

# Do Management Practices Matter in Further Education?

*Sandra McNally, Luis Schmidt, Anna Valero*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: <https://www.cesifo.org/en/wp>

# Do Management Practices Matter in Further Education?

## Abstract

Further Education colleges are a key way in which 16-19 year olds acquire skills in the UK (much like US Community Colleges), especially those from low income backgrounds. Yet, little is known about what could improve performance in these institutions. We design and conduct the world's first management practices survey in these colleges (based on the World Management Survey) and match this to administrative longitudinal data on over 40,000 students. Value added regressions with rich controls suggest that structured management matters for educational outcomes (e.g. upper secondary qualifications), especially for students from low-income backgrounds. In a hypothetical scenario where a learner is moved from a college at the 10th percentile of management practices to the 90th, this would be associated with 8% higher probability of achieving a good high school qualification, which is nearly half of the educational gap between those from poor and non-poor backgrounds. Hence, improving management practices may be an important channel for reducing inequalities.

JEL-Codes: I200, J240.

Keywords: management practices, further education.

*Sandra McNally*  
*School of Economics*  
*University of Surrey / United Kingdom*  
*s.mcnally@surrey.ac.uk*

*Luis Schmidt*  
*STICERD, London School of Economics*  
*London / United Kingdom*  
*L.E.Schmidt-Rivera@lse.ac.uk*

*Anna Valero*  
*Centre for Economic Performance*  
*London School of Economics*  
*London / United Kingdom*  
*A.A.Sivropoulos-Valero@lse.ac.uk*

Acknowledgements: We would like to thank Daniela Scur, John Van Reenen, Tomaz Teodorovicz and the WMS team for their advice in the development and implementation of this survey in line with the WMS methodology, and John Van Reenen for his very helpful advice on this paper. We are grateful to Diego Correa and the Integral Research team for conducting the survey. We thank David Corke, Charlynne Pullen, Maggie Gallier, Eddie Playfair, Helen Wood, Kirsti Lord and Steve Frampton for their help developing the survey questions, and David Corke and Morag Gallagher for enabling us to recruit colleges from the AoC network. We thank Jenifer Ruiz-Valenzuela for helping us to test whether management practices are linked to measures of principal effectiveness that are reported in her work. We are grateful to participants at the Empirical Management Conference, the CVER seminar, the AoC Labour Market Observatory Series, the European Association of Labour Economics conference, the CESifo Economics of Education conference, the Business School, University of Stavanger, and the School of Business and Economics of Maastricht University for their useful feedback, especially Eric Taylor and to Chiara Cavaglia, Andy Eyles, Claudia Hupkau, and Guglielmo Ventura for their help with the various datasets. We thank Richard Garrett and Helen Wood at DfE for their support for this project and acknowledge funding from the Department for Education as part of the programme of research under the Centre for Vocational Education Research. The authors also gratefully acknowledge financial support from the CEP and the ESRC under grants ES/T007702/1 and ES/S001735/1. The research proposal was submitted to review by the Research Ethics Committee at LSE and was approved in January 2019.

## Introduction

In the UK, as in other countries, despite the increase in graduates, there are too many people with low levels of skill and too few with intermediate and higher-level technical skills.<sup>2</sup> Over half of every school leaving cohort enter Further Education (FE) colleges that have the provision of such skills as their core mission. A disproportionate number of students from poorer families enter these institutions and hence they also provide an important route for improving social mobility. Yet we know relatively little about what mechanisms improve efficacy in colleges at the post-secondary or tertiary level, in the UK or elsewhere.

We evaluate the role of management practices in FE colleges in England, contributing to an understanding of what influences educational outcomes at this level of education as well as to the broader debate on the role of management practices for improving performance in the public sector. This is the first study to evaluate management practices in colleges<sup>3</sup> – and although its findings are inevitably specific to the institutional context of the UK, it also has relevance to institutions with similar aims in other countries. Furthermore, we examine explicitly, and within FE colleges, whether better management practices help students from disadvantaged backgrounds. The school resources literature suggests that reductions in class size and increases in school expenditure tend to help those from low socio-economic background disproportionately (e.g. see review by Gibbons and McNally, 2013). We investigate whether this is the case here. As FE colleges disproportionately enrol those from such backgrounds (and therefore have scope to affect social mobility), this is a very pertinent issue.

The further education sector has been described as the ‘Cinderella sector’ in England, because it is so often overlooked in public debate about education.<sup>4</sup> This is related to its focus on vocational education and skills, which tend to be treated with less interest in the media than universities or schools; and it is reflected in a funding squeeze over recent years (Britton et al. 2019). Yet the FE sector is a vital part of the national and local infrastructure for education and skills, which has in turn an important part to play for increasing productivity growth and

---

<sup>2</sup> This is regularly discussed in policy reports such as those commissioned by the government - Wolf (2011) and the Augar Review (2019) - as well as reports on economic growth such as those by the LSE Growth Commission.

<sup>3</sup> While there are publications on various aspects of the operation of Further Education Colleges (see for example, Hodgson, 2015), to our knowledge there is no systematic evaluation of management practices that is comparable to our study.

<sup>4</sup> See, for example <https://www.ft.com/content/56729f80-e523-11df-8e0d-00144feabdc0>

improving social mobility (Besley and Van Reenen, 2013). FE and Sixth Form colleges<sup>5</sup> enrol about half of every cohort after they complete their compulsory full-time education at age 16. The share of those from disadvantaged families enrolling is about twice what it is in other educational settings.<sup>6</sup> FE colleges also cater for individuals at all levels and ages who wish to gain vocational qualifications and are thus important for adults who wish to train and reskill as well as for young people engaged post-compulsory education. They have similarities to two-year community colleges in the United States – with one difference being that they typically enrol young people at the beginning of their upper secondary education (at age 16) rather than at the end.

In this paper, we investigate whether management practices in FE colleges are an important correlate of performance, principally measured by student performance and progression but also using other institutional-level outcomes. We collect our own data using the methodology of the World Management Survey (WMS) (Bloom and Van Reenen, 2007) and link this survey to administrative data for outcomes and other important characteristics of these institutions and the people who attend them. The WMS was first applied to the manufacturing sector in a handful of countries and has now been carried out across 35 countries worldwide and in a variety of sectors including schools (Bloom et al., 2015a), universities (McCormack et al., 2014) and healthcare (Bloom et al., 2015b).<sup>7</sup> Across these different settings, management practices have been found to be a key driver of performance. As discussed by Hwa and Leaver (2021), the WMS is increasingly seen as the gold standard for standardized comparisons of school management, with the instrument (and associated data) being used dozens of times since its inception in 2009. However, this is the first time the WMS has been conducted and analysed in the context of Further Education colleges (or, to the best of our knowledge, any similar institutions in other countries).

In general, much less is known about what influences performance within FE institutions compared to schools and universities. To some extent this is driven by the paucity of data available for this sector (Augar Review, 2019), at least up until recently. Aucejo et al. (2020) investigate the value added of FE colleges. They find that although the institution attended does

---

<sup>5</sup> Our study covers both FE colleges and Sixth Form colleges, which we often refer to collectively as “FE colleges” unless there is a need to distinguish between them in the analysis or discussion. Sixth Form colleges cater for 16-18 education only and tend to have more focus on academic post-16 options (A-levels).

<sup>6</sup> The typical alternative for students (aged 16-18) not enrolling in FE and Sixth Form Colleges is to stay in the same school they were in for their 11-16 secondary education (in ‘sixth form’). This is only an option for individuals who are attending a 11-18 school (not available everywhere) and if they want to stay on an academic track having done sufficiently well in the national exams at age 16 (GCSE).

<sup>7</sup> For an overview of this body of research over the past 18 years, see Scur et al., (2021) and <https://worldmanagementsurvey.org/>.

influence progression and earnings, overall variation is relatively modest compared to returns to field of study. Nonetheless, the authors find that a one standard deviation increase in college value-added leads to an increase in daily earnings after college attendance of around 3% for young learners. Ruiz-Valenzuela et al. (2017) is among the few studies to investigate what influences performance within FE colleges. The authors investigate the role of principals and find that while particular principals appear to affect educational outcomes of the students in their colleges,<sup>8</sup> this cannot be explained by observable characteristics of the principals themselves. We attempt to get inside the ‘black box’ of what happens in these institutions by collecting detailed and standardised data on organisational processes in operations, targeting, monitoring and people management, together with institutional details and information on college leadership which we then relate to the performance of these institutions. In line with the broader literature using WMS, we develop an international management index for further education institutions in England. We use double-blind telephone interviews with principals to collect information on management practices for 79 colleges (25% of the survey population) in 2018-2019.

Given that the focus of this study is on 16-19 education, our WMS survey has been closely modelled on that used for schools (Bloom et al., 2015a), and has been adapted for the FE sector with the help of experts and practitioners. Bloom et al. (2015a) find a strong correlation between management practices and school performance, which is largely driven by the strength of governance (i.e. strong accountability to an outside body) and the degree of school leadership (i.e. developing a long-term strategy for the school). Their findings are supported by studies of school practices in the US such as Dobbie and Fryer (2013) and Angrist et al. (2013). More recently, Fryer (2017) has conducted a randomised control trial of management practices in Texas schools. Specifically, he investigates the effect of principal management training on school productivity, where this consists of 300 hours of training on lesson planning, data-driven instruction, and teacher observation and coaching. He finds effects that are driven by principals who implement the training well and who do not subsequently leave the school.

Although our setting has some similarities with schools, they are not the same and the outcome measures are very different. FE colleges are much larger and have a broader mission that encompasses 16-19 education (vocational and academic), large-scale provision for unemployed people and English for speakers of other languages and adult training at all levels.

---

<sup>8</sup> In contrast, Janke et al. (2019) find little evidence of top manager impacts on the performance of hospitals.

Like schools, there is a rigorous accountability framework in place, but they have had additional challenges in recent years with more severe funding cuts and many policy changes affecting aspects of their work. The outcomes of relevance are whether individuals are able to achieve qualifications at various levels and whether they are able to progress to university. Unlike for schools, there is no test score relevant in this context (and hence it is difficult to compare our results to Bloom et al. (2015) for schools). The outcomes of interest here are important for success in the labour market and we can contextualise the magnitude by comparing “effects” to the value added of Further Education Colleges on the same outcome measures (Aucejo et al. 2020).

We contribute to the literature in the following ways. First, we investigate to what extent management practices influence institutional and individual outcomes in the FE sector in England. The importance of this sector is increasingly recognised for its role in improving the skills base in England and for social mobility (Augar Review 2019). But we know much less on what influences performance in this sector compared to schools or universities. Second, because we can link the survey to individual level (administrative) data on educational histories, we can estimate value added regressions (which is unusual in this literature), and explore heterogeneity across different types of learner. Although our estimates are still correlational, the fact that we can control for many obvious confounders in our individual-level analysis brings us closer to a causal estimate of the effect of management practices on an individual’s future achievement and progression. Third, we investigate whether management practices are disproportionately important for students from disadvantaged backgrounds. We can investigate this within colleges in fixed effect regressions (thus controlling for any systematic differences between colleges that could be correlated with overall management practices). Thus, a causal interpretation can be given to the estimate under weaker identification assumptions than for the average effect of management practices.

Our first key finding is that structured management practices do appear to matter for educational achievement and progression to university education at the individual level, and these results are robust to alternative specifications. Our main result is that a 1 standard deviation increase in the management score is associated with around a 2 percentage point increase in the likelihood of achieving “Level 3” (upper secondary) qualification, or pursuing a degree by age 20. Our second main finding is that good management practices are more important for achieving secondary qualifications for students from low-income backgrounds, controlling for prior ability. This is robust to including institutional fixed effects and also to including interactions between the socio-economic background of students and other



institutional characteristics. In a hypothetical scenario where a learner is moved from a college at the 10<sup>th</sup> percentile of management practices to the 90<sup>th</sup>, this would be associated with 8% higher probability of achieving a good high school qualification (at level 3), which is nearly half of the educational gap between those from poor and non-poor backgrounds. The labour market returns to level 3 qualifications have been estimated to be at least 6% (Machin et al., 2018), suggesting that improving college management practices has the potential to help reduce inequality and improve social mobility.<sup>9</sup>

Good management practices also have the potential to enable students to progress to higher education. We show that this effect is driven by students who enter FE colleges with good results in national exams at age 16. Furthermore, this effect is driven by institutions that where higher education is a main area of college provision. This result shows that well managed FE colleges have potential to be engines of social mobility at this higher level, at least for those students who are already well-prepared when they enter.

Having established that management practices appear to matter in the further education sector, we go on to explore whether there are any particular principal or college characteristics that tend to be associated with better management practices. We find that spatial measures of competition (from other colleges) appear to matter, but other features – including observable characteristics of college principals - do not explain differences in management practices. This gives confidence that the management practices score is not simply reflecting other attributes of colleges.

We explore further the link between management scores and measures of effective college leadership and find that there is some evidence that these two dimensions are related, but that management practices do not appear to simply reflect more effective leadership. In particular, we merge our survey with measures of principals’ “ability” (Ruiz-Valenzuela et al., 2017), and find that these do not generally explain differences in management practices though it is difficult to draw firm conclusions given the small sample size in this piece of analysis. We also explore the extent to which measures of effective principal leadership and accountability (based on additional questions in our survey) relate to management practices and find that these are highly correlated. Both are positively associated with educational outcomes, but overall there appears to be a stronger relationship between management practices and progression. We

---

<sup>9</sup> We extend the analysis in order to explore directly whether management practices matter for the early labour market outcomes of learners (at age 20). We find little evidence of such a relationship in the short timeframe available in the data. The lack of effect could also be because management practices are influencing the probability of entering higher education (inducing negative selection into the labour market at age 20).

interpret these findings as evidence that management practices are capturing something distinct from effective leadership. While we expect the leaders of organisations to influence management practices, such practices reflect processes on the ground which can be thought of as a type of technology (Bloom et al., 2016), evolving slowly as particular leaders come and go.<sup>10</sup>

The remainder of this paper is structured as follows. We describe the institutional setting and data (Section 1). We then report on how management practices relate to educational outcomes within a regression framework (Section 2). We evaluate the implications for labour market outcomes (Section 3). We then explore how management practices vary across FE colleges and how this correlates with observable characteristics (Section 4). We then bring our findings together and discuss implications for policy in the conclusion (Section 5).

## **1. Institutional setting and data**

### ***1.1 England's Further Education Sector***

FE colleges are major providers of education for adults and for young people in post-compulsory education in England. The latter have become an increasing part of their budget over time (around 40%) and tends to be the most reliable as government funding for adult skills has fallen over many years. Given the broad mission of FE colleges, we focus our study on management practices around areas that are mostly relevant to their provision for young people (16-19 year olds). We also include Sixth Form colleges in our study. These are much smaller institutions and cater exclusively for young people. They have a stronger focus on academic education (A-levels). FE and Sixth Form colleges have a high degree of autonomy from the government, although they are subject to the same regulatory regime as schools for the provision of education for young people. Thus, like schools, they are subject to visits from the national inspectorate (Ofsted) and are scored according to the same criteria. These reports are in the public domain and a key part of the accountability infrastructure.

About half of each cohort attend FE and Sixth Form colleges in the year following national exams at the end of compulsory schooling (GCSEs). Disadvantaged groups are overrepresented in this sector: 15% of students in FE and Sixth Form colleges were eligible to receive Free School Meals (FSM) in their final year of lower secondary education (i.e. age

---

<sup>10</sup> This appears to be the case in firms. Bandiera et al. (2020) develop an index of CEO behaviour which is related to firm productivity, and find this to be correlated with WMS scores. However, the authors also find that these two variables are independently correlated with firm productivity suggesting that they do indeed capture distinct drivers of performance. For further discussion on the link between managerial human capital and management practices, see Valero (2021).

16),<sup>11</sup> which is double the share of those staying in schools for their upper secondary education. Typically, those attending FE colleges will be doing a vocational qualification and some may choose to combine this with A-levels. Since 2015, all new entrants need to repeat GCSE English and/or maths if they failed to get a good grade in the national exams at age 16. The other half of the cohort remain in their secondary school for their further education. Although it is – in theory – compulsory to stay in some form of education or training until the age of 18, there is some drop-out before that time. Nationally, about 4 per cent of each cohort are classified as ‘not in education training or employment’ at age 18.

Whereas academic qualifications in further education (A-levels) are well understood, these are pursued only by about 40 per cent of students as their main qualification. This is the main route to university studies. Other qualifications are broadly classified as ‘vocational’ although they vary in their pre-requisites, length, duration, and field of study. Post-16 vocational education has frequently been criticised as having too many options to choose from and too few progression pathways to higher levels of education and skills (see, for example, Hupkau et al. 2017). Many attempts have been made at reform in the further education sector and this is ongoing, occurring against a backdrop of funding cuts and high levels of policy churn.

This is the environment and policy context in which we conduct our survey of management practices, the first of its type in this sector. The complex and pressurised environment makes this a very interesting sector to explore, apart from its obvious importance in the national educational infrastructure.

### ***1.2 Measuring management practices in the FE sector***

In order to measure management practices in FE colleges, we adapted the World Management Survey (WMS) methodology, first described in Bloom and Van Reenen (2007). The WMS consists of a series of open-ended questions on organisational processes which need to be scored between 1 and 5, where a high score indicates that a college has adopted structured management practices.

The core survey as originally developed for the manufacturing sector consists of 18 questions of management under the following groupings: operations, management, targets and incentives. This has been previously adapted for schools (Bloom et al., 2015a) through tailoring

---

<sup>11</sup> Eligibility for Free School Meals is a standard measure of socio-economic disadvantage and is based on whether the family is eligible for various types of income support. Eligibility for free school meals only applies to students up to the end of their lower secondary education at age 16.

and adding to the operations questions. Here we have taken the 20 school questions as our starting point, and adjusted these so that they are appropriate for FE and Sixth Form colleges in England.

This process involved changing the language in some of the questions so that they are relevant for the sector (e.g. using “learner” rather than “student” which is the norm in this sector) and adding one question in the operations section to measure effectiveness of matching learners to courses which is a particular issue in FE colleges where students select their courses. In order to inform the survey design, we discussed the questions and piloted the survey with representatives from the Association of Colleges and other sector stakeholders including former college principals. We also discussed our questions with members of the WMS core team to ensure consistency with previous surveys. Our main measure of management practices, the management score, is the average of the set of 21 questions, but we also analyse the different groupings separately. Our core survey questions are grouped into four categories as follows (the full set of questions is set out in Appendix Table A1.1):

### **Operations**

*Matching learners to courses:* Learners are recruited and retained in well-matched qualifications/courses.

*Standardisation of teaching processes:* Materials and practices are standardised and aligned in order to be capable of moving learners through learning pathways over time, and ensuring courses are meeting the needs to individuals and employers.

*Personalisation of instruction and learning:* Flexibility in teaching methods and learner involvement ensuring all individuals can master the learning objectives.

*Data-driven planning and learner transitions:* College uses assessment to verify learning outcomes at critical stages, makes data easily available, and uses it intelligently to adapt learner strategies and course offerings accordingly.

*Adopting educational best practices:* College incorporates teaching best practices and the sharing of these resources into the classroom.

### **Monitoring**

*Continuous improvement:* College implements process documentation and continuous improvement.

*Performance tracking:* College performance is measured with appropriate methods and frequency, and communicated effectively with staff, governors and other stakeholders.

*Performance review:* Performance is reviewed with appropriate frequency and follow-up.

*Performance dialogue:* The quality of performance review conversations.

*Consequence management:* The extent to which differing levels of college performance (not only individual teacher performance) lead to different consequences.

### **Target Setting**

*Target balance:* System tracks meaningful targets tied to learner outcomes, in particular, the extent to which colleges set their own internal targets and use these in a sensible way.

*Target interconnection:* College and individual targets are aligned with each other and the overall system goals.

*Time horizon of targets:* College has a rational approach to planning and setting the targets. Target stretch: Targets are appropriately difficult to achieve.

*Clarity and comparability of targets:* Performance measures are understandable and performance is openly communicated.

### **People / Talent Management**

*Rewarding high performers:* Good teacher performance is rewarded proportionately.

*Fixing poor performers:* College is able to deal with underperformers.

*Promoting high performers:* Promotions and career progression are based on performance.

*Managing talent:* College identifies and targets needed teaching, leadership and other capacity in the college.

*Retaining talent:* College will go out of its way to keep its top talent.

*Attracting Talent/Creating a distinctive employee value proposition:* College has a clear employee value proposition.

We targeted principals or vice principals as the interviewee with college-wide knowledge of processes and systems in place, and this was consistent with the schools survey (Bloom et al., 2015a) where head teachers or principals were interviewed. Also, consistent with the schools survey, we included an extra section of questions on principal leadership which were scored. This included three questions on leadership vision and strategy, clearly defined accountability and clearly defined leadership and teacher roles. We also added some specific questions relevant to FE that have not been asked in other sectors (e.g. extent of engagement with HEIs and employers) and background characteristics of principals and colleges to explore as determinants of potential differences in management practices.

The interviews were focused on management practices governing 16-19 provision (as processes may well vary according to different areas of college activity). However, we also asked about whether processes are centralised across the college, or whether there is flexibility across subject areas or learning.

### *Conducting the Survey*

Once we designed the survey instrument, we passed it to a survey firm who have conducted the WMS previously and are well trained in its methods. The interviewers received rigorous training, including from external experts in the WMS methodology, and members of the research team also met with the interviewers in order to explain the institutional context and motivation of the study.

We attempted to reach the full population of FE colleges in England that have at least one learner aged 16 to 19 years old (310 colleges), with a number of emails and endorsements in sector newsletters from the UK Government's Department for Education and Association of

Colleges. We achieved a sample of 79 colleges over the period February to September 2019 (while 83 interviews were conducted, there were three instances where colleges had two individuals keen to take part separately, and we had to drop one college from our sample as we were unable to obtain outcomes data for it). Overall, this represents a response rate of 25%,<sup>12</sup> which is relatively high relative to comparable management surveys in the education sector.<sup>13</sup>

Once a principal agreed to take part, they were emailed some information on the study, and a short online survey which included consent and basic information questions, to save time in the subsequent management survey which would be conducted on the phone. At the end of the online survey, participants were able to schedule a time for the interviewers to call them and conduct the full survey.

As in previous WMS surveys we followed several steps to obtain a high-quality response. First, a “double blind” methodology is used to reduce biases. Interviewers are not given any metrics on the institution’s performance in advance of the interview. These are matched in from independent sources after the interviews are finished. In addition, as in similar surveys in other sectors, interviews are conducted without informing the principals that their answers would be evaluated against a scoring grid. This allows the gathering of information on actual management practices as opposed to the principal’s aspirations of what should happen.

Second, the interviewers were trained by individuals who had conducted the WMS before in other contexts and a number of calibration exercises are conducted to ensure consistency of scoring. All interviews are reviewed and approved by the survey manager, and around 80% of interviews are double-scored: by the interviewer conducting the interview and their manager listening in.

Third, we collected data on the interview process itself (such as the time of day and the day of the week), characteristics of the interviewee and the identity of the interviewer. These variables can be used as “noise controls” in the regression analysis to improve the precision of our estimates by reducing some of the measurement error.

### ***1.3 Measures of college performance***

#### *Institution level outcomes*

We obtained a number of measures of performance at the institution level, either ready-made (from the National Student Survey, the FE Choices learner/employer satisfaction surveys

---

<sup>12</sup> More specifically, we interviewed representatives of 65 Further Education Colleges (N=216; response rate 30%) and 14 Sixth Form Colleges (N=93; response rate 15%).

<sup>13</sup> For example, in schools, Bloom et al. (2015a) achieved a response rate of 8% in England, and around 20% in the US and Canada.

and the national Inspectorate, Ofsted), or constructed based on data on learners from administrative data in the Individualised Learner Record (ILR). Measures from these data sources are used to construct publicly available information on the performance metrics of FE institutions. Appendix A.2.1 describes in detail the data sources and cleaning steps for the institutional data, but key information is summarised here.

In the National Student Survey (NSS), we use the overall satisfaction measure that represents the share of learners that are satisfied with their course. Learner and employer satisfaction from the FE Choices survey are measured as the percentage of learners that would recommend the college to their friends, and the percentage of employers that would recommend this training provider to another employer, respectively. For Ofsted scores, we use the most recent data available from Ofsted inspections and reverse the raw Ofsted Effectiveness scores (1-4, where 1 is outstanding and 4 is inadequate) in our analysis so that higher values reflect higher scores.

Administrative data from the Individual Learner Record (ILR) is used to construct institutional measures of achievement that are appropriate for young people in further education (i.e. age 16-19). These are the proportion of learners who achieved at least one aim at Level 2 (equivalent to GCSE grade A\*-C) or Level 3 (A-Level or vocational equivalent) respectively. These outcomes are important both for educational progression and in the labour market (as shown by Machin et al. 2020; McIntosh, 2006 and Patrignani et al. 2017).

#### *Individual level outcomes*

Since the main focus of our management survey is on 16-19 provision at FE colleges, our individual level analysis centres on the outcomes of young learners. We construct individual level outcomes using administrative records from the National Pupil Dataset (NPD) linked to the Individualized Learner Record (ILR, 2013-2016) and Higher Education Statistics Agency (HESA, 2015-2016). In combination, these data permit us to track individuals from the last year of secondary school through further education and higher education. We are able to observe the entire population of individuals that enter the colleges, whether or not they complete their course. We use the most recent individual-level data available to us and enable us to observe medium-term outcomes for learners who attended these colleges (up to age 20).

We focus on the population of learners from the 2012 GCSE cohort (age 16 at that point) that entered further education in 2013 (age 17). Using the sources described above, we construct three key medium-term educational outcomes: these are indicators for whether, by the age of 20, an individual has achieved Level 2 or higher, Level 3 or higher, or is enrolled in

higher education, respectively. We then merge this dataset to our WMS survey based on the institution that each learner attended in 2013 (age 17).

Due to lags between learners being at specific colleges and the measurement of their educational outcomes, and the fact that our management survey was conducted in 2019, for the purpose of the individual-level analysis, we focus on the subsample of colleges where the principal has been working at the college for at least 6 years. This implies that these principals were at their respective colleges in 2013 (the year the 2012 GCSE cohort entered further education). Assuming that there is some persistence in the management practice scores (consistent with a view of management as an organisational technology, Bloom et al., 2016), this allows us to attribute at least some influence on the management scores collected in the WMS survey (in 2019), to the current principal who was also working at the college six years previously, and therefore consider that the management practices we measure in 2019 are likely to be correlated with the quality of management practices when the students whose outcomes we measure were at the college.<sup>14</sup> Our final sample for the individual analysis consists of the 45 colleges where principals have been at the college for over six years, and 40,000-45,000 learners (depending on the outcome variable), and our results are robust to alternative cut-offs.

As an extension to our main analysis, we also explore whether college management practices are related to early labour market outcomes. As in our main regressions, we use the sample of young learners that entered FE and Sixth Form colleges in academic year 2013 (age 17), and focus on their labour outcomes (employment and wages) in year 2017 (age 20).

#### ***1.4 Other data***

##### *Individual level controls*

The individual level data includes demographics (gender and ethnicity), family background information (Free School Meal eligibility and English as first language) and previous attainment in national exams during compulsory schooling (KS4 GCSE points). For more detail on the individual level data, see Appendix A.2.2.

##### *Regional characteristics*

We geocoded the colleges in our sample, together with the wider population of colleges and schools offering sixth form provision in order to obtain some geographic characteristics

---

<sup>14</sup> We note that because of these lags in the data, it is likely that our measures of management quality might be subject to more measurement error than if we had conducted the WMS in 2013. Over the six-year period, even colleges with the same principals might have improved or worsened in their practices due to external or internal factors. Therefore, to the extent we are able to find relationships between management practices and individual learner outcomes, these could be attenuated.



based on regional economic indicators or spatial measures of competition. Figure 1 plots the colleges that we interviewed on a map, together with the wider population of colleges and shows that our sample is evenly spread across England.

In order to test whether management quality in colleges is influenced by the extent of competition they face (a result that has been established in other sectors), we included a question on the extent of competition faced by the college in the survey. For a more objective measure, we also calculated alternative competition variables based on the location of colleges / sixth form colleges or schools with post-16 provision. In line with recently published work in a related sector, University Technical Colleges (Machin et al., 2020), we calculated two measures of spatial competition: (i) the number of other colleges within a radius of 20km, and (ii) the number of other colleges and schools with post-16 provision. In our analysis, we constructed a “high competition” dummy equal to 1 where the number of competitors is greater than the 75<sup>th</sup> percentile of the sample on each of the competition measures in turn (self-reported, colleges or colleges and schools).

We also obtained geographic characteristics of the regions where our surveyed colleges are located: regional GDP per capita and population density in 2017, from Eurostat.

#### *Measures of the “effectiveness” of principals*

Here we draw upon work in Ruiz-Valenzuela et al. (2017), where the effectiveness of principals is estimated using panel data on learners in colleges over time, to recover principal and college “fixed effects” in institution-level learner outcomes (at level 2 and level 3). More detail on the methodology is set out in the Data Appendix A.2.3. We gathered the most recent estimate of principal’s fixed effect for each of the colleges, and were able to merge such information for 61 out of the 79 colleges in our WMS sample. To capture a measure of relative effectiveness and abstract from the units of measurement in Ruiz-Valenzuela et al. (2017), we standardise the principal fixed effects.

### ***1.5 Descriptive Statistics***

Table 1 describes the key variables used in the analysis and begins with the management practice scores (which are scaled from 1-5). The average college has a relatively high management score, of 4.24, though there is substantial variation as shown in Figure 2, which plots the distribution of management scores across our sample. Within the overall score, FE colleges do particularly well in terms of the operations practices, which has a higher mean and smaller standard deviation than the other management practice groupings. And they appear to do worse in terms of people management – where there is a lower mean and a thicker lower

tail. This might be explained by the fact that colleges are restricted in the extent they are able to use pay to incentivise teachers.<sup>15</sup> And this finding is consistent with the schools sector, where people management practices tended to be worse than non-people management practices, in the UK and internationally (Bloom et al., 2015). Appendix Figure B1 plots the distribution of management practice scores across the separate groupings illustrating these patterns.

We compare the college sample to the distribution of management scores in UK secondary schools (sourced from Bloom et al., 2015a) and university departments (sourced from McCormack et al., 2014) in Figure 3. This shows that on average and based on these samples, FE colleges appear to be better managed than both schools and university departments, but the difference is larger in the case of schools. We explore the extent to which the FE college sample is selected (based on observables) below.

At the end of the interviews, the principals were asked to rate the management practices at their college. Overall, we find a positive correlation between these self-assessed scores and the WMS scores (as shown in Appendix Figure B2) which suggests that there is an alignment between what principals consider to be good practice and the scoring methodology in the WMS. These data also show that a large number of principals underestimate their relative management practice scores: a high share (around half) of those who rated themselves lower than the average actually had higher than average WMS management scores.<sup>16</sup>

Next, we describe the key outcomes for which we explore the relationship with management practices at the institution level. On average across our sample, colleges tend to perform relatively well. 78% of learners in colleges say that they are satisfied in the National Student Survey, and 80% and 82% of learners and employers are satisfied with these colleges according to the FE Choices surveys. On average, the Ofsted effectiveness score is 2.9 (representing a “good” rating here as the scores are reversed), and 79% and 80% of learners achieved at least one aim at levels 2 and 3 respectively. Regarding the colleges themselves, on average there are nearly 9,000 learners in these colleges across all types of provision, and 3,600 are aged 16-19 (which is the focus of our survey). 18% of colleges in the sample are sixth-form colleges.

Table 1 also describes the characteristics of the college principals. 46% of the principals are women. Regarding their training and background, 33% have a FE management specific

---

<sup>15</sup> In fact full-time FE teaching professionals in the UK earn, on average, around £2,500 less than secondary school teachers. ONS (2018).

<sup>16</sup> Conversely, a lower share tended to overestimate their relative positive. A quarter of those that considered their college to be above average turned out to have lower than average WMS scores.

qualification, 72% have a teaching specific qualification and 51% has some experience in Industry. 57% of principals have been working at their college for over 6 years.

We report our favoured spatial measure of competition here: 25% of the sample face a high density of colleges in their surrounding area (this is by construction, as this variable is an indicator which equals one for colleges that are in the top quartile in terms of the number of other colleges within a 20km radius). Figure B3 in the appendix plots the spatial competition measures on a map – these spatial measures are positively correlated with self-reported measures. Finally, the average college is in a NUTS1 region where population density was 916 people per square kilometre (8 of the colleges in our sample are in London which drives up this average), and where GDP per capita was nearly 31,000 in 2017.

As discussed, in our individual level analysis, we focus on a subset of colleges where the principal has been at the college for at least 6 years. Assuming that there is some persistence in management practices (which we measure in 2019), this allows us to correlate the management practices of the college they attend at age 17 with the outcomes of individuals we observe by age 20 in the ILR data, for more detail see the Data Appendix A.2.2. The descriptive statistics for this subsample are in Table B1, and we summarise the key outcomes, prior achievement and demographic variables at the individual level in Table B2.

The sub-sample of colleges where we can investigate individual-level outcomes (because the principal has been at the college for at least 6 years) has almost identical outcome measures to the full sample (See Table 1 and Appendix Table B1). The overall management score is also very similar (with an average of 4.28 versus 4.24 in the full sample).

### *Balance Checks*

We first compare our sample of colleges interviewed and the population of FE colleges that have at least one learner aged in the 16-19 age group (see panel A of Appendix Table B3). There is little evidence that our sample of colleges performs better than the average college, though they do seem to do slightly better in terms of Ofsted scores (this relationship is significant at the 10 per cent level). Our sample of colleges does seem to be positively selected in terms of size – colleges in the sample are 42% larger, have 37% more aims, and 38% higher expenditure in absolute terms. Normalising aims and expenditure by the number of learners however shows that on a per learner basis, there is no evidence of selection. It appears to be the case that larger colleges participated in our survey, but they do not appear to be better

resourced in per learner terms.<sup>17</sup> Given the positive selection in terms of size, we ensure that we control for college size in regressions that link college outcomes to management practices.

## 2. Management practices and educational outcomes

Are management scores associated with educational outcomes, as has been found to be the case in other educational contexts? In the next sections we first estimate basic institution level regressions, and then more disaggregated analysis at the individual level where we are able to control for many obvious confounders.

### 2.1 Management practices and educational outcomes at the institution level

We begin by estimating a version of the education production function as follows:

$$Y_{jk} = \alpha_0 + \alpha_1 M_j + \alpha_2' Z_j + \alpha_3' R_k + u_{jk} \quad (1)$$

where  $Y_{jk}$  are the outcomes for college  $j$  in region  $k$ ,  $M_j$  is the standardised management score (the z-score of the average of 21 individual management questions),  $Z_j$  are characteristics of the college – in particular its overall size (which is a proxy for school resources), a measure of prior achievement of incoming cohorts and demographics.  $R_k$  are geographic characteristics.

The results are in Table 2. Each panel relates to a different institutional-level outcome and in general there is a positive relationship between management practices and educational outcomes. These are significant in the case of (NSS) share satisfied, employer satisfaction and the share of learners achieving at least one aim at level 2 or level 3, even in the more saturated specifications where prior attainment of learners is controlled for. Figure 4 shows the scatter plots corresponding to columns (1) and (4) for the NSS share satisfied and proportion of learners achieving level 3 aims.

These results suggest that management practices might help explain differences in college performance. We note however, that these specifications have few degrees of freedom and as such the significance of the relationships tends to be lost when we consider alternative specifications in the robustness checks such as when we include extensive controls for survey noise like interviewer dummies, day of the week that the interview was held, etc. (see Appendix

---

<sup>17</sup> In panel B of Appendix Table B3 we replicate the same exercise for the subsample of colleges which we use in our individual level analysis and the findings are similar: there is balance in the key outcomes of level 2 and level 3 achieved, and some negative selection in the student satisfaction and number of courses per learner.

Table B4).<sup>18</sup> Overall, this analysis suggests that there is a positive association between management practices and aggregated outcomes at the college level and motivates our further, better-identified analysis at the individual level which follows.

## ***2.2 Management practices and educational outcomes at the individual level***

In our individual level analysis we estimate logit regressions of the form:

$$Y_{ijk} = \beta_0 + \beta_1 M_j + \beta_2' X_i + \beta_3' Z_j + \beta_4' R_k + u_{ijk} \quad (2)$$

where  $Y_{ijk}$  are three core binary outcomes for individual  $i$ , in college  $j$ , and region  $k$ : whether or not a learner has achieved qualifications at level 2 (equivalent to GCSE), level 3 (upper secondary level), or is pursuing a degree, by age 20.<sup>19</sup> As before,  $M_j$  is the standardised management score and  $Z_j$  are college characteristics, in particular, college size, and  $R_k$  are key geographic characteristics.<sup>20</sup> We now control for a series of individual characteristics, including prior achievement and demographics, within the vector  $X_i$ . The coefficient of interest is  $\beta_1$  which gives the effect of a standard deviation increase in the management scores on the probability of achieving the specified educational outcomes.

Our results are summarised in Table 3. Again, there appears to be a positive relationship between the probability of achieving each of the three educational outcomes we measure and management practices. With the likelihood of achieving level 2 or level 3 qualifications (panels A and B), the raw correlations reported in column (1) gain significance when controlling for college size (column 2), and the coefficients on management practices unchanged. The relationship is about double in magnitude for level 3 aims, suggesting that management practices matter in particular at more advanced educational levels. Controlling for individual demographics and prior achievement, i.e. estimating college value-add, halves the magnitude of the effects (column 3) but they are still significant. Finally, adding geographic controls has little impact. On the likelihood of pursuing a degree by age 20, the coefficients on management practices are similar in magnitude to level 3, but the results are only significant in the value-

---

<sup>18</sup> In particular, Figure 4 reveals the presence of an outlier college that has particularly low management scores and also does worse in its outcomes. The significance of some coefficients for some of the institutional level outcomes reduces when this college is excluded.

<sup>19</sup> In further analysis (not reported here), we analyse the relationships between management practices and whether learners have achieved non-graduate tertiary level qualifications (at levels 4 and 5) by age 20. These qualifications are not the focus of our analysis as the number of learners pursuing these qualifications is much smaller (about 4% of the cohort). The coefficients are small, positive and insignificant.

<sup>20</sup> College size proxies overall resources because funding is highly correlated with the number of students.

add specifications.<sup>21</sup> The results for level 3 and pursuing a degree are represented graphically in Figure 5.

The upshot of this analysis is that a 1 standard deviation increase in the management score (0.36 in terms of the raw scores on this sample, as shown in Table B1) is associated with around a 2 percentage point increase in the likelihood of achieving at least a level 3 qualification (for which the mean in the sample is 66%), or pursuing a degree by age 20 (for which the mean is 23%). Proportionally, this represents a higher impact on those pursuing a degree. The relationship is smaller in magnitude, but still significant in terms of level 2 aims (just under 1 percentage point, and on average 86% of the individuals in the sample achieve this level). The magnitude of the association between management practices and these outcomes is relatively high when we consider the effect of attending a college which is of higher value added in general. Aucejo et al. (2020) find that a one standard deviation increase in college value added increases the likelihood of having achieved level 3 or attending university by about four percentage points.

#### *Robustness*

We find that these results are robust to a number of alternative specifications, including estimating a basic linear probability model, dropping the outlier college (this actually increases the magnitude of our results), adding additional college level controls to include characteristics of the principals<sup>22</sup> and adding the full set of survey noise controls (see Table B5).<sup>23</sup> While our focus in this analysis is on colleges where principals have at least 6 years of tenure at college, and on the 2012 GCSE cohort of learners (that enrolled in the colleges in 2013) for whom we observe educational outcomes by age 20, our results are not sensitive to alternative feasible cut-offs (results available on request). For example, results for the 2012 cohort are very similar on the sample of colleges where principals were in the college for 7 or 5 years; and remain positive and significant (for level 3 outcomes) when considering the 2013 GCSE cohort (for whom we are able to track educational outcomes up until the age of 19).

---

<sup>21</sup> Constraining the sample in Panels B and C to the learners in Panel A (those that take level 2 courses at the college) also yields similar results though the size of the coefficient in Panel C is slightly smaller. This restriction involves removing higher ability learners – those that entered only to pursue qualifications at level 3 or higher - from the sample.

<sup>22</sup> These include dummies for gender, qualifications (FE management specific qualification, teaching qualification), experience in industry and dummy for principals with 6 years or more of tenure at the college.

<sup>23</sup> We also note that in equivalent regressions on the full sample of colleges (i.e. to include colleges where principals' tenure is below 6 years), the coefficients are still positive for all outcomes, but they are not significant at conventional levels. This is not surprising because it seems likely that management practices will change with the leadership team over time and therefore the management scores as at 2019 are likely to be a more noisy measure in the enlarged sample.

### *The relative importance of different types of practices*

We also explore whether any particular categories of management practices might be driving these results and find that across outcomes, the magnitudes of the coefficients on management scores for the different management practices groupings are similar, but significance varies (Table 4). For level 2 and level 3 outcomes, the coefficients on the targeting z-score is the most precisely estimated. With respect to the likelihood of pursuing a degree by age 20, people management practices seem to matter most. Consistently across outcomes, operational practices seem to matter less. These findings are consistent with the international evidence on schools (Bloom et al., 2015), which found that people management practices had the strongest relationship with pupil outcomes from all the management practice groupings (followed by targets).<sup>24</sup> People management practices also appear to be worse on average in our sample (Table 1), and are an area where colleges (and public sector institutions) are restricted from implementing the types of incentives that are considered “best practice” in the WMS.

### ***2.3 Do management practices matter more for disadvantaged learners?***

Having established that management practices appear to matter for individual level educational outcomes, we next explore whether they matter more for students coming from a disadvantaged background (as measured using eligibility for free school meals (FSM) while at school). We expect this to be the case, as such learners may have less access to support at home such that they benefit more from well-organised and structured practices in the classroom. This hypothesis is supported by the literature linking school resources to student outcomes. Higher school expenditure or lower class size appear to matter more for students from poorer backgrounds (Gibbons and McNally, 2013). On the other hand, it is also possible that a lack of complementary inputs from home could reduce impacts. For example, Fryer’s (2017) evaluation of a management training intervention in schools found treatment effects to be weaker for disadvantaged students.

To test our hypothesis, we include an interaction between the management z-score and FSM status, and also test the robustness to including college fixed effects and adding interactions between FSM status and other institutional indicators. The advantage of this strategy is that we can control for any systematic differences between colleges that could in principle be correlated with management practices overall. Our results have a causal

---

<sup>24</sup> See Bloom et al. (2015), online appendix, Table B4.

interpretation provided that there is no omitted variable that is correlated with the interaction between FSM status and the management score. The results are shown in Table 5. Column (1) replicates our basic results (now using OLS specifications)<sup>25</sup> with the full set of controls for other demographic characteristics and prior achievement, reporting the coefficient on FSM status which is negative, showing that those from poor socio-economic backgrounds are less likely to achieve good educational outcomes.<sup>26</sup> Column (2) includes the interaction term which is positive and significant for level 2 or level 3 outcomes, though there is no evidence of heterogeneity in terms of the likelihood of pursuing a degree by age 20. Column (3) then shows that these significant interaction terms survive when college fixed effects are included (the zManagement score drops out in these specifications). This is robust to interacting the FSM status dummy with other institutional characteristics included in the controls (college size, regional GDP and population density, as in column (4)), as well as restricting panels B and C to the same sample as in Panel A.

These results imply that raising the management practices in a college from the 10<sup>th</sup> percentile to the 90<sup>th</sup> percentile (an increase of 2.49 standard deviations) is associated with an 8 percentage point higher likelihood of achieving a level 3 qualification for learners on Free School Meals.<sup>27</sup> It is nearly half of the gap in the raw probability of FSM and non-FSM students achieving a level 3 qualification by age 20 (which is 19 percentage points). This type of “effect” would apply for every FSM student – as an improvement in management practices in a college would affect future cohorts too. This evidence suggests that the management practices in colleges are likely to be especially important for improving intermediate educational outcomes for disadvantaged learners, and therefore for improving social mobility. We explore what this implies for labour market outcomes in Section 3 below.

---

<sup>25</sup> Note that the results reported here are based on a linear probability model (column (2) of Appendix Table B5), because of the complexities of determining statistical significance of interaction terms in non-linear models (Norton, Wang and Ai, 2004). We note that the coefficients in logit estimations are very similar – the equivalents for Column (2) of Table 5 are reported in Appendix Table B6 where we summarise heterogeneity analysis across demographic characteristics and prior achievement.

<sup>26</sup> It might seem curious that FSM students are not less likely to enter university compared to other students. But this is because exam results at age 16 are included in these regressions. Without including controls, FSM students are less likely to enter university at age 20 by 9.5 percentage points compared to other students.

<sup>27</sup> In order to calculate this we use the coefficients in column (3), since college fixed effects specifications do not allow us to observe the main effect of college management practices which varies at the institution level. Panel B column (3) shows that the main coefficient is 0.016, and the interaction term coefficient is 0.017. The total “effect” for FSM learners is therefore 0.033. This implies that a 1 standard deviation increase in management practices is associated with a 3.3 percentage point increase in the probability of FSM learners achieving level 3. In our thought experiment of moving from the 10<sup>th</sup> to the 90<sup>th</sup> percentile of management scores (2.49 standard deviations), this is  $3.3 \times 2.49 = 8.2$  percentage points.



## ***2.4 Other sources of heterogeneity***

In Table 6, we investigate whether there is heterogeneity in the effect of management practices according to a measure of prior achievement - whether learners entered with ‘good’ grades at GCSE.<sup>28</sup> Although this does not influence outcomes at levels 2 or 3, it is important for influencing whether learners pursue a university degree by age 20. In fact, the effect of management practices on this outcome is driven by these learners. The coefficient on the interaction between whether learners have good GCSE grades and the management score of the college remains positive and significant (at the 10% level) when controlling for college fixed effects and including other interactions with institutional characteristics (college size, regional GDP and population density). We find no evidence of heterogeneity across other key demographic characteristics of learners available in our data (e.g. gender).

We also consider heterogeneity in the effect of management practices by institutional characteristics and find that the relationship between management practices and the probability of pursuing a degree appears to be driven by colleges where higher education (HE) is a main area of educational provision (the results are in Table 7). Around two thirds of surveyed colleges reported that HE was a main area of provision,<sup>29</sup> and such colleges tend to be on average larger, with fewer nearby competitor colleges, and in poorer areas. This analysis suggests that where HE is a key area of provision for a college, management practices appear to matter more for progression to HE qualifications. Perhaps the increased focus on HE at such colleges raises its prominence as an outcome for learners (either at the same college or elsewhere), and could explain why management practices appear to be more strongly associated with an increased likelihood of pursuing a degree in such colleges.<sup>30</sup>

## **3. Management practices, upper secondary education and labour market outcomes**

Our results imply that management practices matter for improving educational attainment, and this is particularly the case with respect to improving the probability that those from disadvantaged backgrounds achieve level 3 – upper secondary – qualifications. We know from

---

<sup>28</sup> To consider this issue, we replace a continuous measure of the GCSE score (1-9) with a threshold indicator showing whether the student achieves ‘good grades’ at GCSE (at least five GCSEs at Grade C or above). This is an important indicator of an individual’s level of preparation before they enter the FE college and will affect what courses they are able to enrol in once in the college. We note that there is no significant relationship between the continuous measure of GCSE scores and management practices and subsequent outcomes.

<sup>29</sup> 51 out of 79 colleges (65%) reported that HE is a main area of college provision in the survey. On the subsample of 45 colleges where the principal has been at the college for six or more years, the share is 69%.

<sup>30</sup> We also calculated various measures of geographic proximity to universities and found no evidence of differential effects of management practices on learner outcomes for colleges closer to universities.

the literature that level 3 qualifications matter for future earnings: a conservative estimate gives a 6% return in lifetime earnings.<sup>31</sup> Our results therefore suggest that management practices can be an important channel for improving labour market outcomes for disadvantaged groups.

We also explore the relationship between management practices and labour market outcomes of learners early on in their careers (by age 20) as permitted by the timeframe of data available to us – bearing in mind that these are not representative of a longer-term effect. This analysis is based on the sample of learners that are not in higher education at age 20. We have seen that good management practices increase the probability of entering higher education by this age. Therefore, assessing whether they have an impact on the school to work transition by that same age is complicated by this selection effect.

In Table 8, column (1), we estimate regressions equivalent to Equation (2) but with labour market outcomes as the dependent variables. Measures for the latter are log daily earnings and an indicator where an individual has been employed for more than 90 days, both measured at age 20 and shown in panels A and B respectively. Column (2) reports the FSM coefficient, and column (3) adds also the interaction with free school meal eligibility and labour market outcomes of learners. The results show a negative correlation between management practices and these outcomes which is precise in the employment regressions. The interaction between management practices and free school meal eligibility is, however, positive and significant for both outcomes in column (3). These remain positive but lose significance once college fixed effects are included in column (4). These outcomes relate to early days in the working lives of young people, and the results presented here are not necessarily representative of the longer-term effect (especially bearing in mind that management practices influence the probability of remaining within the education system). Nevertheless, this analysis provides tentative evidence that those from poorer backgrounds might experience more positive immediate labour market outcomes when they have attended better managed colleges.

#### **4. How management practices vary across FE colleges**

We have established that management practices appear to be positively correlated with performance at the institution level, and with individual learner outcomes, even in value-added

---

<sup>31</sup> See Machin et al. (2018), the working paper version of Machin et al. (2020) which also includes estimates of the returns to level 3 qualifications. These estimates are conservative compared to other papers which have used different approaches to estimate higher returns, although there is substantial heterogeneity according to what is studied at level 3 (Patrignani et al., 2017; McIntosh, 2006; DfE, 2021).

specifications and controlling for observable college characteristics. It is therefore interesting to understand whether there are any particular principal or college characteristics that tend to be associated with better management practices.

#### ***4.1 Do observable characteristics of colleges and their principals explain differences in management practices?***

We begin by exploring whether observable principal or college characteristics help to explain differences in management practices, estimating the following linear regression:

$$M_{jk} = \varphi_0 + \varphi_1'Z_j + \varphi_2comp_j + \varphi_3'R_k + u_{jk} \quad (3)$$

where for college  $j$ , in region  $k$ ,  $Z_j$  includes college size, an indicator of whether the college is a sixth form college and principal characteristics;  $comp$  is a measure of competition faced by the college – our core measure being an indicator for a college being in the top quartile in terms of the number of other colleges within a 20km radius; and as before,  $R_k$  includes our regional covariates.

The results are shown in Table 9. In general, the coefficients are of expected sign, though not significant at conventional levels. Larger organisations tend to be better managed, and sixth form colleges worse managed, but these relationships are not significant. Adding in characteristics of the principal, we see that there is a positive and significant coefficient on the FE-specific management qualification, but not on any other observable characteristics.<sup>32</sup> This finding suggests that this qualification is aligned with what is considered best-practice in the WMS but with a small sample it is difficult to ascertain this robustly. We also note that this relationship remains positive but is not significant once regional covariates are included.<sup>33</sup>

In the saturated specification, the only variable that appears to have explanatory power is our measure of competition – colleges facing a high degree of spatial competition from other colleges have 0.6 of a standard deviation higher management practices, *ceteris paribus* (column 4). The finding that competition is positively related to management practices is consistent with the broader literature on management practices in firms and hospitals (see, for example Bloom

---

<sup>32</sup> In Appendix Table B7 we estimate the simple association between management practices and various different measures of the qualification of the principal and of the workforce more generally. Whether the principal has a FE management-specific qualification is the only variable that comes out as statistically significant (and with a sizeable coefficient).

<sup>33</sup> We also explored pairwise correlations (i.e. not controlling for other variables) between the characteristics of principals and management practices and found that the only variable that was significant (at the 5% level) was the FE-specific management qualification.

et al., 2016 and Bloom et al., 2015b though competition does not appear to be a significant driver for schools in Bloom et al., 2015a).<sup>34</sup>

The fact that management practices are not correlated with many of these observable college and principal characteristics but are correlated with outcomes suggests that the relationships between educational outcomes and management practices which we document above are not driven by some obvious omitted variable at the institution level.

#### ***4.2 Do higher management scores simply reflect better leadership?***

In general, in the literature on management practices a distinction is made between management practices and the characteristics or styles of particular managers (for discussion, see Valero, 2021). While these two dimensions are clearly related, with leaders playing a key role in the way organisations are managed,<sup>35</sup> management practices are also likely to reflect processes on the ground which can be thought of as a technology – evolving slowly, and depending on other organisational and environmental factors. To the extent that the data allow, we seek to explore whether higher management scores can be attributed to attributes of particular principals in some way.

While we have found little evidence of a link between observable features of principals and management practices, these say nothing about underlying effectiveness of principals with respect to learner outcomes. We therefore draw on estimates of principal effectiveness (for level 2 and level 3 outcomes) from Ruiz-Valenzuela et al. (2017) to allow us to explore this. Indeed, this paper found that the principal fixed effects were not correlated with observable characteristics of principals.

At the institution level, we are able to match principal fixed effects to 61 colleges in our sample. We find that there is generally a positive relationship between these and the management score for level 2 effectiveness measures, but this is not significant except when regional covariates are included (see Appendix Table B8). The relationships are negative but not significant for level 3. Moreover, when we add the principal fixed effects into our core

---

<sup>34</sup> We note however, that the significance of this relationship is not robust to alternative measures of competition, including those based on a self-reported measure. We also considered whether lagged values of financial performance might help explain differences in management practices, to understand the extent to which colleges that are facing financial pressures might be induced to have better (or worse) management practices, but found no evidence of any such relationships.

<sup>35</sup> In our individual-level analysis, this provides the justification for our focus on colleges where the principal in place at the time of our management survey was also present at the time of the learners in question being at the college.

individual level regressions, the coefficients on management practices are unchanged (results not reported here).<sup>36</sup>

We can also further explore the relationship between leadership and management practices using questions asked in our survey. In addition to the core questions on operational management practices, we asked three questions on college leadership and accountability. The question on leadership vision and strategy measures the extent to which college leaders have an understanding of the broader set of challenges faced by the college, and the right mindset to address them and the score of this question is combined with the scores on accountability of leaders and the extent to which leadership roles are clearly defined to generate a Leadership score (see Appendix Table A1.2 for details). We find that this score is positively correlated with the core management practices scores (see Appendix Figure B4), and explore whether it appears to have an impact on learner outcomes. The results are reported in the Appendix (Table B9). When the leadership z-score is included in our core value add specifications, we find an effect on learner outcomes of similar magnitude as with management practices, but more noisily estimated. In the case of the probability of studying for a degree by age 20, there is no significant relationship for the leadership scores, and when both measures are included together only the management score survives. Overall this analysis suggests that despite these scores being correlated (and therefore likely to give rise to multicollinearity when included in the same regression), the management practice scores appears to have a more precise and robust relationship with learner outcomes at higher levels of achievement compared with measures of leadership and accountability.

The small sample size prevents us from drawing firm conclusions in this piece of analysis, but on balance this evidence suggests that management scores measured in this study are capturing something distinct from principal effectiveness or effective leadership.

## **5. Conclusion**

Post-secondary institutions are very important for building up the skills base in the UK and internationally, but much less is known about what drives performance in these institutions compared to schools or universities. We have conducted the first World Management Survey in the FE sector, and found that in this context, as in other sectors, there is variation in management scores which is correlated with important outcomes. Linking our survey data on

---

<sup>36</sup> In further analysis not reported here, we found no significant correlations between college fixed effects from Ruiz-Valenzuela et al. (2017) and management practices, though coefficients were positive and larger in magnitude for Level 3 fixed effects.

management practices with individual level administrative data on educational histories we find that structured management practices appear to matter as a predictor of learner achievement, even after controlling for prior achievement and demographic information. These effects are stronger for higher levels of educational achievement and suggest that an increase in management scores of one standard deviation is associated with a two percentage point increase in the probability of achieving level 3 qualifications or being enrolled at a university at age 20. Comparing our results with other work on value added (Aucejo et al. 2020), the “effect” of improving management practices alone is high relative to improving value added in general terms.

When evaluating whether management practices matter for students with different characteristics, we can interpret our estimates as causal under weaker identification assumptions because fixed effect estimates control for all systematic variation between institutions (including the average effect of management practices). We also include interactions between the student characteristic of interest and other institutional-level variables (i.e. not only management practices). Our findings are robust to these stringent controls. The finding of most substantive interest is for students who come from low socio-economic backgrounds – who disproportionately attend FE colleges. At lower levels of achievement (i.e. outcomes at levels 2 and 3), good management practices appear to matter more for learners from disadvantaged backgrounds. This suggests that improving management practices may play a role in improving labour market outcomes, and hence social mobility in the communities they serve. Furthermore, as ‘good management’ is a slow-changing technology, such an effect would apply to multiple cohorts of disadvantaged students, thus potentially having a profound effect on social mobility over time.

Students from disadvantaged backgrounds are much less likely to enter colleges with a sufficient level of preparation to advance all the way to higher education by age 20. Yet we find that well managed FE colleges are likely to help students who enter with a good level of academic achievement by the time they leave school. Thus, FE colleges have potential to improve outcomes at different parts of the ability distribution.

The institutions in our analysis perform well on average in terms of their management practice scores, and seem to be representative of the sector as a whole except for being larger. The analysis here suggests that improving management practices has some role to play for improving performance and therefore the skills base. However, we have not found evidence of largescale underperformance in this sector. This might be because the sector operates within a

strong accountability framework as it is and has been under sustained pressure for several reasons including government-initiated funding cuts.

Our analysis has highlighted some interesting avenues for future research. In particular, we explored whether better management practices simply reflect more effective principals, and found that there is some evidence that these two dimensions are related, but that management practices do not appear to simply reflect more effective leadership. It would be valuable to understand more about the interaction between management practices and leadership styles, given that management practices do not fully capture leadership and we know that principals do matter for outcomes in this sector (Ruiz-Valenzuela et al. 2017). It would also be useful to build the evidence base by evaluating management training programmes in an experimental setting.

## REFERENCES

- Angrist, J.D., Pathak, P.A. and Walters, C.R. (2013). Explaining charter school effectiveness, *American Economic Journal: Applied Economics*, vol. 5(4), pp. 1–27.
- Association of Colleges (2018), “AoC 2018 report on college finances”, September 2018
- Aucejo, E., Hupkau, C., and Ruiz-Valenzuela, J. (2020). Where versus What: College Value-Added and Returns to Field of Study in Further Education. CVER Discussion Paper No. 30. Centre for Vocational Education Research, London School of Economics.
- Augar Review (2019). Independent panel report to the Review of Post-18 Education and Funding. HMSO.
- Autor, D. H. (2019). Work of the Past, Work of the Future. Richard T. Ely Lecture, *AEA Papers and Proceedings*, 109: 1–32
- Bandiera, O., Hansen, S., Prat, A. and Sadun, R. (2020). CEO Behavior and Firm Performance. *Journal of Political Economy* 128, no. 4 (April 2020): 1325–1369.
- Besley, T., and Van Reenen, J. (2013). Investing for Prosperity: A Manifesto for Growth. LSE Growth Commission. <https://cep.lse.ac.uk/new/research/growth/default.asp>
- Bloom, N. and Van Reenen, J. (2007). Measuring and Explaining Management Practices Across Firms and Countries, *The Quarterly Journal of Economics*, 122, Issue 4, November 2007, Pages 1351–1408
- Bloom, N., Lemos, R., Sadun, R. and Van Reenen, J. (2015a). Does Management Matter in Schools? *Economic Journal*, Volume 125 (May), Pages 647–674
- Bloom, N., Propper, C., Seiler, S. and Van Reenen, J. (2015b). The Impact of Competition on Management Quality: Evidence from Public Hospitals, *The Review of Economic Studies*, Volume 82, Issue 2, April 2015, Pages 457–489
- Bloom, N., Sadun, R., & Van Reenen, J. (2016). Management as a Technology? (No. w22327). National Bureau of Economic Research.
- Britton, J., Farquharson, C., and Sibieta, L., (2019). 2019 Annual Report on Education Spending in England. The Institute for Fiscal Studies. UK.
- DfE (2021). Measuring the Net Present Value of Further Education in England 2018/19. <https://www.gov.uk/government/publications/measuring-the-net-present-value-of-further-education-in-england-2018-to-2019>
- Dobbie, W. and Fryer, R.G. (2013). Getting beneath the veil of effective schools: evidence from New York City, *American Economic Journal: Applied Economics*, vol. 5(4), pp. 28–60.
- Fryer, R.G.(2017). Management and Student Achievement: Evidence from a Randomized Field Experiment. Working Paper. Harvard University.
- Financial Times (2010). ‘Cinderella sector’ challenges universities, *Financial Times*, 31 October 2010
- Gibbons, S. and S. McNally. (2013). The Effects of Resources Across School Phases: A Summary of Recent Evidence. CEP Discussion Paper No. 1226. Centre for Economic Performance, London School of Economics



- Hodgson, A. (2015). *The Coming of Age for FE?: Reflections on the past and future role of further education colleges in England*. Institute of Education Press. UK.
- Hwa, Y. and C. Leaver (2021). Management in Education Systems. *Oxford Review of Economic Policy* vol. 37(2), pp. 367-391.
- Hupkau, C., S. McNally, J. Ruiz-Valenzuela, and G. Ventura. (2017). Post-compulsory education in England: choices and implications. *National Institute Economic Review*. No. 240. pp. R42-57.
- Machin, S., S. McNally, C. Terrier, and G. Ventura (2020). Closing the Gap Between Vocational and General Education? Evidence from University Technical Colleges in England. CVER Discussion Paper Number 31. Centre for Vocational Education Research, London School of Economics.
- Machin, S., S. McNally and J. Ruiz-Valenzuela. (2018). Entry Through the Narrow Door: The Costs of Just Failing High Stakes Exams. CVER Research Discussion Paper 014.
- Machin, S., S. McNally and J. Ruiz-Valenzuela. (2020). Entry Through the Narrow Door: The Costs of Just Failing High Stakes Exams. *Journal of Public Economics* 190.
- McIntosh, S. (2006). Further Analysis of the Returns to Academic and Vocational Qualifications. *Oxford Bulletin of Economics and Statistics*, 68: 225-51.
- McCormack, J., Propper, C. and Smith, S. (2014). Herding cats? Management and university performance, *Economic Journal*. Volume 124. Issue 578. Pages 534-564.
- McIntosh, S. (2006). Further Analysis of the Returns to Academic and Vocational Qualifications. *Oxford Bulletin of Economics and Statistics*, 68, 2, 0305-9049
- Norris, E. and Adam, R. (2017) All change. Why Britain is so prone to policy reinvention, and what can be done about it. Institute for Government.  
[https://www.instituteforgovernment.org.uk/sites/default/files/publications/IfG\\_All\\_change\\_report\\_FINAL.pdf](https://www.instituteforgovernment.org.uk/sites/default/files/publications/IfG_All_change_report_FINAL.pdf)
- Norton, E. C., H. Wang, and C. Ai. 2004. "Computing Interaction Effects and Standard Errors in Logit and Probit Models." *Stata Journal* 4 (2): 154–67.
- Patrignani, P., G. Conlon and S. Hedges. (2017). The earnings differentials associated with vocational education and training using the Longitudinal Education Outcomes data, Centre for Vocational Education Research, London School of Economics. Discussion Paper 007.
- Ruiz-Valenzuela, J., Terrier, C. and VanEffenterre, C. (2017). Effectiveness of CEOs in the Public Sector: Evidence from Further Education Institutions, CVER Briefing Note 005, Centre for Vocational Education Research, London School of Economics.
- Scur, D., Sadun, R., Van Reenen, J., Lemos, R. and Bloom, N. (2021). The World Management Survey at 18: lessons and the way forward, *Oxford Review of Economic Policy*, Volume 37, Issue 2, Summer 2021, pp.231-258.
- Valero, A. (2021) "Education and Management Practices", *Oxford Review of Economic Policy*, Volume 37, Issue 2, Summer 2021, pp.302-322.
- Wolf, A. (2011). Review of Vocational Education: the Wolf report. 2011.  
<https://www.gov.uk/government/publications/review-of-vocational-education-the-wolf-report>

**Tables**  
**Table 1: Descriptive Statistics**

Variable	Mean	SD	N
<b><i>Management practices</i></b>			
Management score	4.24	0.34	79
Operations score	4.59	0.36	79
Monitor score	4.35	0.44	79
Target score	4.12	0.45	79
People score	3.96	0.46	79
Leadership score	4.21	0.55	79
<b><i>Institution level outcomes</i></b>			
Share satisfied (NSS)	0.78	0.12	68
Learners satisfaction (FE Choices)	0.80	0.08	65
Employers satisfaction (FE Choices)	0.82	0.10	63
Ofsted Effectiveness (reversed score)	2.90	0.73	79
Proportion learners, achieved at least 1 aim, level 2 (16-19)	0.79	0.06	79
Proportion learners, achieved at least 1 aim, level 3 (16-19)	0.80	0.08	79
<b><i>College characteristics</i></b>			
Number of Learners	8837.46	6443.38	79
Log Number of Learners	8.79	0.85	79
Number Learners aged 16-19	3555.24	2033.55	79
Sixth Form	0.18	0.38	79
<b><i>Principal characteristics</i></b>			
Female	0.46	0.50	79
FE management specific qualification	0.33	0.47	79
Teaching qualification	0.72	0.45	79
6 + years of tenure at college	0.57	0.50	79
Experience: Industry	0.51	0.50	79
<b><i>Geographic, region level variables</i></b>			
High density: colleges	0.25	0.44	79
Region: GDP per capita (2017)	915.70	1594.84	79
Region: Population density (2017)	30811.06	8230.03	79

*Notes: Management score* is the average of all 21 WMS management scores. The scores of separate groupings of scores as indicated in Appendix Table A1.1, are then given. Leadership score is the average of the scores across the 3 leadership and accountability questions in Table A.1.2. *Outcomes*: Share Satisfied is the percentage of students satisfied with their course. Learner satisfaction is the percentage of learners that would recommend the college to their friends. Employer satisfaction denotes the percentage of employers that would recommend this training provider to another employer. Ofsted Effectiveness scores is the reversed raw score awarded to the institution (1-4, where 1 is outstanding and 4 is inadequate). Proportion learners, achieved level 2 is the proportion of learners aged 16 to 19 years old that achieve at least one aim level 2 out of the total learners aged 16 to 19 years old with at least one active aim level 2 (similar for level 3 achieved). *College characteristics*: Number of learners is the total number of learners recorded in the ILR. Sixth Form is a dummy variable that takes value one if the college is a Sixth Form. *Principal characteristics* variables are dummies=1 if the characteristics listed apply. *Geographic variables*: High density: colleges is a dummy that takes value of one if the college is in the top 25% of colleges with higher number of colleges within 20 kilometres. Population density and GDP per capita are regional measures (NUTS1) obtained from EUROSTAT.

**Table 2: Management Practices and Institution Level Outcomes**

	(1)	(2)	(3)	(4)
<b>A: Share Satisfied (NSS)</b>				
zManagement	0.040** (0.018)	0.040** (0.018)	0.035** (0.017)	0.029* (0.017)
Observations	68	68	68	68
Adjusted R-squared	0.100	0.086	0.086	0.123
<b>B: Learner satisfaction (FE Choices)</b>				
zManagement	0.020** (0.009)	0.019** (0.009)	0.016 (0.010)	0.017 (0.010)
Observations	65	65	65	65
Adjusted R-squared	0.061	0.082	0.081	0.063
<b>C: Employer satisfaction (FE Choices)</b>				
zManagement	0.030** (0.012)	0.029** (0.012)	0.030** (0.013)	0.033** (0.014)
Observations	63	63	63	63
Adjusted R-squared	0.076	0.094	0.081	0.089
<b>D: Ofsted (reversed score)</b>				
zManagement	0.121 (0.102)	0.109 (0.103)	0.102 (0.104)	0.108 (0.105)
Observations	79	79	79	79
Adjusted R-squared	0.015	0.016	0.061	0.037
<b>E: Share of Learners who achieved at least one aim level 2</b>				
zManagement	0.011* (0.006)	0.013** (0.006)	0.011** (0.006)	0.010* (0.005)
Observations	79	79	79	79
Adjusted R-squared	0.018	0.062	0.178	0.220
<b>F: Share of Learners who achieved at least one aim level 3</b>				
zManagement	0.007 (0.010)	0.011 (0.010)	0.014* (0.007)	0.013* (0.008)
Observations	79	79	79	79
Adjusted R-squared	-0.005	0.055	0.445	0.448
Size		x	x	x
Previous achievement			x	x
Demographics			x	x
Log population density				x
Log region GDP per capita				x

*Notes:* Robust standard errors in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. Each panel reports results for a different outcome variable as defined in the notes to Table 1. Column (1) shows the raw correlation between management zscore and the outcome variables of interest. Column (2) add the total number of learners in the college (in logs) as control. Column (3) adds the average GCSE score (in standard deviations) of the 2015 KS4 that each of the FE colleges in our sample in 2016. Column (4) adds the percentage of learners that are women, percentage of learners eligible with Free School Meals and percentage of learners that speak English at home of the 2015 cohort that enter each of the FE colleges in 2016. Log population density and log region GDP per capita (2017) are at the NUTS1 level.

**Table 3: Management Practices and Learner Outcomes**

	(1)	(2)	(3)	(4)
<b>A: Achieved Level 2 plus by age 20 (mean=0.86)</b>				
zManagement	0.010*	0.011**	0.006*	0.008**
	(0.006)	(0.006)	(0.004)	(0.003)
Observations	38501	38501	38501	38501
Clusters	44	44	44	44
<b>B: Achieved Level 3 plus by age 20 (mean=0.66)</b>				
zManagement	0.032*	0.036**	0.018**	0.019***
	(0.017)	(0.015)	(0.009)	(0.007)
Observations	44584	44584	44584	44584
Clusters	45	45	45	45
<b>C: Pursuing degree by age 20 (mean=0.23)</b>				
zManagement	0.025	0.033	0.017*	0.019**
	(0.021)	(0.020)	(0.010)	(0.008)
Observations	44584	44584	44584	44584
Clusters	45	45	45	45
Size		x	x	x
Previous achievement			x	x
Demographics			x	x
Log population density				x
Log region GDP per capita				x

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by age 20. Coefficients reported are the marginal effects using a logistic regression. Column (1) shows the raw correlation between management z-score and the outcome variables of interest. Column (2) add the total number of learners (in logs) as control. Column (3) adds the individual GCSE score (in standard deviations). Column (4) adds gender, saturated variables of ethnicity, Free School Meals eligibility and English spoken at home (these are missing for some learners (see appendix Table B2), and in such cases we mean code the variable and add a dummy for missing status). Log population density and log region GDP per capita are regional measures (NUTS1) from 2017.

**Table 4: Management Practices Groupings and Learner Outcomes**

Management practices in z-score:	(1) Operations	(2) Monitoring	(3) Targeting	(4) People
<b>A: Achieved Level 2 plus by age 20 (mean=0.86)</b>				
z-score	0.005 (0.004)	0.008** (0.004)	0.006*** (0.002)	0.007* (0.004)
Observations	38501	38501	38501	38501
Clusters	44	44	44	44
<b>B: Achieved Level 3 plus by age 20 (mean=0.66)</b>				
z-score	0.015* (0.008)	0.016* (0.009)	0.016*** (0.005)	0.019** (0.007)
Observations	44584	44584	44584	44584
Clusters	45	45	45	45
<b>C: Pursuing degree by age 20 (mean=0.23)</b>				
z-score	0.013 (0.009)	0.017* (0.010)	0.012 (0.008)	0.017** (0.007)
Observations	44584	44584	44584	44584
Clusters	45	45	45	45

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by 20. Coefficients reported are the marginal effects using a logistic regression. Each column replicates the specification in column (4) Table 3, but with average management scores across the groupings of practices as labelled in the columns (see Appendix Table A.1.1 for details on the specific practices within each category). As in Table 3, dependent variables are indicated by each panel A-C.

**Table 5: Do FE colleges help disadvantaged students?**

	(1)	(2)	(3)	(4)
<b>A: Achieved Level 2 plus by age 20 (mean=0.86)</b>				
zManagement	0.007** (0.003)	0.005* (0.003)		
FSM eligible	-0.048*** (0.005)	-0.051*** (0.004)	-0.052*** (0.004)	1.034* (0.553)
FSM * zMan		0.016*** (0.004)	0.015*** (0.004)	0.015*** (0.004)
Observations	38501	38501	38501	38501
Clusters	44	44	44	44
<b>B: Achieved Level 3 plus by age 20 (mean=0.66)</b>				
zManagement	0.018** (0.008)	0.016* (0.008)		
FSM eligible	-0.064*** (0.006)	-0.067*** (0.007)	-0.070*** (0.007)	0.609 (0.553)
FSM * zMan		0.017** (0.006)	0.013** (0.005)	0.015*** (0.005)
Observations	44584	44584	44583	44583
Clusters	45	45	44	44
<b>C: Pursuing degree by age 20 (mean=0.23)</b>				
zManagement	0.017** (0.008)	0.019** (0.009)		
FSM eligible	0.003 (0.006)	0.005 (0.006)	-0.001 (0.006)	-1.087* (0.557)
FSM * zMan		-0.010 (0.007)	-0.005 (0.006)	-0.006 (0.006)
Observations	44584	44584	44583	44583
Clusters	45	45	44	44

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by 20. Coefficients are estimated using OLS. Columns (1) and (2) include controls for college size, previous achievement, demographics, and regional GDP per head and population density in logs. Column (3) includes college fixed effects and controls for demographics and prior achievement. Column (4) adds interactions of FSM with size, region GDP per head and population density. The share of the Panel A with Free School Meal (FSM) eligibility is 16%, and in Panels B/C it is 14.8%: those where this information is missing are coded as zero and a dummy is included in the regression.

**Table 6: Heterogeneity by prior achievement**

	(1)	(2)	(3)	(4)
<b>A: Achieved Level 2 plus by age 20 (mean=0.86)</b>				
zManagement	0.010** (0.004)	0.011* (0.006)		
Passed GCSE	0.155*** (0.009)	0.155*** (0.008)	0.142*** (0.009)	-1.518* (0.805)
GCSE * zMan		-0.004 (0.006)	-0.005 (0.006)	-0.008 (0.006)
Observations	38501	38501	38501	38501
Clusters	44	44	44	44
<b>B: Achieved Level 3 plus by age 20 (mean=0.66)</b>				
zManagement	0.024*** (0.009)	0.025** (0.010)		
Passed GCSE	0.420*** (0.011)	0.421*** (0.012)	0.389*** (0.013)	-1.841** (0.859)
GCSE * zMan		-0.002 (0.009)	-0.006 (0.011)	-0.015** (0.007)
Observations	44584	44584	44583	44583
Clusters	45	45	44	44
<b>C: Pursuing degree by age 20 (mean=0.23)</b>				
zManagement	0.023** (0.010)	0.007 (0.005)		
Passed GCSE	0.334*** (0.020)	0.325*** (0.016)	0.278*** (0.015)	0.749 (1.381)
GCSE * zMan		0.044** (0.017)	0.026* (0.013)	0.024* (0.012)
Observations	44584	44584	44583	44583
Clusters	45	45	44	44

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by 20. Coefficients are estimated using OLS. All columns include controls for demographics and previous achievement (GCSE binary outcome = 1 if learner passed at least five GCSEs at grade C or above). Columns (1) and (2) include controls for college size, demographics, and regional GDP per head and population density in logs. Columns (3) and (4) add college fixed effects. Column (4) adds interactions of the GCSE dummy with size, region log population density and region log GDP. The share that passed GCSE (achieved 5 A-C) is 43.7% in this sample, and in Panels B/C it is 36.3%: those where this information is missing are coded as zero and a dummy is included in the regression.

**Table 7: Heterogeneity by colleges reporting HE as a main area of provision**

	(1)	(2)	(3)
<b>A: Achieved Level 2 plus by age 20 (mean=0.86)</b>			
zManagement	0.007** (0.003)	0.007** (0.003)	0.007 (0.008)
Main provision HE		-0.002 (0.012)	-0.002 (0.012)
Main provision HE * zMan			0.000 (0.009)
Observations	38501	38501	38501
Clusters	44	44	44
<b>B: Achieved Level 3 plus by age 20 (mean=0.66)</b>			
zManagement	0.018** (0.008)	0.018** (0.008)	0.014 (0.015)
Main provision HE		-0.001 (0.023)	-0.001 (0.024)
Main provision HE * zMan			0.006 (0.018)
Observations	44584	44584	44584
Clusters	45	45	45
<b>C: Pursuing degree by age 20 (mean=0.23)</b>			
zManagement	0.017** (0.008)	0.017** (0.008)	-0.015 (0.011)
Main provision HE		0.014 (0.025)	0.015 (0.022)
Main provision HE * zMan			0.044** (0.016)
Observations	44584	44584	44584
Clusters	45	45	45

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by age 20. Coefficients are estimated using OLS. Share of learners enrolled in a college with Higher Education as a main provision is 75.6% (31 out of 45 colleges). All columns include controls for college size, previous achievement, demographics, and region GDP per head and population density in logs.



**Table 8: Management Practices and labour market outcomes for young learners.**

	(1)	(2)	(3)	(4)	(5)
<b>A: Log daily earnings at age 20 (mean=2.853)</b>					
zManagement	-0.028*	-0.027*	-0.031*		
	(0.014)	(0.014)	(0.015)		
FSM eligible		-0.135***	-0.141***	-0.129***	-0.364
		(0.018)	(0.020)	(0.019)	(1.627)
FSM * zMan			0.035*	0.018	0.023
			(0.019)	(0.017)	(0.016)
Observations	34559	34559	34559	34558	34558
Clusters	45	45	45	44	44
<b>B: Employed more than 90 days at age 20 (mean=0.725)</b>					
zManagement	-0.015***	-0.014***	-0.016***		
	(0.005)	(0.005)	(0.005)		
FSM eligible		-0.118***	-0.121***	-0.114***	-1.836***
		(0.008)	(0.008)	(0.008)	(0.550)
FSM * zMan			0.013**	0.009	0.012*
			(0.006)	(0.008)	(0.006)
Observations	44584	44584	44584	44583	44583
Clusters	45	45	45	44	44

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Coefficients are estimated using OLS. Columns (1) to (3) include controls for college size, previous achievement, demographics, and regional GDP per head and population density in logs. Column (4) includes college fixed effects and controls for demographics and prior achievement. Column (5) adds interactions of FSM with size, region GDP per head and population density. The share of the Panel A with Free School Meal (FSM) eligibility is 12.4%, and in Panel B it is 14.8%: those where this information is missing are coded as zero and a dummy is included in the regression.

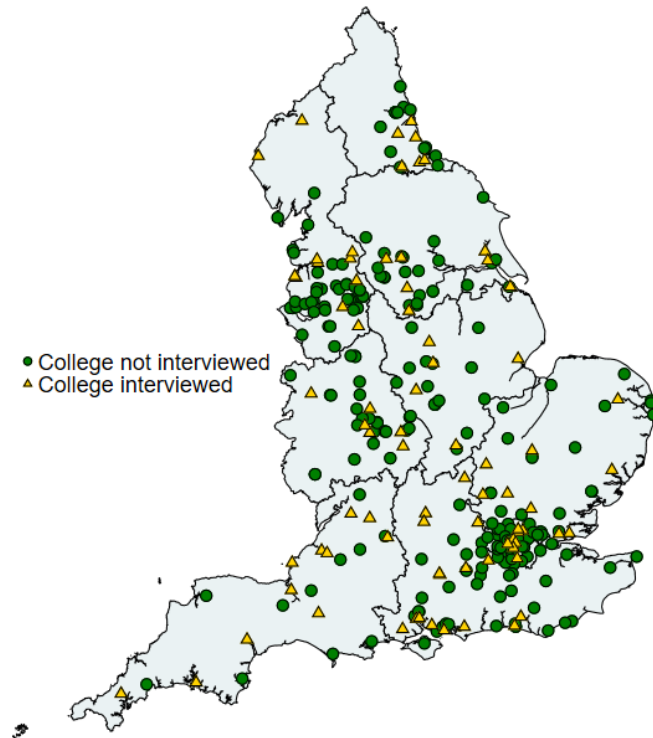
**Table 9: Management Practices and College/ Principal Characteristics**

	(1)	(2)	(3)
Log Number of Learners	0.129 (0.140)	0.049 (0.148)	0.048 (0.148)
Sixth Form	-0.146 (0.324)	-0.203 (0.326)	-0.212 (0.332)
Female		-0.287 (0.240)	-0.313 (0.240)
FE management specific qualification		0.395* (0.226)	0.272 (0.225)
Teaching qualification		-0.284 (0.240)	-0.246 (0.237)
6 + years of tenure at college		0.122 (0.219)	0.148 (0.222)
Experience: Industry		0.062 (0.226)	-0.078 (0.237)
High density: colleges			0.635** (0.261)
Region: log pop density in 2017			-0.333 (0.302)
Region: log gpd per head in 2017			0.108 (1.090)
Observations	79	79	79
Adjusted R-squared	-0.003	-0.001	0.013

*Notes:* Robust standard errors in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. Dependent variable is the standardised management score. High competition variables are dummies=1 if spatial competition measure is above the 75% percentile.

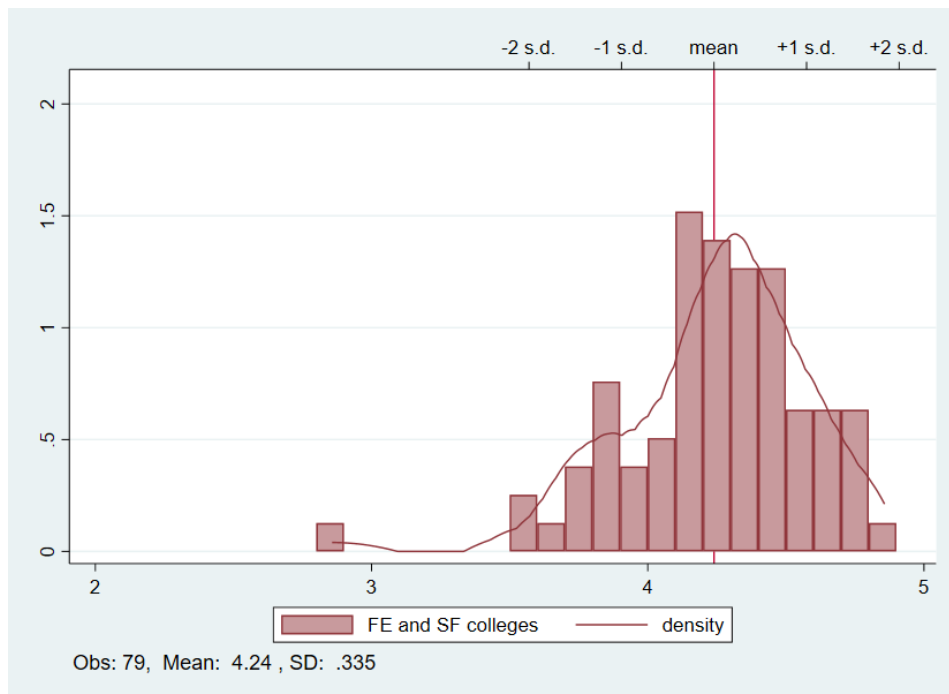
## Figures

**Figure 1: Map of population of colleges and sample of colleges interviewed.**



*Notes:* Analysis based on geocoded location of colleges in the population and our sample.

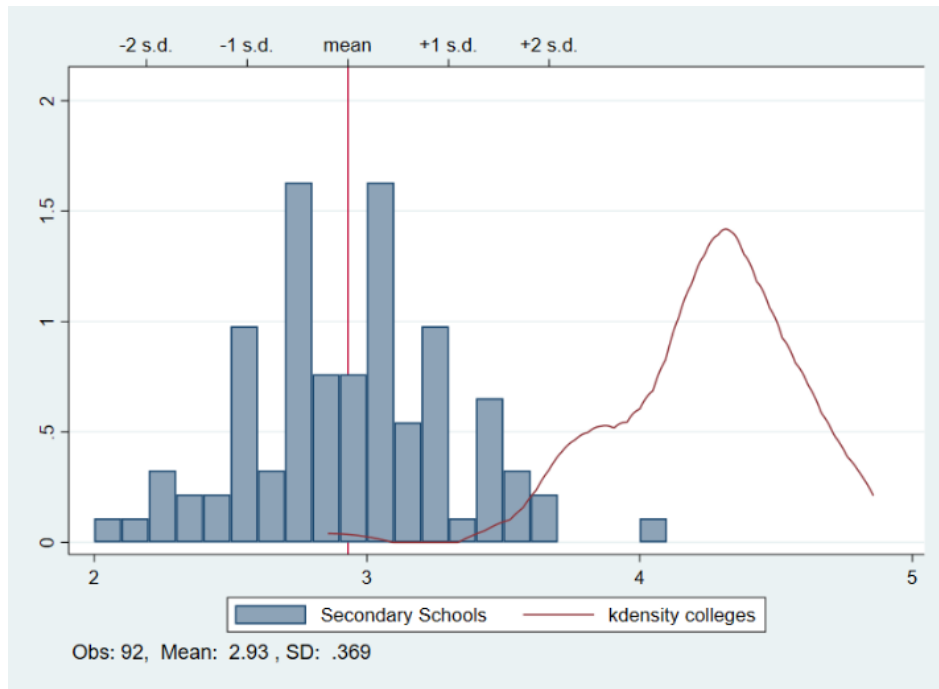
**Figure 2: Histogram of Management Scores in FE Colleges**



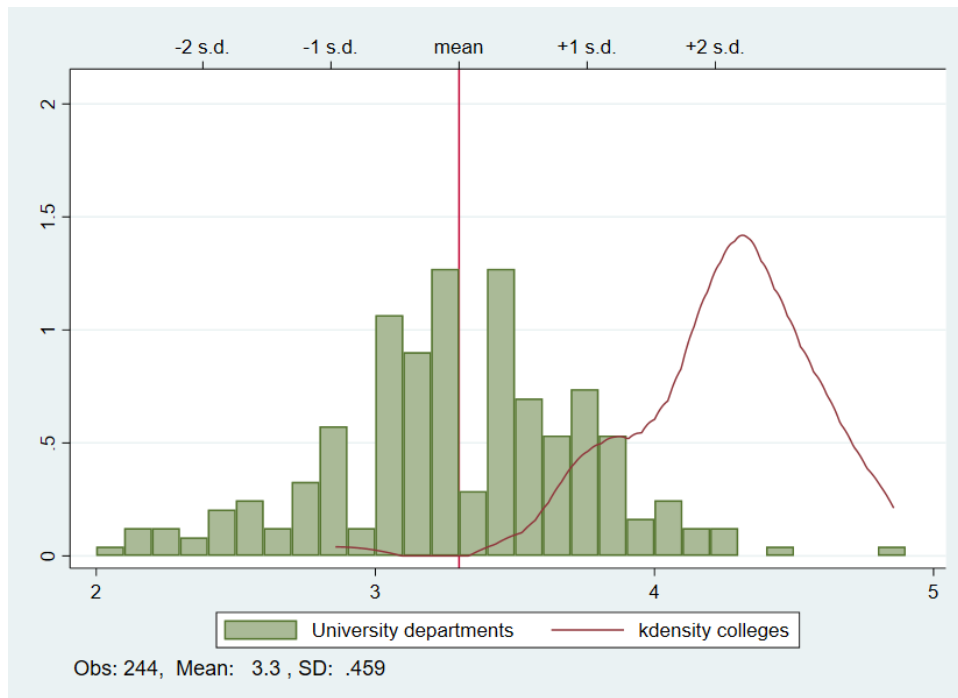
*Notes:* Histogram of management scores based on sample of 79 colleges.

**Figure 3: Histogram of Management Scores in FE Colleges vs Schools**

**Panel A: FE versus Schools**



**Panel B: FE vs University Departments**

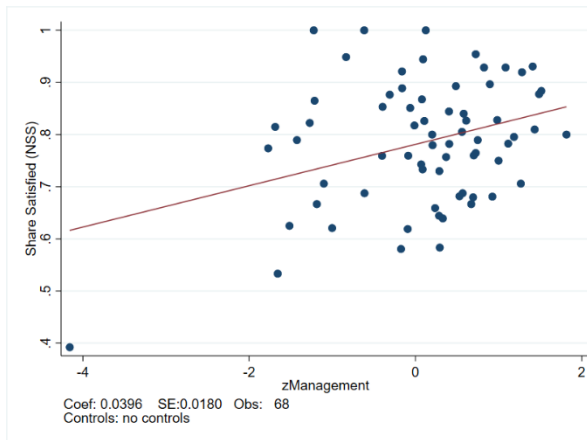


*Notes:* Schools data from the UK sourced from Bloom et al. (2015a), Universities Data sourced from McCormack et al. (2014).

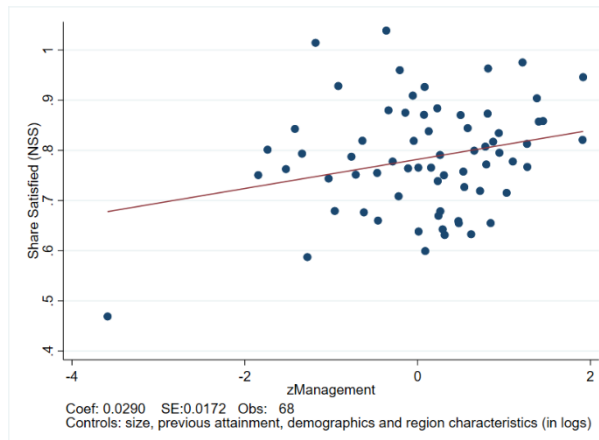
**Figure 4: Management Practices and Institution Level Outcomes**

**Panel A: NSS Share Satisfied**

**Raw correlation**

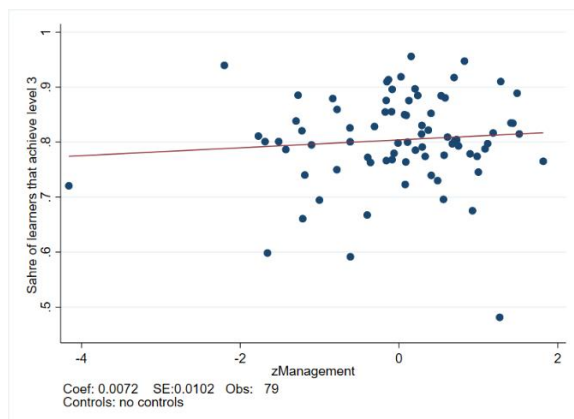


**Controls**

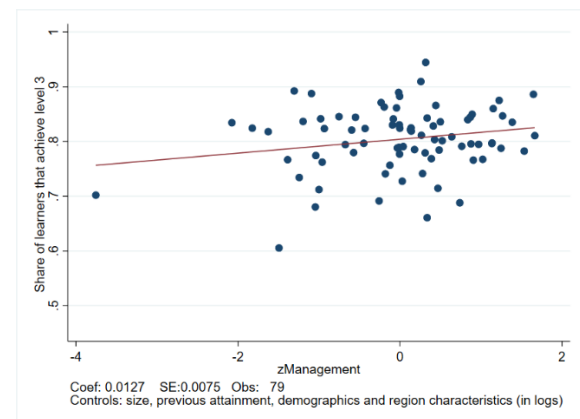


**Panel B: Proportion of learners achieving level 3 aims**

**Raw correlation**



**Controls**

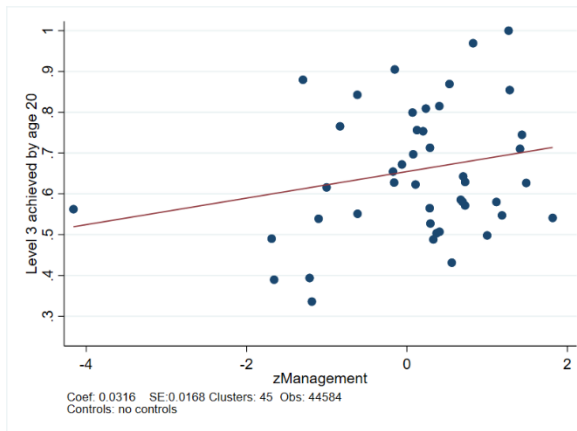


*Notes:* Scatter plot of simple correlation between NSS share satisfied (Panel A) and the proportion of learners that achieve Level 3 aims (Panel B) versus management score (standardised), corresponding to Table 2 columns (1) and (4) respectively

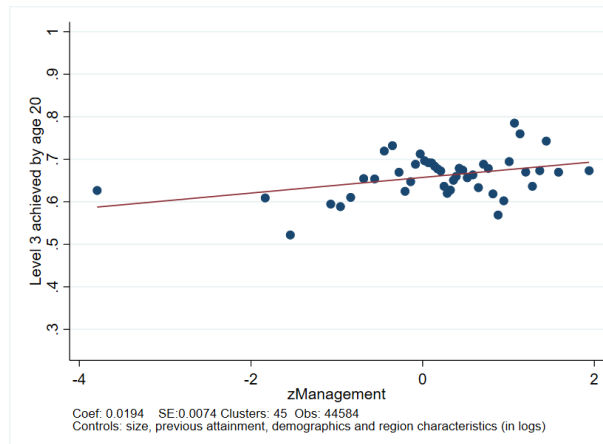
**Figure 5: Management Practices and Individual Level Outcomes**

**Panel A: Achieved Level 3 plus by age 20**

**Raw correlation**

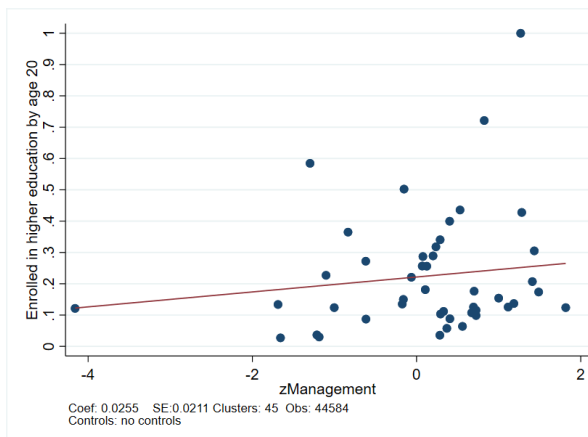


**Controls**

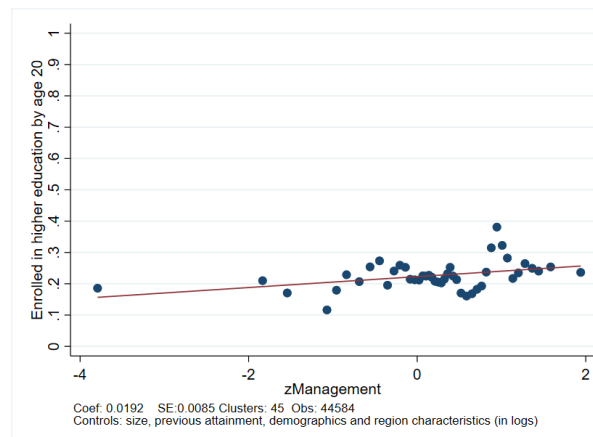


**Panel B: Pursuing a degree by age 20**

**Raw correlation**



**Controls**



*Notes:* Bin scatter plot of simple correlation between the probability of achieving level 3 aims (panel A) and studying for a degree (panel B), and the management score (standardised), corresponding to Table 3 columns (1) and (4) respectively

## APPENDICES

### Appendix A: Data Appendix

#### A1. World Management Survey of FE Colleges

The 21 management practices questions are outlined in Table A1.1 Table A1.2 sets out the questions on leadership and autonomy.

**Table A1.1: Management Practices Questions**

<b>Operations</b>		
Q0	Matching learners to courses	Measures how well learners are recruited and retained in well-matched qualifications/courses
Q1	Standardisation of teaching processes	Measures how well materials and practices are standardised and aligned in order to be capable of moving learners through learning pathways over time, and ensuring courses are meeting the needs to individuals and employers
Q2	Personalisation of instruction and learning	Measures for flexibility in teaching methods and learner involvement ensuring all individuals can master the learning objectives
Q3	Data-driven planning and learner transitions	Measures if the college uses assessment to verify learning outcomes at critical stages, makes data easily available, and uses it intelligently to adapt learner strategies and course offerings accordingly
Q4	Adopting educational best practices	Measures how well the college incorporates teaching best practices and the sharing of these resources into the classroom
<b>Monitoring</b>		
Q5	Continuous improvement	Measures attitudes towards process documentation and continuous improvement
Q6	Performance tracking	Measures whether college performance is measured with the right methods and frequency, and the extent to which this is communicated with staff, governors and other stakeholders.
Q7	Performance review	Measures whether performance is reviewed with appropriate frequency and follow-up
Q8	Performance dialogue	Measures the quality of review conversations
Q9	Consequence management	Measures whether differing levels of college performance (not only individual teacher performance) lead to different consequences
<b>Target Setting</b>		
Q10	Target balance	Measures whether the system tracks meaningful targets tied to learner outcomes, in particular, the extent to which colleges set their own internal targets and use these in a sensible way.
Q11	Target interconnection	Measures whether the college and individual targets are aligned with each other and the overall system goals
Q12	Time horizon of targets	Measures whether the college has a rational approach to planning and setting the targets
Q13	Target stretch	Measures whether targets are appropriately difficult to achieve
Q14	Clarity and comparability of targets	Measures how easily understandable performance measures are and whether performance is openly communicated
<b>People / Talent Management</b>		
Q15	Rewarding high performers	Measures whether good teacher performance is rewarded proportionately
Q16	Fixing poor performers	Measures whether the college is able to deal with underperformers
Q17	Promoting high performers	Measures whether promotions and career progression are based on performance
Q18	Managing talent	Measures how well the college identifies and targets needed teaching, leadership and other capacity in the college
Q19	Retaining talent	Measures whether the college will go out of its way to keep its top talent
Q20	Attracting Talent/Creating a distinctive employee value proposition	Measures how strong the teacher value proposition is to work in the individual college

**Table A1.2: Principal Leadership and Accountability Measures**

<b>Principal Leadership and Accountability Measures</b>	
Q21 Leadership vision and strategy	Measures whether college leaders have an understanding of the broader set of challenges that the college, system and key actors face and the right mindset to address them by checking whether the vision is clearly defined, set with the collaboration of a wide range of stakeholders, broadly communicated, linked to learner outcomes and built upon a keen understanding of community / business needs
Q22 Clearly defined accountability for college leaders	Measures whether college leaders are accountable for delivery of targets (including quality, and learner outcomes), are held responsible through governance structures, with both college and individual-level consequences for good and bad performance, and are autonomous in order to make decisions that will directly affect the outcomes of these targets
Q23 Clearly Defined Leadership and Teacher Roles	Tests how clearly the roles, responsibilities and required attributes of teachers, learners and staff are defined within the college

***Sample of surveyed colleges in final analysis***

We conducted the World Management Survey in Further Education (FE) and Sixth Form (SF) colleges in England (WMS) between February and September 2019. 80 colleges and 83 principals were interviewed (three colleges were keen to have two separate participants in the survey). One of the colleges was established in 2017 we were unable to obtain outcome data for it. Therefore, our final sample used in our data analysis consisted of 79 colleges. We keep the interview with the most senior manager in colleges where more than one person was interviewed.

**A2.1. Data for institutional level analysis**

We constructed a dataset containing institution-level outcomes and college characteristics as follows.

***Institution level outcomes:***

National Student Survey (NSS, 2011-2019): The NSS collects students’ data on course satisfaction across educational institutions in the United Kingdom. We use the most recent overall satisfaction score available for each college in our sample. These data are made available by the Office for Students.

FE Choices learner satisfaction survey (2015-2019): this survey collects data on learner satisfaction across further education colleges in the United Kingdom. We use the most recent score for each college, which is the percentage of respondents that would recommend the college/training organisation to their friends or family. These data are made available by the Department for Education (DfE).

FE Choices employer satisfaction (2016-2019): this survey collects data on employer satisfaction across further education colleges in the United Kingdom. We use the most recent score for each college. For the academic year 2016, this denotes the percentage of respondents that would recommend the college/training organisation to another employer. For the academic year 2017 onwards, this represents the percentage of employers that would ‘extremely likely’ or ‘likely’ to recommend the college/training provider to another employer. These data are made available by the Department for Education (DfE).



Office for Standards in Education (Ofsted, 2016-2019): These data contain the outcomes of inspections carried over the years 2016 to 2019. We use the overall effectiveness score which reflects the inspectors' assessment of the quality of teaching, learning and assessment, personal development, behaviour and welfare and outcomes for pupils. We use the most recent data available for each college.

Individual Learner Records (ILR) 2017: This administrative dataset comprises all the courses taken by learners in the publicly funded further education sector in the year 2017. It contains also the level and type of the courses, and the age of the learners. We aggregate these data to construct institutional-level outcomes for learners aged 16 to 19, as well as institution characteristics such as the size of the college. These data are provided by the Department for Education (DfE).

***Learner characteristics and institutions attended:***

National Pupil Dataset (NPD): Details on learners' characteristics, previous achievement and the institution they attend at age 17 comes from Spring Census (2015), KS4 (2015) and NPD-Linked ILR (2016) respectively.

These three administrative datasets contain three relevant groups of variables for our analysis: (i) individual learner characteristics such as gender, free school meal eligibility (FSM) and whether they speak English language at home (ENG); (ii) learner attainment during compulsory schooling (GCSE exams); and (iii) further education institutions that learners attend at age 17.

We use the 2015 KS4 (GCSE) cohort, and link these learners to the further education institution they attend in 2016 in the NPD-Linked ILR. We then merge the individual characteristics (gender, FSM eligibility, ENG, and KS4 achievement) to the institution that the learners of the 2015 cohort attend in 2016 (age 17). To construct institution level demographic characteristics, we collapse the measures of individual characteristics of the 2015 cohort by the institution they attend in 2016.

These data are provided by the Department for Education (DfE).

***Geographic data:***

Spatial competition: We geocode colleges (and schools) based on their postcodes. We use college postcodes from the ILR 2017 and as reported in our survey. We obtain postcodes for secondary schools through EDUBASE, and retain secondary schools with some 16 to 18 provision that are open in year 2020 for the purposes of our analysis, and we drop independent schools, British schools overseas, pupil referral units, free schools, special schools, and alternative provision schools. Using this information, we calculate the number of relevant competitor institutions within a 20km radius of each college in our sample.

Region characteristics: We obtain the region-level (NUTS1) population density and gross domestic product per head in 2017. This data is sourced from EUROSTAT.

***Financial data:***

College accounts: We obtain college level financial data as well as other college level characteristics such as number of teachers. We use the most recent data available for the colleges in our sample over the years 2016-2019. These data are provided by the Education and Skills Funding Agency (ESFA).

### **A.2.2. Data for individual level analysis, young learners**

We create an individual level dataset that enables the analysis of the relationships between long-term learner outcomes and management practices at the institutions they attended. We focus our analysis on the 2012 KS4 cohort and the further education institution they spend most of their time at age 17, in year 2013. We observe educational outcomes up to 4 years after individuals leave compulsory education.

We measure educational outcomes up until age 20, including whether or not level 2 or higher was achieved by age 20, level 3 or higher was achieved by age 20, and enrolment in higher education by age 20. The individual level data also includes demographics (gender and ethnicity), family background information (free school meal eligibility and English as first language), learner attainment during compulsory schooling (KS4 GCSE points) and information on courses taken in FE colleges and subsequent enrolment in higher education. The separate sources are described in more detail here:

National Pupil Dataset-Linked Individual Learner Records and Key Stage 5 (2013-2016): These data include information on all the courses taken by learners in publicly funded institutions of further education from year 2013 to 2016. It includes extensive information on the FE colleges curricula, including level of the courses, the learning hours, type of qualification, learners' outcomes in the course and the learning hours. We use this data to construct the outcomes level 2 or higher achieved by age 20 and level 3 or higher achieved by age 20.

Spring Census (2012): These data contains learners' ethnicity, gender, eligibility for free school meals and English spoken at home. We use these data as covariates in our regressions.

Key Stage 4 data (2012): This contains the results of the GCSE exams for the cohort of 2012 and 2013. We standardise this measure and use it as our measure of prior achievement in our regressions.

Higher Education Statistics Agency (HESA, 2015-2016): This provides information on attendance to higher education institutions. We use this dataset to construct the outcome that individuals are observed in higher education by age 20.

Her Majesty's Revenue's and Customs (HMRC) 2004 to 2017: This administrative dataset records data on employment spells and earnings from individuals in the 2013 cohort in the young learners' sample. We use these data to construct the outcome variables of individuals' daily earnings and whether they have been employed 90 days or more in 2017 (aged 20).

### **A.2.3. Measures of principal effectiveness**

Ruiz-Valenzuela et al. (2017) employ a panel dataset of principals in FE colleges in England over the period 2003 to 2015, and combine it with institution-level data on education performance coming from the Individualised Learner Records (ILR), the National Pupil Database (NPD) and the Higher Education Statistics Agency (HESA), in addition to Staff Individualised Records (SIR) and Ofsted reports. Using this dataset, the authors estimate

principal and college “fixed effects”. The detailed methodology behind the estimation is given at Ruiz-Valenzuela et al. (2017).

The basic approach builds on Bertrand and Schoar (2003), and exploits the fact that we observe principals and colleges over time. We therefore see a given principal in different institutions, and a given institution managed by different principals.

In that paper, institution level regressions are estimated, where the level 2 and level 3 outcomes are the fraction of learners enrolled in each course respectively, that achieved that level of qualification. The “fixed effects” recovered for principals and colleges respectively are therefore in those same units, and we standardise these for the purposes of our analysis. We capture the most recent estimate of principal fixed effects for each of the colleges, achieving 61 observations (out of a total of 79 in our WMS sample).

#### **A.2.4. Sample selection**

Our group of interest are young learners enrolled in General Further Education (or Tertiary) colleges in England. The population of FE and Sixth Form colleges comprises 310 institutions, 80 of which were part of our management survey. For the institutional data, we use both learner-level administrative data as well as institution-level survey data. To construct the learner outcomes, we use the 2017 ILR and focus in learners age 16 to 19. We keep the learners with at least one active course and with information available about the level and outcome of the course. As a result of this procedure, one of the institutions interviewed for our study did not have data on the student outcomes and we drop it from our analysis. This leads to a final sample of 79 colleges. We construct achievement rates for courses level 2 or level 3, and gather data on student satisfaction (NSS), learner and employer satisfaction (FE Choices) and Ofsted inspections.

In the individual dataset, we use the 2012 KS4 cohort that permits the analysis of educational outcomes up to four years after the last year of compulsory education (by age 20). We drop the learners that are not aged 16 in year 11 (i.e. repeated a grade and the learners that stayed in the same school after grade 11 or that have a missing grade in the GCSE. Sometimes learners are enrolled in more than one college at age 17, but for the purpose of our analysis we will assign the learners to the college where they spend most of their learning time.

In the adult-learners’ we keep the learners aged 25-59 and pursuing a course of level 2 or higher. We use the learners that enter college in academic year 2013 or 2015, and focus on their labour market outcomes in year 2017. As with the young learners, sometimes adult learners are enrolled in more than one college in the same academic year. For the purpose of our analysis, we keep the college where they spend most of their time.

We focus our analysis in the colleges where principals have at least 6 years of tenure at college because this allows us to create a more plausible link between the management score collected in the WMS in 2019 and the college in 2013, which is when the 2012 GCSE cohort enrolled in further education. Overall, we have a sample of 277,520 young learners in the population of FE and Sixth Form colleges in England, of which 80,899 are in institutions interviewed for our survey and 44,584 are in colleges where the principals have at least 6 years of tenure at the college.

## **Appendix B: Further Analysis**

### **Appendix Tables**

**Table B1: Descriptive statistics, sample of colleges where principal tenure  $\geq 6$  years**

Variable	Mean	SD	N
<b><i>Management practices</i></b>			
Management score	4.28	0.36	45
Operations score	4.63	0.41	45
Monitor score	4.43	0.44	45
Target score	4.11	0.52	45
People score	3.99	0.46	45
Leadership score	4.27	0.55	45
<b><i>Institution level outcomes</i></b>			
Share Satisfied (NSS)	0.77	0.13	42
Learner satisfaction (FE Choices)	0.80	0.08	40
Employer satisfaction (FE Choices)	0.82	0.12	39
Ofsted Effectiveness (reversed score)	2.93	0.75	45
Proportion learners, achieved at least 1 aim, level 2 (16-19)	0.80	0.05	45
Proportion learners, achieved at least 1 aim, level 3 (16-19)	0.80	0.09	45
<b><i>College characteristics</i></b>			
Number of Learners	9814.84	7136.75	45
Log Number of Learners	8.93	0.76	45
Number Learners aged 16-19	3664.00	2053.58	45
Sixth Form	0.11	0.32	45
<b><i>Principal characteristics</i></b>			
Female	0.47	0.50	45
FE management specific qualification	0.38	0.49	45
Teaching qualification	0.69	0.47	45
6 + years of tenure at college	1.00	0.00	45
Experience: Industry	0.60	0.50	45
<b><i>Geographic, region level variables</i></b>			
High density: colleges	0.29	0.46	45
Region: Population density (2017)	975.17	1666.57	45
Region: GDP per capita (2017)	31375.07	8493.57	45

*Notes: Management score* is the average of all 21 WMS management scores. The scores of separate groupings (Appendix Table A1.1), are then given. Leadership score is the average across the 3 leadership and accountability questions in Table A.1.2. *Outcomes:* Share Satisfied is the percentage of students satisfied with their course in the NSS. Learner satisfaction is the percentage of learners that would recommend the college to their friends. Employer satisfaction is the percentage of employers that would recommend this training provider to another employer. Ofsted Effectiveness is the reversed raw score awarded to the institution (1-4, where 1 is outstanding and 4 is inadequate). Proportion learners, achieved level 2 is the proportion of learners aged 16 to 19 years old that achieve at least one aim level 2 out of the total learners aged 16 to 19 years old with at least one active aim level 2 (similar for level 3 achieved). *College characteristics:* Number of learners is from the ILR. Sixth Form is a dummy variable that takes value one if the college is a Sixth Form. *Principal characteristics* variables are dummies=1 if the characteristics listed apply. *Geographic variables:* High density: colleges is a dummy that takes value of one if the college is in the top 25% of colleges with higher number of colleges within 20 kilometres. Population density and GDP per capita are regional measures (NUTS1) obtained from EUROSTAT.

**Table B2: Individual Level Descriptive Statistics**

Variable	Mean	SD	N
<b>Outcomes</b>			
Achieved Level 2 by age 20	0.86	0.34	38501
Achieved Level 3 plus by age 20	0.66	0.47	44584
Pursuing degree by age 20	0.23	0.42	44584
<b>Prior achievement (KS4 data)</b>			
Average GCSE total points (standardised across the population)	-0.30	0.84	44584
<b>Characteristics (Census data)</b>			
Eligible for Free School Meals (FSM)	0.15	0.36	42951
Female	0.48	0.50	42951
Speak English at home	0.93	0.26	42951
Ethnicity: White	0.84	0.36	44584

*Notes:* N refers to the population of learners for which each variable was available. The prior achievement measure is standardised, and therefore represents standard deviations from the mean of the population of GCSE scores in the 2012 cohort. It shows that our sample on average achieved 0.3 standard deviations worse GCSE scores than the mean.

**Table B3: Balance checks**

Dependent variables in rows:	Coefficient on Interviewed dummy	s.e.	N
<b>A: Full sample</b>			
<b>Outcomes</b>			
Share satisfied (NSS)	-0.025	(0.017)	218
Learners satisfaction (FE Choices)	0.012	(0.012)	216
Employers satisfaction (FE Choices)	0.019	(0.017)	211
Ofsted (reversed score)	0.054	(0.093)	304
Share of learners that achieve level 2	0.003	(0.012)	308
Share of learners that achieve level 3	-0.009	(0.014)	308
<b>Size and resources</b>			
log Number Learner (All)	0.419***	(0.111)	309
log Number Aims (All)	0.370***	(0.100)	309
log Number Aims by Learner (All)	-0.049	(0.031)	309
log total expenditure	0.381***	(0.093)	301
log total expenditure over total learners	-0.016	(0.077)	293
<b>B: Colleges where principal has 6 or more years tenure at college</b>			
<b>Outcomes</b>			
Share satisfied (NSS)	-0.040**	(0.020)	218
Learners satisfaction (FE Choices)	0.008	(0.014)	216
Employers satisfaction (FE Choices)	0.015	(0.020)	211
Ofsted (reversed score)	0.088	(0.115)	304
Share of learners that achieve level 2	0.014	(0.015)	308
Share of learners that achieve level 3	-0.012	(0.017)	308
<b>Size and resources</b>			
log Number Learner (All)	0.527***	(0.137)	309
log Number Aims (All)	0.450***	(0.124)	309
log Number Aims by Learner (All)	-0.077**	(0.038)	309
log total expenditure	0.431***	(0.115)	301
log total expenditure over total learners	0.090	(0.095)	293

*Notes:* Robust standard errors in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. N refers to the population of colleges that have at least one learner in the 16-19 age group, for which the relevant variables are available.

**Table B4: Robustness on Institution Level Regressions**

	(1)	(2)	(3)	(4)	(5)
<b>A: Share Satisfied (NSS)</b>					
zManagement	0.029*	0.012	0.029*	0.042**	0.032
	(0.017)	(0.015)	(0.017)	(0.018)	(0.020)
Observations	68	67	68	68	68
Adjusted R-squared	0.123	0.039	0.123	0.256	0.182
<b>B: Learner satisfaction (FE Choices)</b>					
zManagement	0.017	0.012	0.017	0.021*	0.017*
	(0.010)	(0.013)	(0.010)	(0.012)	(0.010)
Observations	65	64	65	65	65
Adjusted R-squared	0.063	0.010	0.063	0.130	0.386
<b>C: Employer satisfaction (FE Choices)</b>					
zManagement	0.033**	0.024	0.033**	0.037**	0.034**
	(0.014)	(0.016)	(0.014)	(0.017)	(0.016)
Observations	63	62	63	63	63
Adjusted R-squared	0.089	0.036	0.089	0.040	0.204
<b>D: Ofsted (reversed score)</b>					
zManagement	0.108	0.002	0.108	0.067	0.023
	(0.105)	(0.078)	(0.105)	(0.132)	(0.166)
Observations	79	78	79	79	79
Adjusted R-squared	0.037	0.029	0.037	-0.035	0.044
<b>E: Share of Learners who achieved at least one aim level 2</b>					
zManagement	0.010*	0.012*	0.010*	0.006	0.009
	(0.005)	(0.007)	(0.005)	(0.006)	(0.006)
Observations	79	78	79	79	79
Adjusted R-squared	0.220	0.223	0.220	0.201	0.420
<b>F: Share of Learners who achieved at least one aim level 3</b>					
zManagement	0.013*	0.009	0.013*	0.012	0.008
	(0.008)	(0.010)	(0.008)	(0.008)	(0.008)
Observations	79	78	79	79	79
Adjusted R-squared	0.448	0.448	0.448	0.432	0.615
College size	x	x	x	x	x
Previous achievement	x	x	x	x	x
Demographics	x	x	x	x	x
Log population density	x	x	x	x	x
Log region GDP	x	x	x	x	x
Additional college controls			x		
Survey noise controls				x	
College size weights					x
Sample	All Colleges	Drop outlier	All Colleges	All Colleges	All Colleges

*Notes:* Robust standard errors in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. *Additional college controls* include a dummy for Sixth Form, dummies for gender of the principal/vice principal, FE management specific qualification, teaching qualification, experience in industry and 6 years or more of tenure. *Noise controls* include job post and tenure of interviewee, interviewer dummies, day of week, time of day, interview duration and an interview reliability measure.

**Table B5: Robustness on Individual Level Regressions**

	(1)	(2)	(3)	(4)	(5)
<b>A: Achieved Level 2 plus by age 20 (mean=0.86)</b>					
zManagement	0.008** (0.003)	0.007** (0.003)	0.009** (0.005)	0.007** (0.003)	0.007** (0.003)
Observations	38501	38501	37638	38501	38501
Clusters	44	44	43	44	44
<b>B: Achieved Level 3 plus by age 20 (mean=0.66)</b>					
zManagement	0.019*** (0.007)	0.018** (0.008)	0.028*** (0.008)	0.017** (0.008)	0.017** (0.008)
Observations	44584	44584	43711	44584	44584
Clusters	45	45	44	45	45
<b>C: Pursuing degree by age 20 (mean=0.23)</b>					
zManagement	0.019** (0.008)	0.017** (0.008)	0.022** (0.011)	0.020*** (0.008)	0.020*** (0.008)
Observations	44584	44584	43711	44584	44584
Clusters	45	45	44	45	45
College size	x	x	x	x	x
Previous achievement	x	x	x	x	x
Demographics	x	x	x	x	x
Log population density	x	x	x	x	x
Log region GDP	x	x	x	x	x
Additional college controls				x	
Survey noise controls					x
Estimation	Logit Main sample	OLS Main sample	Logit Drop outlier	Logit Main sample	Logit Main sample

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by 20. *Additional college controls* include a dummy for Sixth Form, dummies for gender of the interviewee, and whether or not the principal has a FE management specific qualification, Teaching qualification or Experience in industry. *Noise controls* include job-post and tenure of interviewee, interviewer dummies, day of week, time of day, interview duration and an interview reliability measure.



**Table B6: Heterogeneity Regressions – Logistic Specification**

		(1)	(2)	(3)	(4)
		FSM	Female	Passed GCSE	HE as a main provision
<b>A: Achieved Level 2 plus by age 20 (mean=0.86)</b>					
Variable * zManagement	Coefficient	0.007	-0.003	-0.005	0.002
	s. error	0.002	0.004	0.005	0.008
	z-stat	3.937	-0.964	-1.098	0.253
	Observations	38501	38501	38501	38501
<b>B: Achieved Level 3 plus by age 20 (mean=0.66)</b>					
Variable * zManagement	Coefficient	0.011	-0.007	-0.009	0.012
	s. error	0.006	0.007	0.007	0.014
	z-stat	2.014	-1.161	-1.348	0.827
	Observations	44584	44584	44584	44584
<b>C: Pursuing degree by age 20 (mean=0.23)</b>					
Variable * zManagement	Coefficient	-0.003	0.001	0.054	0.037
	s. error	0.007	0.005	0.020	0.014
	z-stat	-0.499	0.303	2.737	2.526
	Observations	44584	44584	44584	44584

*Notes:* Columns (1) - (4) show the coefficients for each of the dependent variables in Panel A, B and C, and for an interaction of the management score and learners' free school meal eligibility, gender, GCSE results, and college's Higher Education provision respectively. Regressions in Columns (1), (2) and (4) include zmanagement as independent variable, and controls for college size, previous achievement, demographics, and region GDP per head and population density in logs. The regression in Column (3) does not include previous achievement as control since it uses the indicator for passing GCSE. The sample mean of Free School Meal (FSM) eligibility is 15.4%, the share of females is 47.7%, the share that passed GCSE is 43.7% and the share of learners that is enrolled in a college with Higher Education as a main area of provision is 75.6% (31 out of 45 colleges). The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by 20. Coefficients are estimated using a logistic regression and margins using the methodology proposed in Norton, Wang and Ai (2004). Clustered standard errors (at college level) in parentheses.

**Table B7: Management Practices and principal and work force characteristics**

<b>Dependent variable:</b>				
<b>zManagement</b>	(1)	(2)	(3)	(4)
<b>Panel A: Principal characteristics</b>				
FE management specific qualification	0.407** (0.202)			
Teaching qualification		-0.198 (0.233)		
Masters in Business Administration (MBA)			0.062 (0.225)	
Vocational qualification				-0.020 (0.257)
Observations	79	79	79	79
Adjusted r-square	0.025	-0.005	-0.012	-0.013
<b>Panel B: Work force characteristics</b>				
% teachers with QTS	0.448 (0.528)			
% teachers with experience in industry		0.525 (0.477)		
Ofsted Teaching (reversed score)			0.297 (0.274)	
Observations	75	47	79	
Adjusted r-square	-0.001	-0.008	0.015	

Notes: Robust standard errors in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level

**Table B8: Management Practices and Principal Effectiveness**

<b>Dependent variable: zManagement</b>	(1)	(2)	(3)	(4)
<b>A: Level 2 effectiveness</b>				
Level 2: Within Estimate (std)	0.064 (0.125)	0.065 (0.128)	0.136 (0.123)	0.229** (0.109)
Observations	61	61	61	61
Adjusted R-squared	-0.012	-0.045	-0.033	0.111
<b>B: Level 3 effectiveness</b>				
Level 3: Within Estimate (std)	-0.031 (0.113)	-0.037 (0.124)	-0.061 (0.119)	-0.076 (0.124)
Observations	61	61	61	61
Adjusted R-squared	-0.016	-0.049	-0.049	0.063
<b>C: Weighed average Level 2 and 3</b>				
Level 2 and 3 weighted average (std)	0.018 (0.120)	0.019 (0.121)	0.066 (0.118)	0.130 (0.100)
Observations	61	61	61	61
Adjusted R-squared	-0.017	-0.050	-0.048	0.075
Size and sixth form dummy		x	x	x
High density: colleges			x	x
Log population density				x
Log region GDP per capita				x

*Notes:* Robust standard errors in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. Panels A and B each have a different estimate of principal effectiveness as the independent variable, corresponding to the level 2, level 3 fixed effects (standardised) respectively. Panel C then takes the average of the fixed effects, weighted using the share of learners enrolled in each level. Column (1) is a raw correlation, and controls are added as indicated. The “high density: colleges” variable is a dummy=1 if the spatial college competition measure is above the 75% percentile.

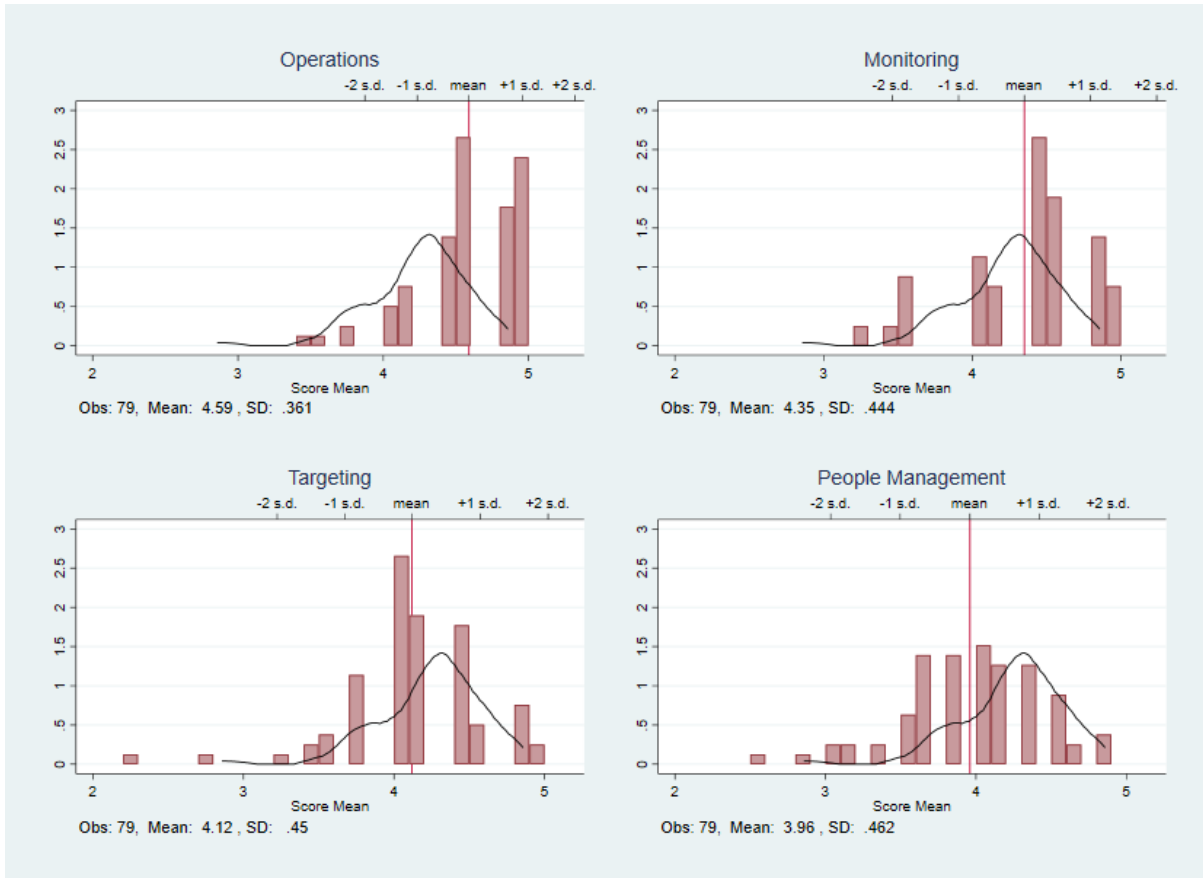
**Table B9: Leadership vs Management Practices, and Educational Outcomes**

	(1)	(2)	(3)
<b>A: Achieved Level 2 plus by age 20</b>			
zManagement	0.008** (0.003)		0.005 (0.005)
zLeadership		0.008* (0.004)	0.005 (0.007)
Observations	38501	38501	38501
Clusters	44	44	44
<b>B: Achieved Level 3 plus by age 20</b>			
zManagement	0.019*** (0.007)		0.014 (0.011)
zLeadership		0.018** (0.007)	0.010 (0.012)
Observations	44584	44584	44584
Clusters	45	45	45
<b>C: Pursuing degree by age 20</b>			
zManagement	0.019** (0.008)		0.018* (0.009)
zLeadership		0.012 (0.008)	0.003 (0.009)
Observations	44584	44584	44584
Clusters	45	45	45
Size	x	x	x
Previous achievement	x	x	x
Demographics	x	x	x
Log population density	x	x	x
Log region GDP per capita	x	x	x

*Notes:* Standard errors clustered at the college level in parentheses, \*\*\* denotes significance at the 1% level, \*\* 5% level and \* 10% level. The sample used in these regressions comprises all learners enrolled at a college where the principal has 6 years or more of tenure. Panel A comprises learners that are enrolled in a level 2 course by age 20, while panel B and C uses a sample of learners enrolled in any course by 20. Column (1) replicates Table 3 column (4). Logit regressions, marginal effects reported. In column (2) the leadership score is used instead of the core management practices score, and in column (3), both the management practices and leadership practices are included in the regressions.

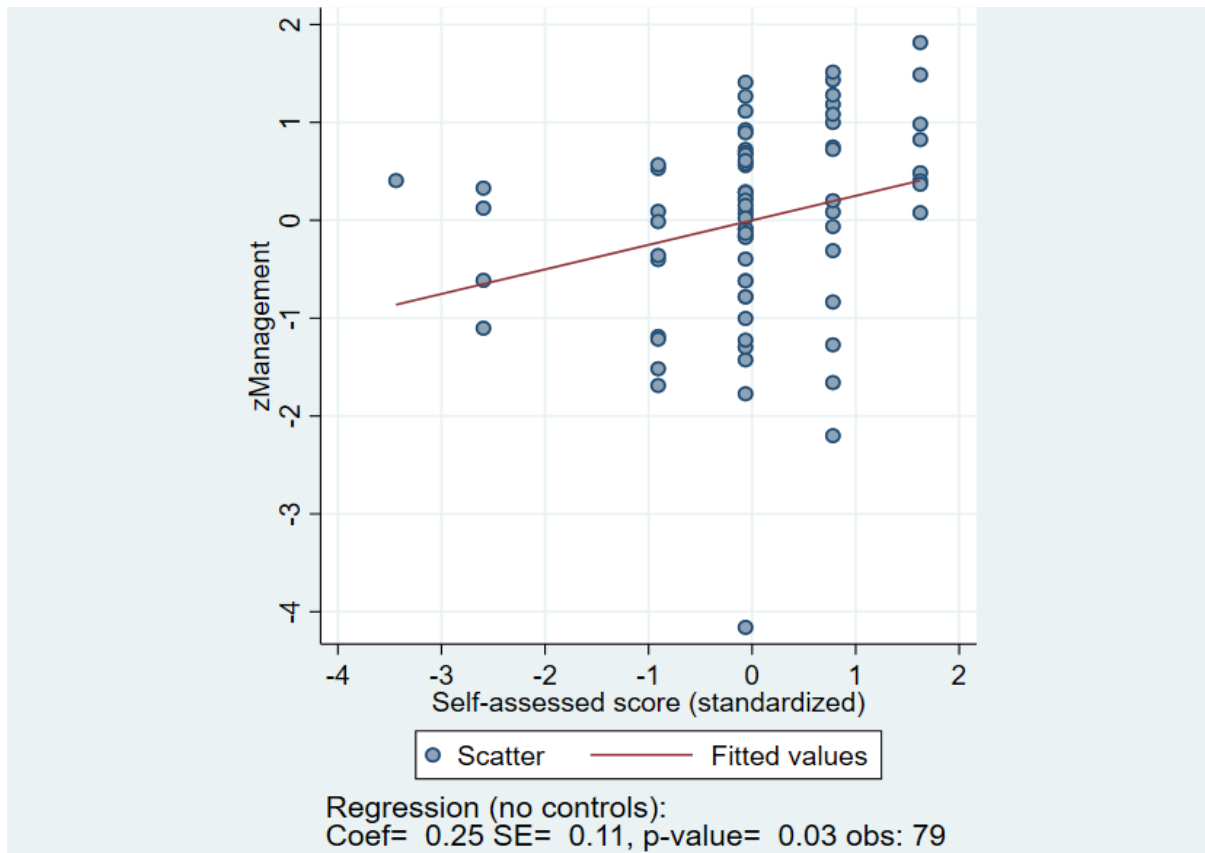
## Appendix Figures

**Figure B1: Histogram of different subgroups of management.**



*Notes:* Histogram of management scores across different groupings (set out in Appendix Table A1.1), based on the sample of 79 colleges.

**Figure B2: Management Scores and Self-Assessed Management Score**

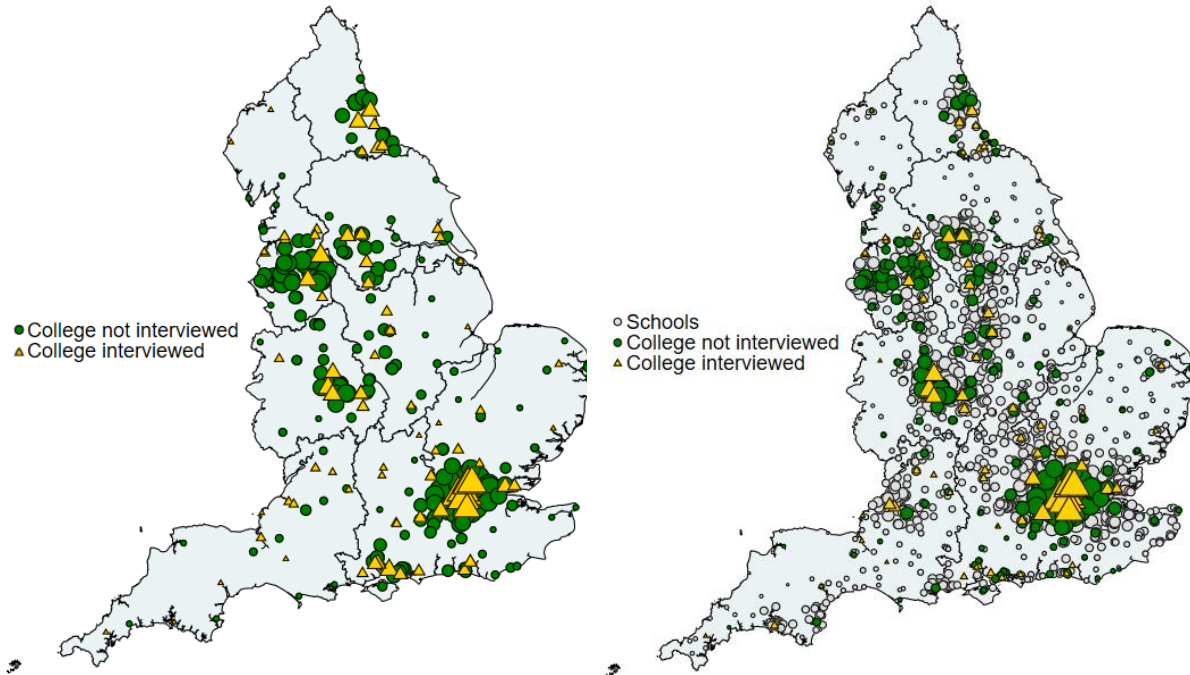


*Notes:* Scatter plot of zManagement scores against the standardised self-assessed score from the survey. The self-assessed score is the answer to the following question, asked after the interview was completed: “Ignoring yourself. How well managed do you think the rest of the college is on scale: 1 to 10, where 1 is worst practice, 10 is best practice.”

**Figure B3: Maps showing spatial competition measures**

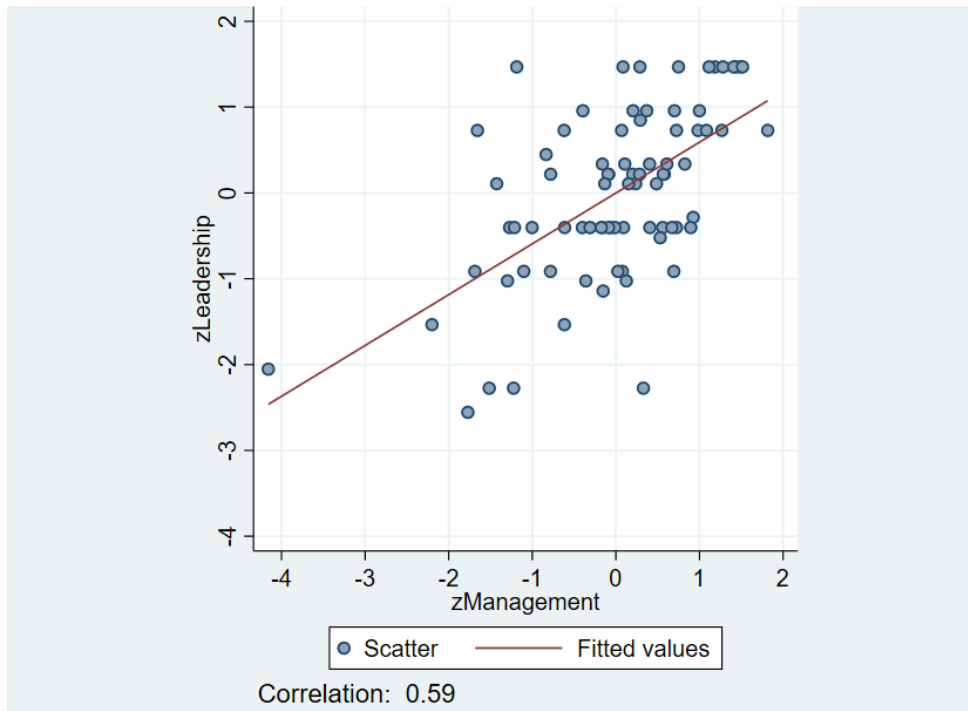
A: Colleges within 20km

B: Colleges and schools within 20km



*Notes:* Analysis based on geocoded location of colleges and schools. Yellow triangles depict colleges in our sample. Green circles show other colleges and grey circles (Panel B) show schools. The size of the triangles / circles gives the relative number of competing institutions within a 20km radius versus other institutions within each map.

**Figure B4: Management Scores and Leadership Scores**



*Notes:* Scatter plot of zLeadership scores (average of Q21-23, standardised) versus zManagement scores (average of Q0-Q20, standardised).