(4), pp. 1217–1223.

Autor, D.H. (2019). 'Work of the Past, Work of the Future', *AEA Papers and Proceedings*, vol. 109, pp. 1–32.

Bertrand, M., Mogstad, M. and Mountjoy, J. (2019). 'Improving Educational Pathways to Social Mobility: Evidence from Norway's "Reform 94"', National Bureau of Economic Research Working Paper 25679.

Brunello, G. and Rocco, L. (2017). 'The Labor Market Effects of Academic and Vocational Education over the Life Cycle: Evidence Based on a British Cohort', *Journal of Human Capital*, vol. 11(1), pp. 106–166.

Carneiro, P., Heckman, J.J. and Vytlacil, E.J. (2011). 'Estimating Marginal Returns to Education', *American Economic Review*, vol. 101(6), pp. 2754–2781.

Carneiro, P. and Lee, S. (2009). 'Estimating distributions of potential outcomes using local instrumental variables with an application to changes in college enrollment and wage inequality', *Journal of Econometrics*, vol. 149(2), pp. 191–208.

Cavaglia, C., McNally, S. and Ventura, G. (2020). 'Do Apprenticeships Pay? Evidence for England', *Oxford Bulletin of Economics and Statistics*, vol. 82(5), pp. 1094–1134.

Dahl, G.B., Kostøl, A.R. and Mogstad, M. (2014). 'Family Welfare Cultures', *The Quarterly Journal of Economics*, vol. 129(4), pp. 1711–1752.

Goldin, C. (2001). 'The Human-Capital Century and American Leadership: Virtues of the Past', *The Journal of Economic History*, vol. 61(2), pp. 263–292.

Hall, C. (2016). 'Does more general education reduce the risk of future unemployment? Evidence from an expansion of vocational upper secondary education', *Economics of Education Review*, vol. 52, pp. 251–271.

Hanushek, E.A., Schwerdt, G., Woessmann, L. and Zhang, L. (2017). 'General Education, Vocational Education, and Labor-Market Outcomes over the Lifecycle', *Journal of Human Resources*, vol. 52(1), pp. 48–87.

Hastie, T. and Tibshirani, R. (1993). 'Varying-Coefficient Models', *Journal of the Royal Statistical Society. Series B*, vol. 55(4), pp. 757–796.

Heckman, J.J. and Pinto, R. (2018). 'Unordered Monotonicity', *Econometrica*, vol. 86(1), pp. 1–35.

Heckman, J.J. and Urzúa, S. (2010). 'Comparing IV with structural models: What simple IV can and cannot identify', *Journal of Econometrics*, vol. 156(1), pp. 27–37.

Hupkau, C., McNally, S., Ruiz-Valenzuela, J. and Ventura, G. (2017). 'Post-Compulsory Education in England: Choices and Implications', *National Institute Economic Review*, vol. 240(1), pp. R42–R57.

Hupkau, C. and Ventura, G. (2017). 'Further education in England: Learners and institutions', CVER Briefing Notes 001.

Imbens, G.W. and Rubin, D.B. (1997). 'Estimating Outcome Distributions for Compliers in Instrumental Variables Models', *The Review of Economic Studies*, vol. 64(4), pp. 555–574.

Kennedy, E.H., Lorch, S. and Small, D.S. (2019). 'Robust causal inference with continuous instruments using the local instrumental variable curve', *Journal of the Royal Statistical Society: Series B*, vol. 81(1), pp. 121–143.

Kirkeboen, L.J., Leuven, E. and Mogstad, M. (2016). 'Field of Study, Earnings, and Self-Selection', *The Quarterly Journal of Economics*, vol. 131(3), pp. 1057–1111.

Kline, P. and Walters, C.R. (2016). 'Evaluating Public Programs with Close Substitutes: The Case of Head Start', *The Quarterly Journal of Economics*, vol. 131(4), pp. 1795–1848.

Mountjoy, J. (2019). 'Community Colleges and Upward Mobility', Social Science Research Network, Rochester, NY.

Musset, P. and Field, S. (2013). *A Skills beyond School Review of England*, Paris: OECD Publishing.

OECD (2017). *Getting Skills Right: Skills for Jobs Indicators.*

Olea, J.L.M. and Pflueger, C. (2013). 'A Robust Test for Weak Instruments', *Journal of Business & Economic Statistics*, vol. 31(3), pp. 358–369.

Roy, A.D. (1951). 'Some thoughts on the distribution of earnings', *Oxford Economic Papers*, vol. 3(2), pp. 135–146.

Ryan, P. (2001). 'The School-to-Work Transition: A Cross-National Perspective', *Journal of Economic Literature*, vol. 39(1), pp. 34–92.

Shavit, Y., Müller, W. and Tame, C. (1998). *From School to Work: A Comparative Study of Educational Qualifications and Occupational Destinations*, Clarendon Press.

Silliman, M. and Virtanen, H. (2019). ‘Labor market returns to vocational secondary education’, ETLA Working Papers 65.

Spiess, C.K. and Wrohlich, K. (2010). ‘Does distance determine who attends a university in Germany?’, *Economics of Education Review*, vol. 29(3), pp. 470–479.

Wolf, A. (2011). ‘Review of vocational education: The Wolf report’, UK Department for Education and Department for Business Innovation and Skills.

Wright, T., Klein, M. and Wieczorek, J. (2019). ‘A Primer on Visualizations for Comparing Populations, Including the Issue of Overlapping Confidence Intervals’, *The American Statistician*, vol. 73(2), pp. 165–178.

Young, A. (2020). ‘Consistency Without Inference: Instrumental Variables in Practical Application’, unpublished manuscript.

Zimmermann, K.F., Biavaschi, C. and Eichhorst, W. (2013). *Youth Unemployment and Vocational Training*, Boston: Now Publishers Inc, illustrated edition edn.

A Additional Tables and Figures

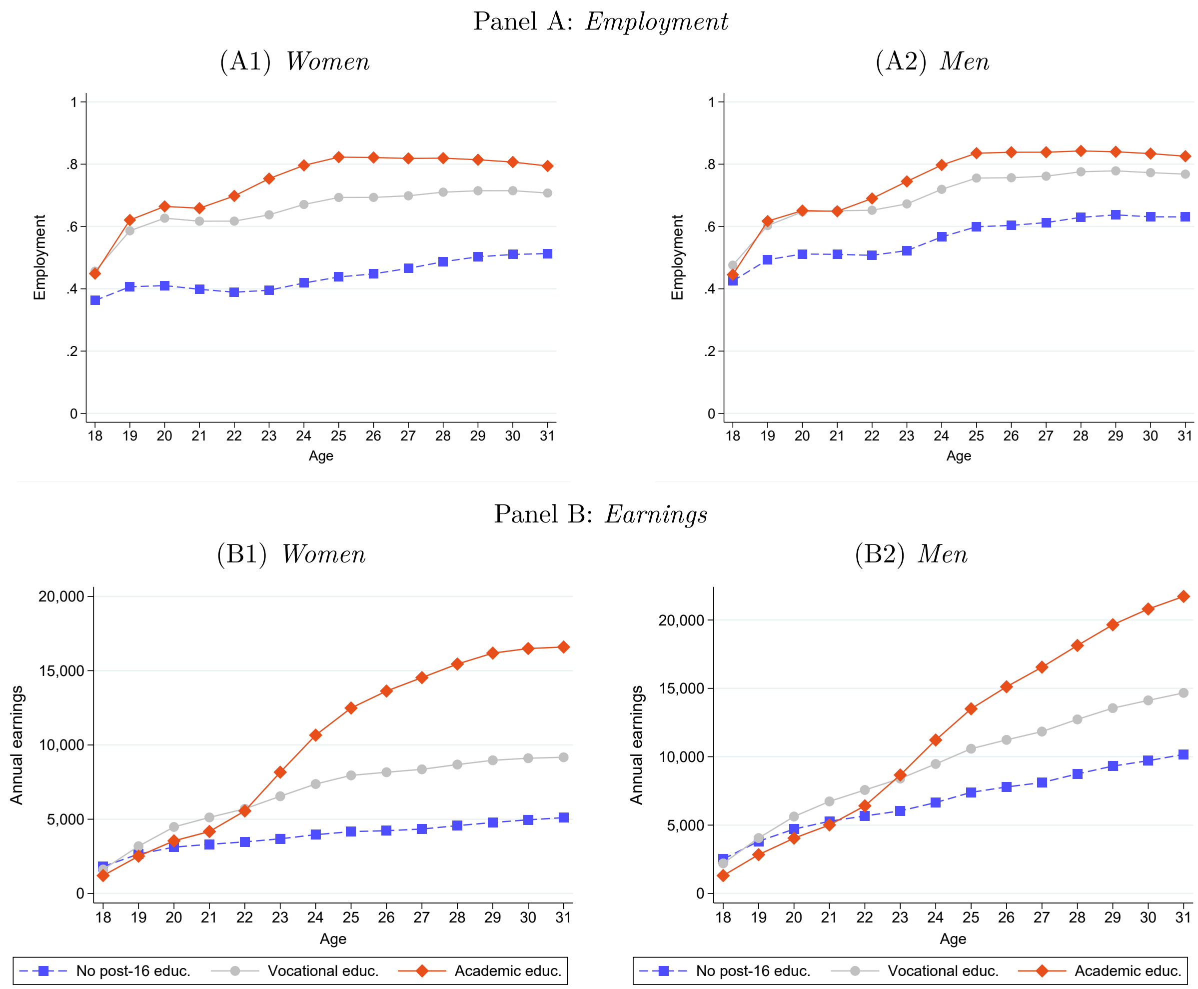


Figure A1. Labour market trajectories by initial enrolment for full sample.

Notes: This reproduces Figure 4 using all students.

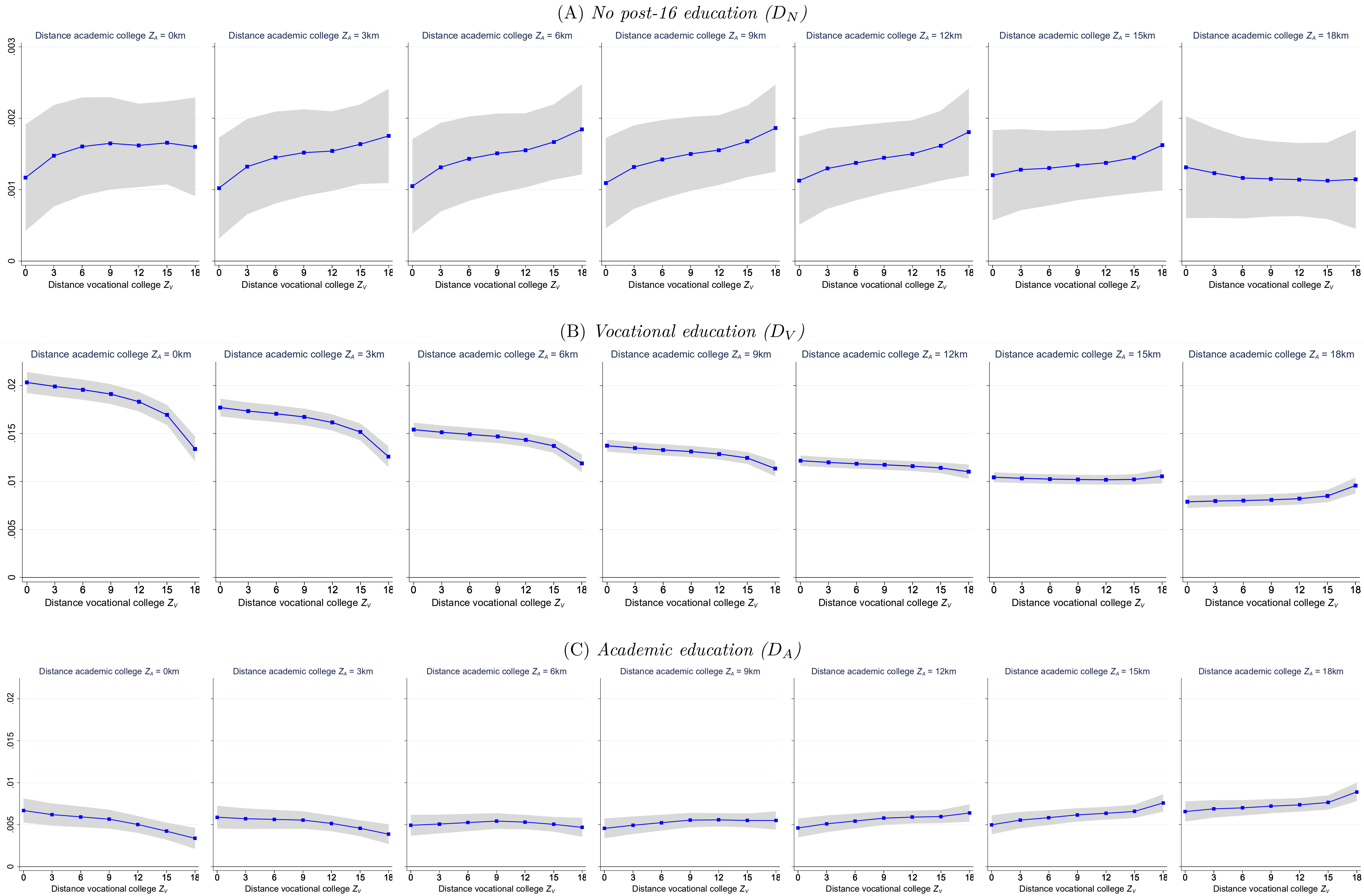
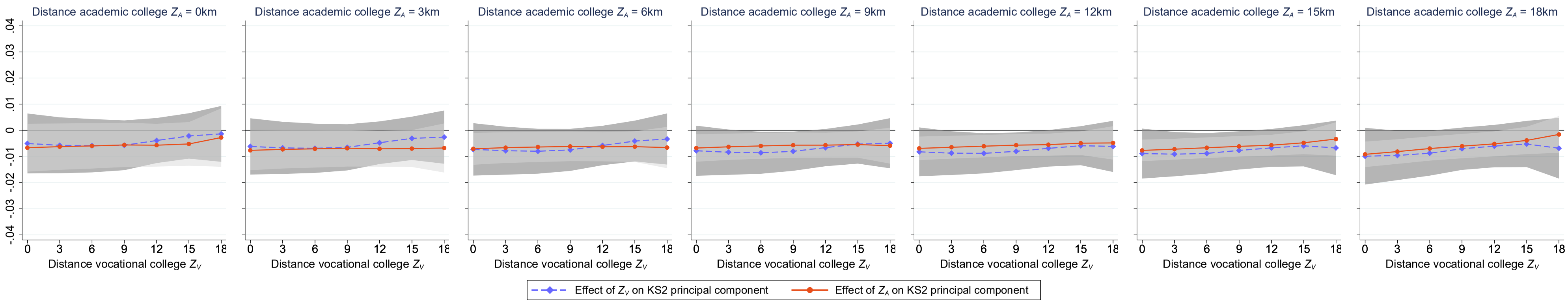


Figure A2. First stages across the distance grid.

Notes: This figure shows the three first stages across all points of the distance grid, estimated using locally weighted OLS regressions analogously to Table 3. At each grid point we weight observations by their proximity to the evaluation point using a two-dimensional Epanechnikov kernel with 15km bandwidth.

(A) *Instrument balance tests for KS2 test scores*



(B) *Comparability of compliers' KS2 test scores*

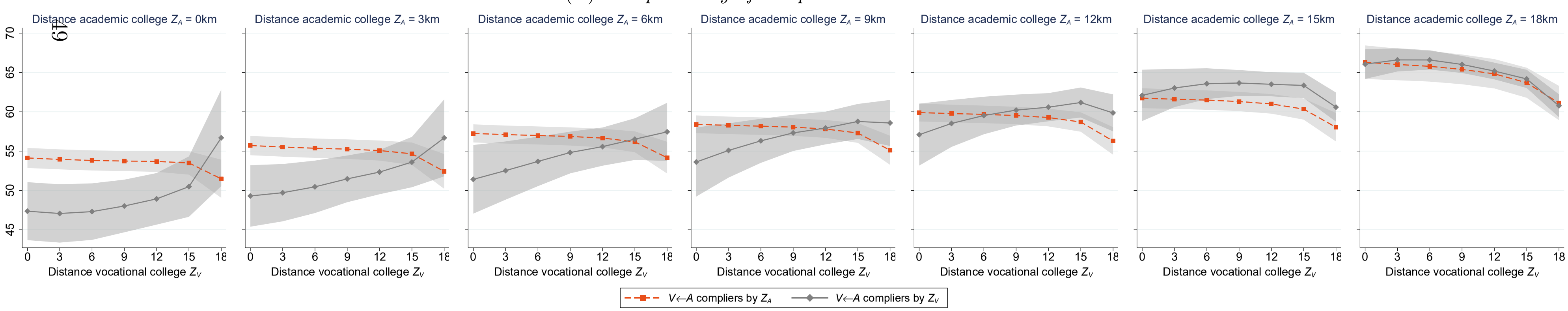
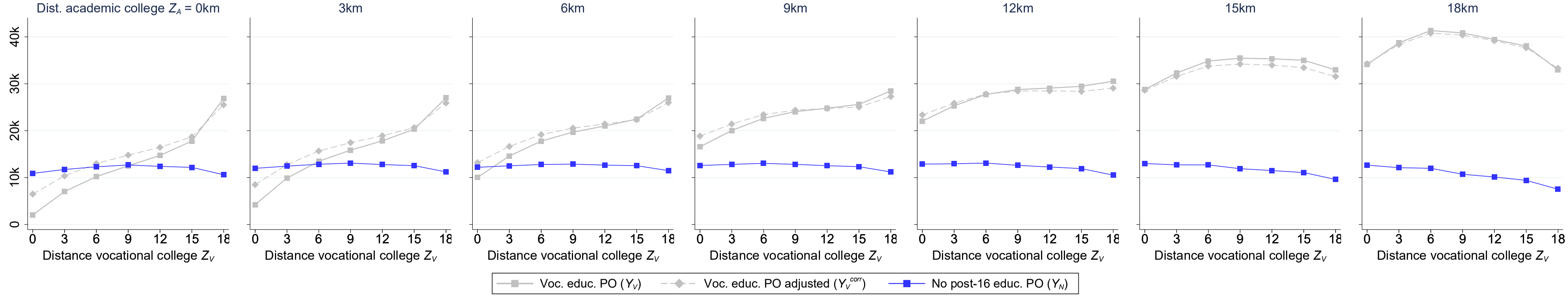


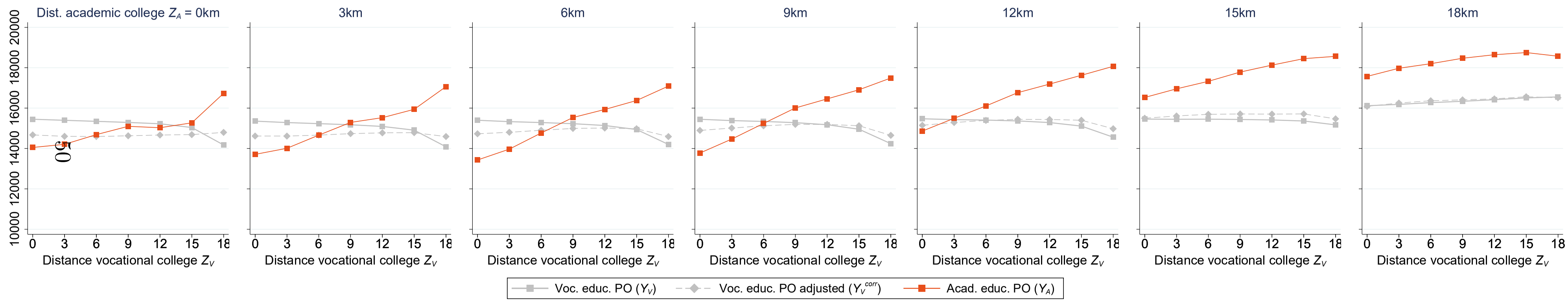
Figure A3. Evaluation of identification assumptions across the distance grid.

Notes: This figure assesses the plausibility of identification assumption A1, independence and exclusion, and A3, complier comparability, across all points of the distance grid. To test A1, panel A shows the results from instrument balance tests, i.e. from regressing the first principal component of the three KS2 test scores on the two distance instruments, Z_V and Z_A, and the control set, $\mathbf{X}$, excluding the KS2 test scores, using locally weighted regressions. We report the coefficients on Z_V and Z_A, including 95% confidence intervals, based on LSOA-level clustered standard errors. To test A3, panel B shows the estimated mean KS2 first principal component (measured in percentiles) for vocational-academic compliers induced by Z_A and vocational-academic compliers induced by Z_V (including 95% confidence intervals), estimated using locally weighted 2SLS regressions, analogously to in Figure 7. For both panels, at each grid point we weight observations by their proximity to the evaluation point using a two-dimensional Epanechnikov kernel with 15km bandwidth.

(A) *Potential outcome estimates for vocational ← no post-16 education compliers*



(B) *Potential outcome estimates for vocational ← academic education compliers*



(C) *Estimated share of vocational ← no post-16 education compliers*

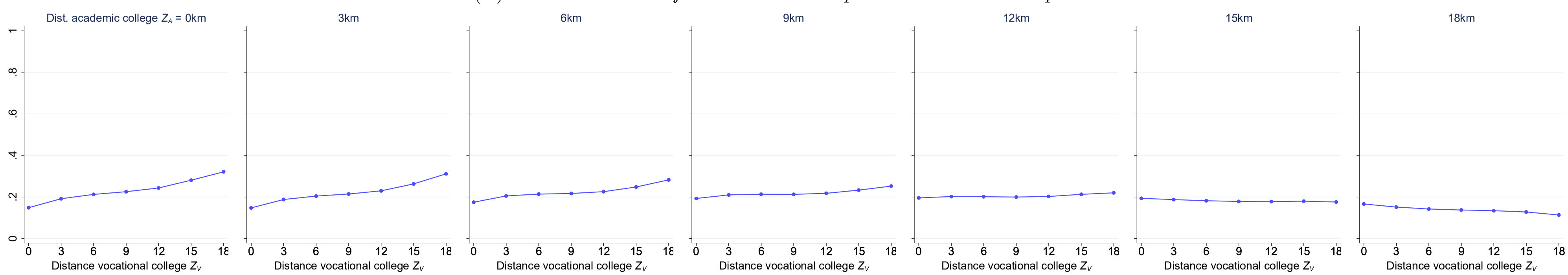


Figure A4. Potential earnings and complier share estimates across the distance grid.

Notes: