Evaluation of the early roll-out of the Early Career Framework (ECF)
Evaluation Study Plan

Evaluator (institution): National Foundation for Educational Research
Principal Investigator: Simon Rutt

PROJECT TITLE	Evaluation of the early roll-out of the Early Career Framework (ECF)
EVALUATOR (INSTITUTION)	National Foundation for Educational Research
PRINCIPAL INVESTIGATOR(S)	Simon Rutt
STUDY PLAN AUTHOR(S)	Simon Rutt, Matt Walker, Connie Rennie, Jose Liht and Jennie Harland
STUDY DESIGN	Matched comparison group quasi-experimental design
NUMBER OF SCHOOLS	1400
NUMBER OF TEACHERS	Around 3800 early career teachers
PRIMARY OUTCOME MEASURE AND SOURCE	• Retention rate of early career teachers (ECTs) within the teaching profession (School Workforce Census)
SECONDARY OUTCOME MEASURE AND SOURCE	• Retention rate of ECTs within their original schools (SWC); • Impact of the ERO on further secondary outcomes including ECTs': self-efficacy; teaching quality; satisfaction with teaching; intentions to remain in teaching (all drawn from survey data)

Study Plan version history

VERSION	DATE	REASON FOR REVISION
1.3 [latest]	27.09.2023	• We originally planned to carry out a Mahalanobis matching procedure to check the sensitivity to matching technique. However, this procedure failed to produce matches to the intervention participants. Consequently, it was removed from both primary and secondary analyses.

1.2	26.05.2023	• Clarified that the same school types would be included in the primary and secondary analysis • Clarified that school turnover will not be used as a covariate for matching for the secondary outcome • Clarified that the primary analysis will take into account both the ECTs who completed the two year period of the intervention and those who opted out during the first year • Clarified there was no plan to run multi-level models on secondary outcomes • Updated timetable
1.1	18.02.2022	• Results from the process of matching ECTs to their SWC records added as Appendix B • Clarified there was no plan to proceed with optional ECT SMS survey • Updated providers' planned hours of delivery • Provided additional detail regarding measurement of teaching quality, self-efficacy and satisfaction • Updated timetable
1.0 [**original**]	17.05.2021	*[leave blank for the original version]*

Table of contents

Background

The challenge of ensuring there are sufficient high-quality teachers employed in schools in England is well documented (e.g., Sibieta, 2018; Sibieta, 2020; Worth *et al.*, 2018; Worth, 2020; DfE, 2019c). Since 2007, teacher numbers have not kept pace with the numbers of pupils in schools (Sibieta, 2020). Over the past decade insufficient numbers of new teachers have joined the profession to meet the demand of increasing pupil numbers (particularly in secondary schools from 2015) and an increasing proportion have left. In 2017, the rate of teachers leaving the profession matched the rate at which teachers entered the profession (9.9 percent respectively) (DfE, 2018).

There are some signs in the last few years that retention of teachers has been improving as fewer teachers left the profession overall in 2018 (9.8 percent (DfE, 2019b) and 2019 (9.2 per cent) (DfE, 2020)), as well as for both primary and secondary teachers (Worth, 2020). However, while recruitment targets for primary teachers were met in 2019-20, recruitment of secondary teachers was below target and was insufficient to meet demand (Worth, 2020). Teacher recruitment numbers for 2021 appear considerably healthier due to Covid-19 and the resulting recession, although there is a concern that these effects may be short-lived (Worth and Faulkner-Ellis, 2021).

Retaining teachers early in their careers is a particular challenge. Teachers in their newly qualified teacher year, and first and second year after qualifying are more likely to leave the profession then at any other time in their career (Worth, 2020). This is the case for teachers across all secondary subjects and for primary teachers and has been a consistent pattern for more than 15 years. Around 15 percent of teachers leave the profession one year after qualifying, and around a further seven percent of teachers leave two years after qualifying (DfE, 2020).The rates at which early career teachers leave the profession are higher for science, mathematics, and languages, and coupled with the recent historical shortage of teachers training in these subjects, creates an additional retention and recruitment challenge in secondary schools (Worth and De Lazzari, 2017)

Retaining teachers that join the profession is a crucial element of the government's Teacher Recruitment and Retention Strategy (DfE, 2019c). At the heart of the strategy is the Early Career Framework (ECF). This underpins an entitlement to a fully-funded, two-year package of structured support for all early career teachers (ECTs)[1], including funded time off timetable in the second year of teaching and additional support from school-based mentors.

There are many reasons why ECTs leave teaching. New teachers commonly experience 'practice shock' when faced with the reality of having their own classes, getting to grips with practice, curriculum, assessments, student behaviour, teaching pedagogy, and intensive workload (Walker *et al.*, 2018; Perryman and Calvert, 2019; Sims and Jerrim, 2020). A study by Hobson *et al.* (2012) indicated that this reality shock can reoccur when the support of the induction year suddenly ceases, as many ECTs still need professional learning opportunities to hone their practice and take on

---

[1] In this study plan we define an ECT as: 'a newly qualified teacher in their first or second year of induction'. Specifically for this evaluation, we mean ECTs who begin their induction in September 2020.

increasing responsibility (Walker *et al.*, 2018). In a recent survey of newly qualified teachers (NQTs), almost three in ten (28 per cent) reported that, after qualifying, workload was greater than they had expected, and around half did not feel supported by their school to manage their workload (Ginnis *et al.*, 2018). ECTs also report working slightly more hours in a typical week than more experienced teachers (Walker *et al.*, 2019). Thus, unsurprisingly, NQTs value being allocated additional time for planning and preparation (Ginnis *et al.,* 2018).

Another reason teachers leave the profession early is due to a lack of collegiality and support – which can leave them feeling isolated and dissatisfied (Buchanan *et al.*, 2013). Conversely, where this support is in place - through mentoring, constructive feedback on observations and informal support from colleagues - this is likely to boost morale and satisfaction, and bolster retention (Walker *et al.*, 2018; Fletcher-Wood and Zuccollo, 2020; Wolstenholme *et al.,* 2012).

The early roll-out (ERO) of the ECF provides a valuable opportunity to evaluate the success with which the induction reforms address some of the key drivers that lead to ECTs leaving the profession.

Study rationale

We agreed with the Education Endowment Foundation's assessment, outlined in the invitation to tender document, that randomisation of schools/ECTs in ERO areas was unfeasible. The main reasons for this included:

- the limited capacity of delivery organisations to recruit the additional schools required for a trial

- the challenge of recruiting schools to a trial given they need certainty regarding their induction arrangements

- the timeframe for commissioning the evaluation, relative to delivery of the ERO, did not provide sufficient time to set up a randomised controlled trial (RCT).

Given the constraints outlined above, a quasi-experimental design (QED) was the preferred approach.

Our design involves:

- measuring the retention rate of ECTs in the profession (primary outcome), and the retention rate within the school in which they began their induction (secondary outcome) using data from the School Workforce Census (SWC)

- using SWC data to construct two comparison groups of schools:
  ➢ to analyse counterfactual retention outcomes in the SWC
  ➢ to sample teachers for gathering counterfactual survey data on further secondary outcomes

- undertaking a 'light-touch' IPE, involving questions in online surveys and telephone/video interviews with ECTs and induction leads to explore the effectiveness of delivery and additional perceptional data on secondary outcomes.
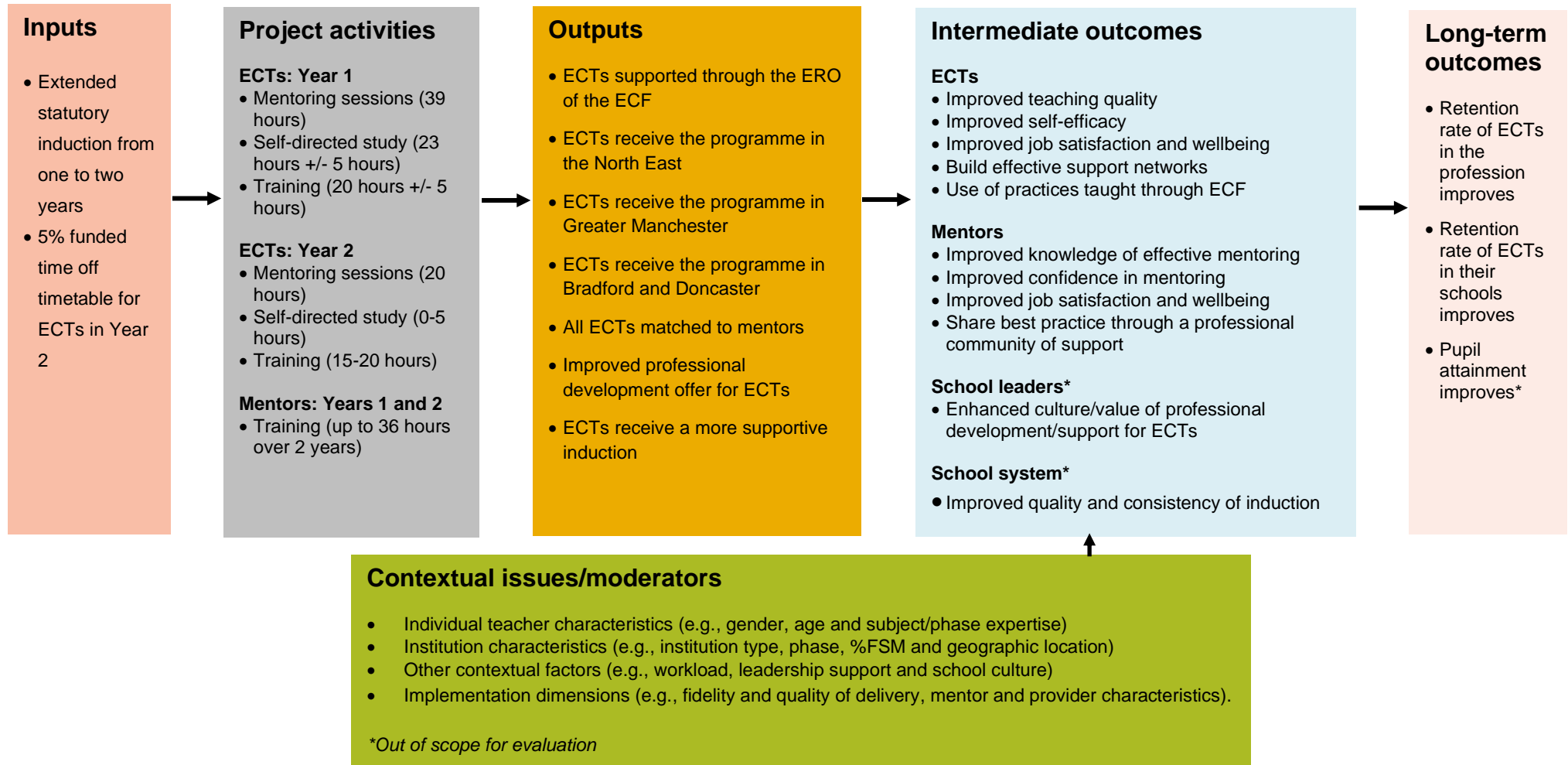
Intervention

Theory of Change

The Theory of Change diagram overleaf provides an overview of the programme. Created by the evaluation team and informed by the Department for Education (DfE), it outlines the main changes the early roll-out of the ECF seeks to make, and the steps that are expected to be involved in making those changes happen. A full description of the programme is provided in the template for intervention description and replication (TiDieR) framework below.

Figure 1 – Evaluation logic model: early roll-out of the Early Career Framework

Rationale and Evidence

Historically, teachers have been most likely to leave the profession within their first two years of service (e.g., Worth, 2020), with workload and stress being major drivers (e.g., Sims, 2017). Early career teachers (ECTs) may not always receive effective support and development opportunities (e.g., DfE, 2019). The government is funding an entitlement for all primary and secondary ECTs (regardless of specialism) to access high quality professional development at the start of their career, which will be underpinned by the Early Career Framework which sets out what new teachers need to learn.

Inputs
- Extended statutory induction from one to two years
- 5% funded time off timetable for ECTs in Year 2

Project activities

ECTs: Year 1
- Mentoring sessions (39 hours)
- Self-directed study (23 hours +/- 5 hours)
- Training (20 hours +/- 5 hours)

ECTs: Year 2
- Mentoring sessions (20 hours)
- Self-directed study (0-5 hours)
- Training (15-20 hours)

Mentors: Years 1 and 2
- Training (up to 36 hours over 2 years)

Outputs
- ECTs supported through the ERO of the ECF
- ECTs receive the programme in the North East
- ECTs receive the programme in Greater Manchester
- ECTs receive the programme in Bradford and Doncaster
- All ECTs matched to mentors
- Improved professional development offer for ECTs
- ECTs receive a more supportive induction

Intermediate outcomes

ECTs
- Improved teaching quality
- Improved self-efficacy
- Improved job satisfaction and wellbeing
- Build effective support networks
- Use of practices taught through ECF

Mentors
- Improved knowledge of effective mentoring
- Improved confidence in mentoring
- Improved job satisfaction and wellbeing
- Share best practice through a professional community of support

School leaders*
- Enhanced culture/value of professional development/support for ECTs

School system*
- Improved quality and consistency of induction

Long-term outcomes
- Retention rate of ECTs in the profession improves
- Retention rate of ECTs in their schools improves
- Pupil attainment improves*

Contextual issues/moderators

- Individual teacher characteristics (e.g., gender, age and subject/phase expertise)
- Institution characteristics (e.g., institution type, phase, %FSM and geographic location)
- Other contextual factors (e.g., workload, leadership support and school culture)
- Implementation dimensions (e.g., fidelity and quality of delivery, mentor and provider characteristics).

*Out of scope for evaluation

4

TIDieR Framework

Brief name

Evaluation of the early roll-out of the Early Career Framework

Why (rationale/theory)

The DfE's ECF sets out what newly qualified teachers in their first two years of teaching need to learn. From September 2021, the government is funding an entitlement for all early career teachers (ECTs[2]) to access high quality professional development at the start of their career, which will be underpinned by the ECF. New teachers will receive development materials, mentoring, and training. Induction will be offered over two years, rather than one year, with additional time away from the classroom made available to ECTs in their second year of teaching. This package is in addition to the DfE statutory induction guidance (2018) which sets out what schools (and other relevant bodies) should provide for newly qualified teachers. All newly qualified teachers must satisfactorily complete an induction period in order to be employed as a teacher in England. DfE statutory induction guidance was updated ahead of the national roll-out[3], to reflect the ECF and new two-year statutory induction.

The first few years of a teacher's career are critical years when the right development opportunities, nurture and support can make or break a sustained future career. ECTs were a major focus in DfE's recruitment and retention strategy and continue to be a policy focus. The ECF has arisen from a desire to improve the wellbeing, job satisfaction and quality of teaching of ECTs – equipping them with the knowledge, skills, practices and support networks to cope with the challenging learning curve to becoming fulfilled and effective teachers. Ultimately, the aim is that this helps to retain more ECTs in the profession.

Historically, teachers have been most likely to leave the profession within their first two years of service. The net year-on-year reduction in the proportion of the NQT cohort that are still in service in the state sector has risen over much of the last decade, and peaked at 15 per cent in 2018/19 (Worth, 2020). In 2019, there were 23,064 full time equivalent newly qualified teachers entering the profession (DfE, 2020). Workload and stress are major drivers for teachers leaving the profession (e.g. Sims, 2017). The demands of marking pupils work, planning lessons, and dealing with pupil misbehaviour are particularly acute for new teachers who are still building their experience and developing effective practices and strategies (e.g., Higton *et al.*, 2017 cited in DfE, 2019a; Barmby, 2006).

The ECF was designed by the DfE in consultation with an expert advisory group. The content of the framework and its underpinning evidence was independently assessed by the EEF and it has been endorsed by a range of other sector bodies. The framework

---

[2] For the purpose of this project, an ECT is regarded as 'a newly qualified teacher in their first or second year of induction'. Specifically for this evaluation, we mean ECTs who will begin their induction in September 2020.

[3] Induction for early career teachers (England). Revised March 2021. To come into force on 1 September 2021:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/972316/Statutory_Induction_Guidance_2021_final__002___1__1_.pdf [14/05/21]

draws together evidence of effective practices in relation to: behaviour management; pedagogy; curriculum; assessment; and professional behaviours. The ECF sets out two types of content. This includes key evidence statements (i.e. 'learn that…'), which set out what new teachers should learn, as well as practice statements (i.e. 'learn how to…'), which provide guidance on the skills that ECTs should be supported to develop. In order to equip teachers with the skills to effectively and efficiently fulfil the core aspects of high quality teaching, the ECF also covers the following areas:

- High expectations (e.g., EEF Toolkit, 2018; Institute of Education Sciences 2008; PISA, 2015).
- How pupils learn (e.g., Deans for Impact, 2015; EEF, 2018; Rosenshine, 2012).
- Subject and curriculum (e.g., Coe *et al.*, 2014; Shanahan, 2005).
- Classroom practice (e.g., Coe *et al.,* 2014; EEF Toolkit, 2018; Rosenshine, 2012).
- Adaptive teaching (e.g., Davis *et al.*, 2004; OECD, 2015).
- Assessment (e.g., Black *et al.*, 2004; Coe, 2013; EEF, 2016).
- Managing behaviour (e.g., Carroll *et al.*, 2017; Coe, 2014; Institute of Education Sciences, 2008).
- Professional behaviours (e.g., Carroll *et al.*, 2017; Cordingley *et al.*, 2015; EEF, 2015).

Evidence suggests that ECTs may not always receive sufficiently effective support and development opportunities during the early years of teaching when there is most to learn to become an established and confident teacher (e.g., DfE, 2019a), leaving some ECTs feeling unsupported and isolated (e.g., Ginnis *et al.*, 2018; Buchanan *et al.*, 2013). A Government consultation (DfE, 2018) revealed the need to strengthen the support for ECTs to improve morale, satisfaction and retention by better equipping them in the early stages of their career. Consultees advocated the value of high quality mentoring in particular. The benefits of professional learning opportunities through mentoring, opportunities to practise learning and receive constructive feedback, seeing effective practice being modelled, and informal support from colleagues have been widely established in literature (e.g., Rosenshine, 2012; Deans for Impact, 2016; Walker *et al.*, 2018; Fletcher-Wood and Zuccollo, 2020; Wolstenholme *et al.*, 2012).

The National Standards for school-based initial teacher training (ITT) mentors (Teaching Schools Council, 2016) were commissioned in response to the Carter review (Carter, 2015) that recommended mentoring needed more prominence in teacher training. The Standards outline the need for additional mentoring expertise in deconstructing and articulating practice, and coaching, as well as additional mentoring capacity beyond training and accreditation into teachers' early careers. The ECF is underpinned by these standards and includes a commitment to improve the availability and quality of mentoring through mentor training, mentoring materials and access to a professional community to share best practice. Through this support, mentors will develop their: understanding of the ECF in order to support ECTs in understanding and implementing the evidence-based practices of the ECF; and approaches to tailoring support to individual mentee needs and experiences.

The ECF induction programme has also been developed to comply with the Standard for teachers' professional development (DfE, 2016) in focusing on improving pupil outcomes, being underpinned by evidence and expertise; including opportunities for

collaboration and expert challenge; being sustained over time; and supported by school leadership.

The government took a phased approach to introducing the ECF, starting with an early roll-out (ERO) from autumn 2020 in the North East, Greater Manchester, Bradford and Doncaster. The ERO was designed to provide learning to inform the national roll-out from September 2021. The ERO areas were identified by the DfE as representing a diverse range of schools in varying circumstances. As such, they were viewed by the DfE as a suitable sample from which to help inform the development of support for all schools nationally to implement the ECF. The ERO was expected to support up to 2,500 ECTs across these three areas – representing approximately 80 per cent of the total number of NQTs expected to be recruited by schools in these areas.

The ERO will offer the Full Induction Programme (FIP) to participating schools – which includes:

- two years of new, government-funded, training
- freely available development materials based on the ECF
- additional government funding for five per cent time away from the classroom for teachers in their second year
- a dedicated mentor and new training for these mentors
- government funding to cover mentors' time with the mentee in the second year of teaching.

National roll-out of the ECF will provide schools with the option of the Core Induction Programme (CIP) materials, supplemented by the school's own training, or the FIP which will include external training, or some form of combination.

Who (recipients)

The ERO of the ECF is being delivered to all schools in which an NQT can undertake statutory induction in the North East, Greater Manchester, Bradford and Doncaster. Eligible local authorities are: Bolton, Bradford, Bury, Darlington, Doncaster, Durham, Gateshead, Hartlepool, Manchester City Council, Middlesbrough, Newcastle, North Tyneside, Northumberland, Oldham, Redcar and, Cleveland, Rochdale, Salford, South Tyneside, Stockport, Stockton-on-Tees, Sunderland, Tameside, Trafford and Wigan. In each participating school it is anticipated that all NQTs will take part in the ERO. Each NQT must be assigned a mentor from their school. The mentor will also need to undertake training (see further details below).

What (materials)

The DfE commissioned four expert teacher training providers to: i) develop a standalone core induction programme (CIP)[4], with materials that schools can draw on to deliver their own ECT and mentor training, and; ii) deliver the full induction programme (FIP)[5], which is the full provider-led programme offering training for early

---

[4] Self-directed study materials; mentor session materials and training session outlines only. These are available online: https://www.early-career-framework.education.gov.uk/ [14/12/2020]

[5] Self-directed study materials; mentor session materials and training for ECTs and training for mentors delivered by an external organisation.

ECTs and their mentors alongside the professional development materials provided as part of the CIP.

The FIP includes:

- a week-by-week sequence for covering the content of the ECF through mentoring sessions, self-directed study, and training over the two-year induction period
- self-directed study materials covering evidence-based practices in relation to each of the ECF statements covering five themes: behavior management; pedagogy; curriculum; assessment and professional behaviours
- mentor session materials aligned to the ECF statements that structure the mentor-mentee meetings with clear outcomes specified
- ECT training session outlines, with specified outcomes and content.

The materials must be applicable to teachers in all phases and subject areas, however the providers can decide how to achieve this, for instance with content that is universally applicable or by providing sub-sets of materials applicable to specific phases and subject specialisms. Providers can also decide the type of study materials to include, for instance, exemplification of practices in video clips, research summaries and reading materials. Providers also design the range of materials and tools for mentoring sessions. In order for the materials to be suitable to all ECTs, regardless of their prior experience and route into teaching, the materials should be appropriately pitched in terms of the level of information and challenge they provide.

What (procedures/activities/processes)

The ECF statements are designed to align to the Teachers' Standards. The FIP should only cover content referred to in the ECF. The ECF induction programme is delivered through the following activities:

- Mentor sessions: including a range of different types of materials and tools and these should adhere to the National Standards for school-based initial teacher training mentors.
- Self-directed study: covers all of the ECF statements only and is evidence-informed. The content is scheduled into a weekly sequence that links together the different activities.
- ECT training: this includes opportunities for networking and observing good practice
- Mentor training: this includes opportunities for networking.

The ECF is not, and should not be used, as an assessment framework. ECTs will continue to be assessed against the Teachers' Standards only. The ECF will underpin an entitlement to training and support for ECTs and should not be seen as an additional assessment tool.

The ECF provides guidance on mentor selection, stipulating that mentors should be suitably experienced with a minimum of two years teaching experience, hold Qualified Teacher Status (QTS) and be a highly regarded 'excellent' teacher in the subject/age range they are mentoring. They should also have professional competencies that align with the National Standards for School-based Initial Teacher Training (ITT) Mentors.

Mentors will have a formal responsibility to work collaboratively with the ECT and other school colleagues to support the ECT in receiving a high quality induction programme.

Who (providers/implementers)

DfE commissioned four providers to deliver the ERO: Ambition Institute; Education Development Trust; Teach First; and a consortium led by UCL. These providers are working with networks of local organisations and partners in the ERO areas to deliver the support, including Teaching School Alliances, Academy Trusts, Universities and Teaching School Partnerships. It is a limitation of this study that the research team did not speak to these organisations directly, and therefore they were out of scope for the evaluation.

In schools, the implementation of the ECF induction programme will be supported by each school's induction lead or coordinator. The induction lead will be expected to appoint suitable mentors, schedule opportunities for the mentor and ECT to meet, and integrate the ECF induction programme with the school's own induction processes, as well as to fulfil statutory induction procedures. In-school mentors will be expected to participate in mentor training, engage with study materials and provide mentoring sessions for their assigned ECT.

How (mode of delivery)

The ECF induction programme is delivered through:

- self-directed study
- mentoring sessions
- ECT training
- mentor training
- funding for five per cent time away from the classroom for teachers in their second year
- funding to cover mentors' time with the mentee in the second year of teaching.

In addition, as part of a related but separate study, an RCT is being undertaken to evaluate the impact of an incentive payment on mentors' engagement with the ERO. The incentive payment is in addition to the funding received by participating schools at the start of the second year of the programme, and will be administered to some of the schools being supported by Ambition Institute and Teach First[6].

Where (location of intervention)

Training is to be delivered through a combination of face-to-face[7] training at local venues, live online remote sessions and recorded webinars, visits to observe good practice in the ECT's own school or a neighbouring school, local networking and peer-to-peer sessions. Mentoring will take place in the ECT's and mentor's school. ECTs and mentors may engage with self-study materials in school or at home, as convenient. Implementation of the ECF strategies will take place in classrooms in

---

[6] For further information, the protocol for this study is available online:
https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Protocols/EEF_school_mentor_incentive_trial_protocol_final.pdf [09/02/21]

[7] In light of social distancing due to the Covid-19 pandemic, planned face-to-face training may be substituted with increased online training.

participating schools. The ERO of the ECF is being implemented in schools in the North East, Greater Manchester, Bradford and Doncaster. In addition, approximately 4,500 new teachers outside of these areas signed up to a one-year funded offer of early career framework-based support. This expansion was designed to provide enhanced support to newly qualified teachers whose Initial Teacher Training (ITT) may have been disrupted due to school closures and social distancing as a result of the Covid-19 pandemic. However, schools participating in this expanded offer do not form part of the evaluation.

When and how much (duration and dosage)

There is an expectation that the induction programme will be delivered entirely within the ECT's timetable reduction – which is 10 per cent in the first year of induction, and five per cent in the second year. The programme should provide sufficient materials and content for the hours available. The maximum total hours of the programme for the ECTs (including self-directed study, mentoring sessions and training) must not exceed 131 hours over the two years. This includes time for three formal assessments and 12 professional progress reviews which must take place as part of the ECT's progression to meeting the Teachers' Standards[8]. Table 1 below provides an overview of the number of hours of support that are expected to be given to ECTs and mentors in each of the two years of the programme.

Table 1: Overview of the duration and dosage of each of the main strands of the ECF induction programme

Session type	Year 1	Year 2
Mentor Sessions	39 hours (equivalent to a weekly one-hour meeting)	20 hours (equivalent to a fortnightly one-hour meeting)
Self-directed study	23 hours (+/- 5 hours)	0-5 hours
ECT training	20 (+/- 5 hours)	15-20 hours
Formal Assessment (school-led activity; 3 meetings over two years)	2 hours	1 hour
Professional progress reviews (school-led activity; 12 meetings over two years)	3 hours	3 hours
Total	**87 hours**	**44 hours**

The schedule for delivering mentor training and support is more flexible, although the maximum total hours of training for mentors should not exceed 36 hours over the two years.

---

[8] The providers are not required to create any materials for the formal assessment and progress reviews as these are conducted by the school against the Teachers' Standards.

The total time dedicated to self-directed study and ECT training must not exceed 43 hours in total in Year 1. The total time dedicated to self-directed study and ECT training must not exceed 20 hours in total in Year 2.

Tailoring (adaptation)

As can be seen from Table 1 above, providers have some limited discretion in how they allocate the time for ECT support across Year 1 and 2 for self-directed study and for ECT training by +/- 5 hours. This must not exceed the overall limit. The number of hours of mentor training is stipulated over the two years, but providers have discretion over how to distribute this over the two years, as well as the length and frequency of sessions. A summary of providers' planned programme delivery can be found in Tables 2 and 3.

Providers may also choose to vary the mode of delivery for the training, which could include face-to-face training, online training, school/lesson visits and networking and peer-to-peer development opportunities. For instance, several providers offered peer-to-peer support for mentors. Two providers stipulated school/lesson visits as part of their provision. All four providers indicated that at least some of the training would be grouped by subject or phase specialism. All providers developed online platforms and planned for blended learning involving both face-to-face and online support.

Providers have discretion to sequence the content of their programmes as they deem appropriate, although this should be evidence-informed and should start by covering the foundational knowledge most needed in the early stages of teaching, before covering more complex content. All four providers planned to sequence their content to cover all areas of the ECF in Year 1, with the focus in Year 2 being more about revisiting, implementing and embedding the learning in some way. Providers planned to take slightly different approaches to achieve this in the second year of the programme. For instance, one chose to focus on practitioner enquiry; another focused on shadowing and optional career development modules. All the providers have broken down their sequences into weekly, step-by-step topics with a specific focus for each week which links the self-directed study, mentoring and training. Furthermore, there may be some scope for tailoring of the delivery of the providers' programmes at the local level as the training is delivered by local delivery partners who have some discretion to adapt the materials based on their expertise and knowledge of local schools and contexts.

The providers' programmes also vary to some degree in the nature of their engagement with school induction leads. For example, two providers offered a direct initial training session for induction leads, while two provided digital training and guidance materials only, and one of these also planned to offer support to induction leads via networks with local partners.

Finally, some of the providers offered distinctive additional aspects to their programmes, such as free access to the IRIS connect lesson video platform and free membership to the Chartered College of Teaching.

Table 2: Overview of planned provider delivery in Year 1

Type of support	DfE guidance on expected number of hours	Planned provider activity			
		Ambition Institute	Education Development Trust	Teach First	UCL Early Career Teacher Consortium
Mentoring sessions	**39 hours (one hour weekly)**	39	39	39	39
Self-directed study	**23 hours (+/- 5 hours)**	24 hours (36 modules, 40 minutes each)	26	27 (4.5 hours per half-term)	22 (22 sessions of 45 mins learning time and 15 mins preparation for mentor meeting)
ECT training	**20 hours (+/- 5 hours)**	19.5 (conferences and clinics)	17	24	21
Mentor training	**Maximum of 36 hours over two years**	26	21	9 (+7 hours mentor induction day in September)	18

Table 3: Overview of planned provider delivery in Year 2

Type of support	DfE guidance on expected number of hours	Planned provider activity			
		Ambition Institute	Education Development Trust	Teach First	UCL Early Career Teacher Consortium
Mentoring sessions	20 hours (one hour fortnightly)	20	20	20	20
Self-directed study	0-5 hours	5	5	5	5
ECT training	15-20 hours	10.5	15	24	15
Mentor training	Maximum of 36 hours over two years	5	15	8.5 (+7 hours mentor induction day in September)	18

Planned strategies to maximise implementation effectiveness include the flexibility given to schools and mentors in how they decide to best support ECTs, and align the programme to existing induction processes. In addition, programme modules were sequenced in order to address key priorities for ECTs and their schools across a typical school year, with providers recommending that this sequence is followed where possible. However, schools were given the opportunity to change the duration or sequencing of programme modules, for example where exceptional and unplanned staffing pressures restrict the availability of mentors to run the programme, or where an ECT required additional support.

Impact evaluation

Research questions

The primary research question is:

1   What is the **impact of the ERO on the retention rate of ECTs in the profession**?

The secondary research questions are:

2   Does the retention rate of ECTs in the profession vary depending on whether they completed one or two years of the intervention?

3   What is the **impact of the ERO on the retention rate of ECTs within their original schools**?

4   What is the **impact of the ERO on further secondary outcomes** including ECTs':

> ➢ self-efficacy

> ➢ teaching quality

> ➢ satisfaction with teaching

> ➢ intentions to remain in teaching?

5   How does the impact vary by compliance (duration)?

6   What are the moderating factors of ERO?

> ➢ Phase

> ➢ Provider

> ➢ School level FSM

> ➢ Workload

Design overview

Table 4: Design

Design		Matched comparison group quasi-experimental design
Unit of analysis		Teacher level
Number of Units to be included in analysis (Intervention, Comparison)[9]		Around 3800 (1900, 1900)
Primary outcome	variable	Teacher Retention
	measure (instrument, scale, source)	Presence in the SWC in November 2022
Secondary outcome(s)	variable(s)	Teacher retention in the same school Self Efficacy Teaching Quality Teaching Satisfaction Intentions to remain in teaching
	measure(s) (instrument, scale, source)	Presence in the same school in the SWC in November 2022 Surveys of ECTs and induction leads

Participants

Treatment participants

The ERO was offered to all schools in which an ECT could undertake statutory induction within the eligible areas, these being the North East, Bradford, Doncaster, and Greater Manchester. Therefore all treatment *schools* will be those in the above areas, who have employed an ECT starting in September 2020, and have agreed to take part in the ERO. The treatment *participants* will be those ECTs in the participating schools who are recipients of the ECF ERO in the years 2020/21 to 2021/22, and who are successfully matched to the SWC records. The list of these ECTs and the schools they work in, is being collected from the four providers (Ambition Institute, Education Development Trust, Teach First, and UCL) by the DfE.

---

[9] Depending on the method used, the number of units included in the analysis can differ from the pool of potential comparison units. For example, when using matching/weighting the pool of comparisons units could represent all schools in England, but only a certain number of units will be included in the analysis after a suitable match is found. Identifying the precise number of units included might not be possible at the design stage. In these cases Evaluators can speculate on the number of units that are expected depending on the method used.

<u>Inclusion and exclusion criteria</u>

In discussion with the EEF, it was decided that participants based in selected institutions would be removed from the treatment and comparison groups and subsequent analysis. Alternative provision settings and post-16 institutions, including school sixth forms and further education colleges, were removed because they were very small in number, which would have made it difficult to accurately estimate impacts on teacher retention. Special schools were also removed as there were practical constraints with regards to the suitability of the research instruments that had been developed as part of the survey work. The school types that were included were:

- Academy converter
- Academy sponsor led
- Community school
- Foundation school
- Free schools
- Local authority nursery school
- Studio schools
- Voluntary aided school
- Voluntary controlled school.

<u>Comparison participants</u>

Comparison schools will be schools that have employed an ECT starting in September 2020, but are not in the areas of the ERO, and have not signed up to take part in the expanded offer of the ECF support[10]. Comparison participants will be the ECTs who are not recipients of the ECF programme, and are teaching in areas outside of those where the ECF was rolled out in the ERO. Therefore teachers in schools in the areas of the ERO that declined to take part in the ERO are excluded. Using statistical matching, the comparison teachers will be similar in characteristics, and will be working in similar schools as those in the treatment group. Two comparison samples will be drawn, the first for the purpose of administering ECT and induction lead surveys, and the second for the impact evaluation. The two comparison samples are described in detail in the sections below.

---

[10] https://www.gov.uk/government/publications/early-career-framework-reforms-overview/early-career-framework-reforms-overview

Initial provider data supplied indicates that there are approximately 1900 ECTs who are taking part in the ERO. Due to potential data quality issues, a potential lag in up to date data for ECTs, and previous matching experience, (e.g., 2019 Teacher Workload Survey), we use the assumption that 20 per cent of the ECT data will not match to the SWC. This would leave a sample of 1520 ECTs with SWC records, and we conservatively assume an equally-sized comparison group of ECTs.

We also conservatively assume that there is zero correlation between covariates and the outcome variable. Regression models of teacher retention tend to be able to explain some variance, but typically have very low explanatory power (see for example, Worth *et al.,* 2017). We assume no intra-cluster correlation: the number of ECTs at a school are typically low anyway, especially for primary schools.

The chart below shows that this design yields sufficient statistical power to detect a retention rate difference of 3.4 percentage points or more. There is limited evidence of impact of programmes on teacher retention. However, quasi-experimental research by Allen and Sims (2017) found that the impact of participating in a science CPD programme was associated with an increase of 3-4 percentage points in department-level rates of science teacher retention in the profession. This suggests that the evaluation design is adequately powered, but some risk of false negatives remains nonetheless.
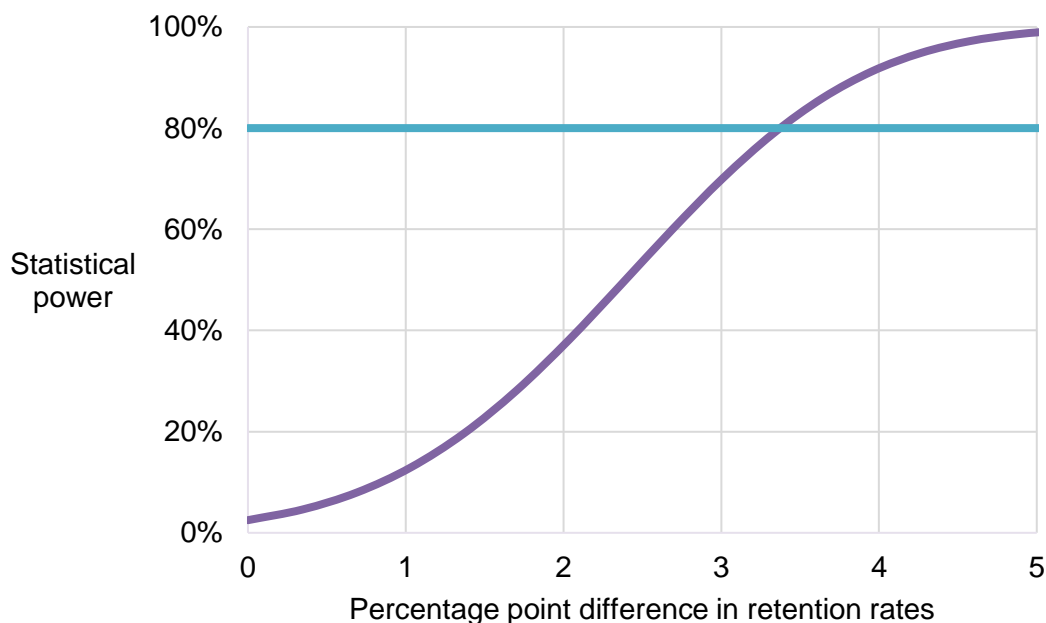
Power - n(intervention)=1527; n(comparison)=1527

Assumptions:
1 ECT per school, ICC = N/A, Correlation with covariates = 0
Confidence level of the test = 95 per cent
3-year teacher retention rate for comparison schools = 74% (from SWC)

Table 5: Sample size calculations

		Study Plan
Minimum Detectable Effect Size (MDES)		3.4 percentage points
Pre-test/ post-test correlations	level 1 (teachers)	0
	level 2 (school)	0
Intracluster correlations (ICCs)	level 1 (teacher)	0
	level 2 (school)	0
Alpha		0.05
Power		0.8
One-sided or two-sided?		Two
Average cluster size		N/A
Number of schools	Intervention	N/A
	comparison	N/A
	Total	N/A
Number of teachers	Intervention	1520
	comparison	1520
	Total	3040

Outcome measures and other data

Table 6: Overview of the outcomes and comparison samples

Outcome	Comparison Sample 1	Comparison Sample 2
Retention in the workforce (SWC)	X	
Retention in the same school (SWC)	X	
Self-efficacy		X
Teaching quality		X
Teacher satisfaction		X
Intentions to remain in teaching		X

Primary outcome

The primary outcome for this investigation is retention in state-funded teaching in England in the year 2022. The ECF is focused on providing additional support to ECTs

in the two years in which they are most likely to leave the profession. Therefore, retention in the profession is the most suitable primary outcome measure. We will measure retention according to whether or not the teacher is present in the SWC in the two years following the start of their participation in the ERO. Specifically, if they are in the SWC in November 2021 and 2022, however the primary outcome is retention after two years (2022). A limitation of this measure is that a teacher who is not present may still be teaching, but outside the state-funded sector in England, e.g., in Scotland, Wales, in further education or in the private sector.

In addition to using the SWC to measure retention in the workforce, the research team considered administering an SMS survey to ECTs, to be conducted in September 2021 and 2022, to measure the same outcomes. An SMS survey would have provided early indicative data. However, these findings would also have been less reliable than the SWC analysis, due to the reduced sample size and the fact that this would not have produced census-level data. For these reasons, it was agreed that an SMS survey would not be administered.

Secondary outcomes
The secondary outcome is retention in the original induction school in the years following the start of their participation in ECF. Specifically, if they are in the SWC and in the same school in November 2021 and 2022. .We will take two approaches to compare the proportion of ECTs who stayed in the same school after one and two years across intervention and comparison. First, the proportion of those staying in the same school will be calculated out of all teachers in the sample (including those who left the profession). Second, we will calculate the proportion out of those who remained in the maintained sector.

Teaching quality

The research team viewed the constructs of 'teaching quality' and 'ECT self-efficacy' as overlapping to some degree. Both measures seek to explore the effectiveness with which ECTs handle the tasks, obligations, and challenges related to their professional activity, with self-efficacy exploring the self-reported confidence with which ECTs engage in these activities. Teacher quality will be assessed primarily using items in the surveys of ECTs and induction leads that will be measured at two time points across years one and two: June-July 2021 (intervention group) and September-November 2021 (comparison group) and June-July 2022 (both groups). This will provide a midpoint and endpoint perspective respectively. It was intended that the midpoint comparison surveys would be administered at the same time as the intervention group surveys, but, due to a delay in receiving the school workforce census (SWC) data needed to draw the comparison group, the comparison group surveys were administered slightly later. This necessitated some small wording changes to some of the questions posed to respondents in the comparison group.

In the absence of recognised high-quality measures of teaching quality, the research team drew on measures of teaching quality drawn from NFER's evaluation of the Teaching and Leadership Innovation Fund (TLIF), commissioned by the DfE. The questions from the TLIF evaluation surveys were themselves drawn and adapted from the OECD's Teaching and Learning International Survey (TALIS). To allow for comparisons to be made between the responses of ECTs and induction leads, where possible the same or similar items were used in both surveys. The resulting combination of items measured a range of metrics associated with teaching quality, including ECTs' subject and pedagogical knowledge and behaviour management skills, see Table 7 below.

Table 7: Survey questions used to measure teaching quality

	ECT surveys	Induction lead surveys
Question number	Q14	Q20
Themes covered	Subject and pedagogical skills, views on teaching quality, and use of assessment strategies and research findings to inform practice.	Covers a range of measures, including ECTs' subject and pedagogical knowledge, adaptive teaching, use of assessment strategies and behaviour management.

N.B. The question numbers relate to the Year 1 surveys.

For ECT question 14 and induction lead question 20, respondents were asked to answer a series of statements on an eight-point scale, from 1, 'strongly disagree, to 8, 'strongly agree'. Alternatively, they could select a single response to indicate they did not know or were not sure. The teaching quality measures will be reported as two mean scores, one for each survey instrument, derived by creating an average across all sub-items, with a higher score indicating higher teaching quality.

Self-efficacy

Self-efficacy was also measured using items included in the surveys of intervention and comparison ECTs. In order to measure ECTs' self-efficacy, we considered including within the surveys the full scales from the short form of the Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran & Woolfolk Hoy, 2001) and the Self-Efficacy in relation to the Early Career Framework (SECF) scale, developed by Hardman *et al.,* 2020. However, we felt there was duplication across scales, and that it would be too burdensome to ask respondents to complete both scales. Instead, we opted to include the full SECF scale due to its closer alignment to the ECF, and the four items relating to the 'Efficacy in Classroom Management' factor from the TSES. This was because we know from the research literature and from our own research that many ECTs struggle with behaviour management (Walker *et al.,* 2018), and we felt there were only limited measures on this in the SECF. For both scales ECTs were asked to rate a series of statements on a scale from 1 to 9. For the TSES these ranged from 1, 'nothing', to 9, 'a great deal', and for the SECF they ranged from 1, 'not at all confident', to 9, 'very confident'.

Table 8: Survey questions used to measure self-efficacy

	ECT surveys		Induction lead surveys
Question number	Q15	Q16	N/A (can only be captured from ECTs themselves)
Themes covered	Draws on four items from the TSES (short form) relating to 'Efficacy in Classroom Management'	Draws on all 16 items from the SECF.	

N.B. The question numbers relate to the Year 1 surveys.

Both scales will be reported as a single mean score derived by creating an average across individual items, with higher scores indicating higher self-efficacy.

Teacher satisfaction

ECTs' and induction leads' satisfaction with teaching, and with the materials and support used to deliver the ECF (intervention group participants only), were measured using a combination of questions/items included within the ECT and induction lead surveys. A summary of the range of questions and the themes they covered can be found in the table below.

Table 9: Survey questions used to measure teacher satisfaction

	ECT surveys		Induction lead surveys
Question number	Q17	Q26 (intervention group only)	Q18 (intervention group only)
Themes covered	Satisfaction with job and enjoyment of becoming a teacher	Overall satisfaction with the ECF programme	Satisfaction with the support the ERO

	ECT surveys		Induction lead surveys
	and working in current school		provider has given the school

N.B. The question numbers relate to the Year 1 surveys.

For ECT question 17, respondents were asked to answer a series of statements on an eight-point scale, from 1, 'strongly disagree, to 8, 'strongly agree'. Alternatively, they could select a single response to indicate they did not know or were not sure. For ECT question 26 and induction lead question 18, respondents were asked to respond to a single statement on a five point scale, from 1, 'very dissatisfied', to 5, 'very satisfied'. The findings from all three questions will be reported as descriptive frequencies.

<u>Intentions to remain in teaching</u>

As part of the surveys, intervention and comparison group ECTs were asked how strongly they agreed with the following statements (question 19): 'I plan to stay in the teaching profession for at least the next three years', and 'I plan to stay teaching at this school for at least the next three years'. Respondents were asked to answer on an eight-point scale, from 1, 'strongly disagree, to 8, 'strongly agree'. Alternatively, they could select a single response to indicate they did not know or were not sure. The findings from question 19 will be reported as descriptive frequencies.

Other data

<u>Workload</u>

Drawing on items from the 2019 Teacher Workload Survey (Walker *et al., 2019*), intervention and comparison group ECTs were asked how strongly they agreed with the following statements (question 18): 'I can complete my assigned workload during my contracted working hours', 'I can complete all my induction-related activities in my allocated 10% timetable reduction', and 'I have an acceptable workload'. Respondents were asked to answer on a five-point scale, from 1, 'strongly disagree, to 5, 'strongly agree'. Using the same scale, induction leads were asked how strongly they agreed with the statement: 'I have an acceptable workload'. The findings from question 18 will be reported as descriptive frequencies. Workload will also be analysed as a moderating factor (see subgroup analysis section).

Dosage data

Monitoring data will be collected by each of the providers. The data will relate to how many hours the ECTs spent partaking in different elements of the programme across Year 1 and Year 2.

The elements and the expected hours per year spent on each element are presented in Table 1.

All four ERO providers have explained that there will be limitations to the completeness /accuracy of the monitoring data they will collect, while one provider has informed us they will not be collecting any dosage data. Another provider has informed us that they are not collecting the number of hours ECTs spend in sessions with their mentors. Instead, they will be collecting the number of mentor sessions attended by ECTs and mentors. For these reasons, data regarding the amount of time mentors spend on training, and ECTs spend on training, in ECT-mentor sessions and on self-directed learning, will also be captured through the induction lead and ECT surveys respectively.

It should be noted that although Table 1 shows a fixed number of hours for each of the programme elements, the actual number of hours expected may vary by provider, as shown in Tables 2 and 3. We will use two measures of dosage, one that consists of raw hours spent, regardless of provider specific expectations, and one that consists of raw hours as a proportion of provider-expected hours.

Moderating data

Phase, school level FSM, and provider will be included in additional models to examine any potential differential impact. Phase and FSM data will be derived from Get Information About Schools while provider data will be supplied by DfE.

Selection mechanism

The ERO was offered to all schools in which an NQT can undertake statutory induction in the North East, Bradford, Doncaster, and Greater Manchester. As it is a new programme, we have little information as to which schools will self-select into the programme, and which will decline to take part. As part of the data collection process we will collect information from the schools that did not take part via analysis of administrative data as well as telephone interviews, identifying the reasons for not doing so. This information will be reported descriptively, as well as the characteristics of the schools that did not take part. Characteristics that make it more likely for a school to have an ECT are, however, characteristics that may influence uptake of the programme. These characteristics include school phase, size, and the proportion of teachers with less than two years' experience (a proxy for turnover). We will descriptively explore these characteristics and others available in the SWC and will use the information to inform the matching process. The Covid-19 pandemic may also have reduced schools' ability to take on NQTs. This will be explored through the IPE methods (namely surveys of and telephone interviews with induction leads).

Selection of the comparison group and identification assumptions

We will draw two comparison groups for this evaluation, for:

1 estimating the counterfactual retention rate in the SWC
2 sampling ECTs for the ECT survey.

The priority for the first comparison group is data quality. When the SWC is recorded it sometimes has gaps in it, that disproportionately affect newly hired teachers, due to school data not being up to date. These gaps are retrospectively filled in at a later stage. Therefore the comparison group for the primary outcome will be selected using retrospective 2020 SWC data in July 2022.

The priority for the second comparison group is timeliness. Therefore, the aim was to construct the comparison group using an early version of the pre-publication 2020 SWC in April 2021. This would have meant the first round of the comparison group surveys could be administered at the same time as the intervention group surveys in June 2021. However, as reported above, due to a delay in receiving the SWC data needed to draw the comparison group, the comparison group surveys were administered slightly later, in September 2021. This necessitated some small wording changes to some of the questions posed to respondents in the comparison group. It also means that some ECTs may have already left teaching at the point the comparison group survey was administered, which may have introduced bias into the responses.

MATCHING/WEIGHTING

Primary analysis comparison group

In order to align the sample for the primary and secondary analysis, we will first eliminate the treatment participants that do not teach in one of the school types delineated above (see *Inclusion and exclusion criteria* section above) We will then match the list of treatment participants to their records in the SWC. Using the population of de-identified teacher records in the SWC from geographical areas outside those of the ERO and included school types, propensity score matching will be used to draw a comparison group in a 1:1 treatment to control ratio. The sample will be drawn using observable teacher and school characteristics at baseline that are known to be associated with teacher retention (Burge *et al.,* 2021). Since the SWC does not identify who is an ECT, the pool of teachers from which the comparison is selected will be limited to those that have qualified and been employed for less than a year from the date in which the 2020 census was conducted. The majority of schools in the ERO pilot areas signed up to the ERO. As a result, the decision was made not to select comparison schools from the same geographic area as the ERO, as those that chose not to participate might have been different in some way to those that did.

Teacher characteristics used for matching will include age and gender. School characteristics will include phase, school size, school proportion of Ever FSM pupils, Ofsted rating, proportion of teachers with less than two years' experience (a proxy for turnover) and local labour market conditions (using local area average wages as a proxy).. Any missing data will be imputed by the method of multiple imputation with chained equations (MICE), which involves iteratively fitting models to predict the missing values and is implemented in the mice R package (van Buren & Groothuis-Oudshoorn, 2011).

S. van Buuren and K. Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in R. Journal of Statistical Software, 45(3):1–67, 2011. URL https://doi.org/10.18637/jss.v045.i03.

The reason for a matching process is to ensure that the analysis is run using a comparison group that is as similar as possible in terms of observable characteristics, in the absence of the ability to randomise participation.

According to Little (2014), there are three main decisions affecting a matched dataset: the choice of measuring distance; the choice of matching strategy; and choice of algorithm to perform matching.

There are many different ways of measuring distance (D_{tc}) between the observable characteristics of study groups, the most common are:

1) Exact[11]:

- $D_{tc} = 0$ if $\mathbf{X}_t = \mathbf{X}_c$

- $D_{tc} = \infty$ if $\mathbf{X}_t \neq \mathbf{X}_c$

2) Mahalanobis[12]:

- $D_{tc} = \sqrt{(\mathbf{X}_t - \mathbf{X}_c)' \mathbf{S_X}^{-1} (\mathbf{X}_t - \mathbf{X}_c)}$

3) Propensity score[13]:

- $D_{tc}(\mathbf{X}_t, \mathbf{X}_c) = |\pi_t - \pi_c|$

The exact method is the most straightforward way, but it is not ideal in our case as we have some continuous observable characteristics, and it is unlikely that the value for these covariates is the same for both study groups. An extension of exact matching is coarsened exact matching (CEM), which allows continuous or ordinal data to be segmented into strata. However, if the strata are too complex, this makes it more likely it will result in failed matches as CEM requires an exact match $(D_{tc} = 0)$.

The Mahalanobis method is not ideal in our case as we have observable characteristics that include several dichotomous variables (e.g., gender) and the Mahalanobis method may not be the most suitable method for such variables (Little, 2014). Using propensity scores overcomes this through collapsing the vector of observable characteristics into a scalar propensity score.

A propensity score is the probability of participating in a given intervention, given a set of observable baseline characteristics. In our case, the relevant propensity is for a school to sign up to the ECF ERO. We have chosen to estimate the propensity scores using a logistic regression model[14]. The outcome of interest in the estimation of propensity scores is the binary indicator of whether a school is part of the main group of interest, i.e. is part of the ERO. Our matching will be a combination of exact

---

[11] $\mathbf{X}_t$ is a vector of observable characteristic values for the treatment group and $\mathbf{X}_c$ for the control group.
[12] $\mathbf{S}^{-1}$ is the covariance matrix of the observations.
[13] π_t is the probability of belonging in the treatment group, given the observable characteristics and π_c is the probability of belonging in the control group.
[14] A propensity score $e(X_i)$ can be estimated from logistic regression of the treatment condition on the covariate x $\log\left(\frac{e(X_i)}{1-e(X_i)}\right) = \boldsymbol{\beta}X_i$ (Pan and Bai, 2015).

matching on school phase and propensity score-based matching on the other covariates.

The last decision to affect a matched dataset is the type of matching algorithm used. We will match ECTs to their 'nearest neighbour(s)' with similar propensity scores. The number of nearest neighbours we select for each treatment ECT will strike a balance between a larger sample size and how close the match is. Initially we propose to use a matching strategy of 1:1 matching without replacement using the nearest neighbour matching algorithm and with the caliper set to 0.2. This is the standard recommendation used in the literature. This assigns a set of nearest propensity scores (neighbours) to a treatment school. Since each treatment ECT is matched based on a minimum distance between its propensity score and the score of its nearest neighbours, the overall heterogeneity of the matched dataset is reduced. However, if we do not achieve a large enough sample size using this method and are able to achieve a matched sample with a higher treatment to comparison sample, or using replacement, we will continue to increase the matching ratio until we achieve a sample large enough to measure the desired effect. The final sample for the comparison group will be no smaller than the number of treatment participants as per the parameters for the evaluation.

Matching imbalance will be investigated looking at the standardised mean differences (SMD) between the two groups[15], before and after matching, as this is the most used technique to explore the balance of covariate distributions between treatment groups (Zhang et al., 2019). We will also identify the extent to which matched groups are similar through data visualisations. SMD and plots will be produced using the MatchIt (Ho et al., 2013) and cobalt (Greifer, 2020) package in R (R Core Team, 2017).

We will be estimating propensity scores as well as creating a matched dataset using the MatchIt (Ho et al., 2013) package in R (R Core Team, 2017). Once a matched sample has been formed, the average treatment effect can be estimated by comparing the outcomes between treatment teachers and comparison teachers through the use of regression models using the lme4 package in R (Bates et al., 2015) as described in the primary analysis and further analysis sections below.

Survey analysis comparison group

To identify a comparison group for the secondary outcomes analysis, we will first match the list of treatment ECTs to their records in an early version of the SWC data in April 2021. We will use a similar propensity score matching approach to the one described above (see Appendix for preliminary information) to construct the comparison group, except for selecting a comparison group of no more than 3,800 ECTs (more than double the proposed sample size of 1,520 comparison group ECTs), due to the cost and time constraints of contacting the ECTs[16]. This large sample will allow for ECTs who refused to sign up, as well as any attrition. Moreover, school turnover as indicated by the proportion of teachers with less than two years'

---

[15] The SMD is given by the difference in mean outcomes between groups divided by the standard deviation of outcome.

[16] Comparison teachers for the survey sample were drawn from a sampling frame of teachers meeting our chosen criteria as there is no census indicator for who is an ECT.

experience will not be used as a covariate for matching for the secondary outcome as it might reduce the pool of eligible ECTs too drastically and reduce the balance for other covariates (see the appendix for preliminary results).

Primary analysis

Primary outcome analysis

To investigate the impact of the ECF on teacher retention in schools in England after the end of programme delivery, we will be analysing the primary outcome using a logistic regression model.[17] The regression model is given by

$$\log\left[\frac{p_i}{1-p_i}\right] = \beta_0 + \beta_k X_{ki}$$

where $\log\left[\frac{p_i}{1-p_i}\right]$ is the link function or logarithmic transformation on the outcome probability of teacher i staying at the end of 1 and then 2 years. The effects of the covariates (X_{ik}) on the logit of the model are given by β_k. Covariates for this model include:

- treatment group (comparison group as the reference group)
- teacher age (lowest age band as reference group)
- teacher gender (female as the reference group)
- school phase (primary as the reference group)
- school size
- school proportion of FSM pupils
- school Ofsted rating (inadequate as the reference group)
- proportion of teachers with less than two years' experience
- local area average pay

The above covariates could potentially impact retention and retention within the same school, and therefore need to be controlled for in the model in order to correct for any potential residual imbalance after matching has taken place. The analysis will take into account both the ECTs who completed the two year period of the intervention and those who opted out during the first year. Additionally, data for dropout during the first year of the intervention will be used to investigate whether receiving the full treatment as planned changes the magnitude of the effect. This will be done by omitting the ECTs opting out and their matches from the additional analysis.

It is the treatment group coefficient that represents the effect of the ECF on the conditional probability of a teacher staying in the profession whilst holding other covariates in the model constant. The outcome will therefore be an odds ratio.

---

[17] An advantage of discrete-time models is that they can be viewed as logistic regression models (Willett and Singer, 1993).

Inference

We will also present a layman summary measure by converting the impact to percentage points, estimated using a statistical marginal effects approach. We will also report confidence intervals and sample sizes, following the EEF reporting guidelines.

Further analyses

Secondary outcome analyses

We will analyse retention in the original induction school in the years following the start of the ECF, specifically; if they are in the same school in November 2021 and 2022, using the same model specification as the primary outcome analysis.

For the self-efficacy, teacher satisfaction, teaching quality and intentions to stay in teaching outcomes, the analysis will consist of comparing the intervention and comparison groups and across survey waves when possible. This will enable us to investigate the effect of the ECF overall and if any potential effects change over time.

Subgroup analyses

We will explore the potential effects of four moderating factors on ECTs' retention outcomes: phase; provider; FSM and workload. For phase and FSM we will use the same sample as the primary analysis, allowing us to compare outcomes between intervention and comparison groups. For provider and workload, we will only explore within the intervention group, as the same information is not available for the comparison group constructed for the primary outcome analysis.

For the phase model, we will include a phase dummy, (primary/secondary) and a phase*intervention interaction term. This will enable us to assess if there are differential effects of the ERO on teachers working in primary schools, versus those working in secondary schools.

For the school level FSM model, we will include school level FSM quintile as a series of dummy variables with the lowest quintile being included as the reference category. Each quintile dummy will be interacted with the intervention dummy. This will allow us to assess any differential effect that having higher proportions of FSM-eligible pupils has on retention.

For the providers model, we will include provider as a series of dummy variables with the largest provider being included as the reference category. This will allow us to assess any differential impacts the provider has on retention within the intervention group.

For the workload model, we will collapse survey responses 'strongly disagree' and 'tend to disagree', forming the reference category. This will allow us to assess any differential impacts that workload has on retention within the intervention group. However, we will not be able to explore comparisons between perceptions of workload and retention outcomes for the comparison group. Therefore, we are limited in making any causal inferences from this analysis. We will also run an additional exploratory analysis using any variables of interest identified by the IPE analysis that may be explored as moderating factors.

Treatment effects in the presence of non-compliance

The main analysis will be followed by a CACE analysis (Complier Average Causal Effect) in order to assess the effect of non-compliance on retention. We will use time up to date of withdrawal from the programme as a measure of compliance (as this was the only compliance data provided by DfE). Teachers may potentially have unobserved characteristics that have an influence on both the compliance with the ECF and retention. Therefore, a two-stage least squares model will be used to calculate the CACE estimate (Angrist and Imbens, 1995). The first stage of the model will be compliance regressed on all covariates that are used in the main primary outcome model and the treatment group variable. The second stage of the model will regress the primary outcome on the covariates used in the main model and will also include a covariate representing the teacher's estimated level of compliance from the first stage of the model. The coefficient of the compliance variable will be the CACE estimate of the compliance effect. In the event that there are no confounding factors affecting compliance and retention, the CACE estimate will be equal to the intention-to-treat estimate. We will use the R package ivpack to perform the CACE analysis on the primary outcome only.

Missing data

Missing data in the SWC will mean that some ECTs participating in the intervention will not be matched to their own records in the SWC. We will present descriptive statistics for both matched and non-matched ECTs participating in the intervention. The data for this comparison will come from the intervention providers.

Missing data in the SWC, be it full records or missing variables, will also mean that the full pool of existing comparison teachers will not be available for matching selection. This 'missingness' will not be investigated. We hold the assumption that missing data in the SWC will affect our ability to match intervention and comparison ECTs pools comparatively and therefore no bias will result from this missingness.

Once matching has taken place for both the intervention and control groups, missing SWC cases at follow-up, from those that have been successfully matched, will be considered as dropout from teaching. We assume, as with matching, that any missingness which is not real dropout will not affect intervention and comparison groups differently and consequently will not bias the calculation of the effect of the intervention.

Effect size calculation

In case of linear models, the numerator for the effect size calculation will be the coefficient of the intervention group from the model. All effect sizes will be calculated using total variance from the model, without covariates, as the denominator i.e. equivalent to Hedges' g. Confidence intervals for each effect size will be derived by multiplying the standard error of the intervention group model coefficient by 1.96. These will be converted to effect size confidence intervals using the same formula as the effect size itself. In the cases where the outcomes are a binary variable, effect size will be presented as odds ratios. In order to better communicate the retention research findings, we will convert odds ratios to relative risks. Although odds ratios are common measures of effect size for binary outcomes, they are not intuitive and are often

misunderstood. Relative risk is a much more intuitive measure that can be easily communicated as a change is percentage points between those exposed to the intervention and those who are not (e.g., the intervention will improve the chances of staying in the profession by 50 per cent). Consequently, we will follow Grant's (2014) suggested method for converting odds ratios to relative risks. This method allows for adjusting the calculation of baseline risks in regards to the covariates included in the logistic regression model.

Implementation and process evaluation (IPE) [18]

In considering the research questions below, the following terms and definitions are used:

- **Compliance**: the extent to which the critical ingredients of the programme are delivered to and/ or received by the target participants. We will use time up to date of withdrawal from the programme as a measure of compliance (as this was the only compliance data provided by DfE).
- **Implementation fidelity**: the degree to which the intervention is delivered as intended or prescribed. For this evaluation, that includes dosage (whether participants receive the expected number of hours of content) and an exploration of intervention fidelity using qualitative data.
- **Moderators**: variables that modify the form or strength of the relation between intervention and outcome (teacher retention). For this study these may individual teacher characteristics (e.g., gender, age and subject/phase expertise), institution characteristics (e.g., institution type and geographic location), contextual factors (e.g., leadership support and school culture) or implementation dimensions (e.g., fidelity and quality of delivery).

Research questions

6 To what extent were each of the ERO's four providers **delivered with fidelity/adherence** to initial intentions?

   6.1.   Were any **adaptations** made?

   6.2.   Was **school and ECT reach** as intended?

7 Was **compliance achieved at the ECT level**?

8 Was **take-up of each strand of activity for the four providers** (e.g., conferences, online professional learning, self-directed study materials and mentoring support) as intended?

9 What was the **quality** of programme implementation?

   9.1.   Were schools, ECTs and mentors **responsive**?

   9.2.   What were ECTs' and induction leads' perceptions of the overall programme, individual elements and the four models?

   9.3.   How were any challenges/barriers addressed?

---

[18] The IPE follows the principles detailed in the Implementation and Process Evaluation Guidance (2019).

9.4.　　What were the key conditions for success?

9.5.　　What **moderating factors** may have influenced the retention rate of ECTs in the profession?

10 What was **'business as usual'** in intervention schools?

10.1.　　What is the **programme differentiation** i.e. how does the intervention compare to usual practice?

11 **What is the counterfactual i.e. what happened in the comparison group**?

11.1.　　What was 'business as usual' in terms of support to ECTs within comparison schools?

11.2.　　Was any new support provided as part of schools' preparation for the ECF national roll-out in 2022?

12 What are ECTs' and other staff **perceptions of softer secondary outcomes for ECTs** (e.g., self-efficacy, teaching quality, satisfaction with teaching and intentions to remain in teaching)?

13 What are the **perceptions of softer outcomes for mentors** (e.g., improved confidence and expertise in mentoring, including ability to assess teacher progress and provide effective feedback, support and challenge)?

14 Why did some schools decide **not to engage** with the early roll-out?

To minimise burden on research participants the research team will undertake telephone/video interviews rather than face-to-face interviews with research participants and limit the number of data collection points. Telephone/video interviews can be less time intensive and burdensome than face-to-face interviews, are more cost effective, and offer research participants' greater flexibility over the scheduling of interviews. Crucially, remotely conducted interviews also represent the most viable form of data collection method during the Covid-19 pandemic.

As part of the IPE, we intend to use provider management information (MI) where possible to access accurate reach and dosage data and minimise data collection burden on schools/ participants. Where possible, we plan to collect the following MI:

For ECTs

- number of hours spent on self-directed study per ECT in Year 1 and Year 2

- number of hours spent on ECT training per ECT in Year 1 and Year 2

- number of hours spent on formal assessment per ECT in Year 1 and Year 2*

- number of hours spent on progress reviews per ECT in Year 1 and Year 2*.

* As these activities are devolved to schools, NFER will collect it from schools via the induction leads survey.

For mentors

- number of hours spent on training per mentor in Year 1 and Year 2.

The primary research participants will be ECTs and staff responsible for overseeing their development - a group that we subsequently refer to as induction leads. Other key stakeholders will include the school mentors and the project managers for the four providers.

A description of the IPE methods used at both the programme/provider level and the individual school level are summarised below. An outline of how each method will contribute to answering the IPE questions, is provided in Table 10 below.

ERO/provider level

1. Hold five face-to-face IDEA workshops (prior to the start of the autumn 2020 school term) to: co-construct and agree the TIDieR framework for the overall programme (workshop 1), and to better understand the specific features of the programmes being delivered by each provider (workshops 2-5); develop an overarching logic model for the ERO; examine online materials; and agree processes and format for sharing of DfE MI data.

2. Conduct telephone/video interviews with the project managers of each of the four providers as well as with the DfE project manager in the spring terms of 2021 and 2022 to explore recruitment and retention, what has been implemented, what has worked well and less well, key challenges/barriers and success factors and perceptions of outcomes.

3. Conduct desk research/analysis of DfE MI data to gather accurate dosage and reach data to reduce data collection burden on schools/participants (July/August 2021 and 2022).

School level

4. Administer 16 telephone interviews with schools (induction leads/senior leaders) between November 2020 and March 2021 to explore recruitment experiences and early delivery challenges/impacts. Four schools (potentially two primaries and two secondaries) would be recruited for each of the four providers. The achieved sample will also include schools in each of the pilot areas, as well as schools in different Ofsted categories and schools with different proportions of pupils in receipt of free school meals.

5. Administer 16 telephone interviews with schools (induction leads/senior leaders) who decided not to take part in the ERO to explore why they found it difficult to engage (and thus what could be done to make the engagement easier in the NRO). The aim will be to recruit a mixture of primary and secondary schools. The achieved sample will also include schools in each of the pilot areas, as well as schools in different Ofsted categories and schools with different proportions of pupils in receipt of free school meals. The interviews will be undertaken between November 2020 and March 2021.

6. Attend one or two remotely delivered induction events for ECTs and/or mentors per provider (maximum of eight) to gain early insights of delivery. The induction events will be attended between October-December 2020.

7. Administer usual practice survey questions (contained within induction lead in midpoint survey), to explore in June/July 2021: i) in intervention schools, how the ECF differs from usual provision and any additional support provided to ECTs and/or mentors; and ii) in comparison schools, business as usual and any additional support provided to ECTs. As it will not be possible to identify a comparison group of schools before April 2021, and we want to collect this information from treatment and comparison group schools at the same time, this activity will need to take place at the end of the 2021 summer term.

8. Administer online surveys for ECTs, in summer 2021 and 2022 to gather: i) in intervention schools, their perceptions of ECF training and support; and in ii) intervention and comparison schools, information on their role, the type and frequency of training and support activities they have engaged in, the time/costs involved, enablers/barriers, and their views on its quality[19]. We plan to survey 1900 treatment and 3800 comparison ECTs and expect response rates of approximately 60% and 30% respectively. To encourage responses, ECTs in the comparison group were offered a £10 Amazon voucher for completion of the survey.

9. Administer online surveys for induction leads, in summer 2021 and 2022 to gather: i) in intervention schools, their perceptions of ECF training and support; and in ii) intervention and comparison schools, information on their role, the type and frequency of activities they have delivered, the training and support they have received, the time/costs involved and their perceptions on its quality and outcomes[3]. We plan to survey 700 treatment and 1400 comparison schools (one response per school) and expect response rates of approximately 60% and 30% respectively. To encourage responses, induction leads in the comparison group

---

[19] The survey will also be used to collect data on secondary outcomes for the Impact Evaluation, see Section 3.4.2

were offered a £10 Amazon voucher for completion of the survey. In addition, induction leads in both the intervention and comparison group were offered feedback reports based on selected responses to the induction leads questionnaire, allowing them to compare their responses to those of other schools in the same phase.

10. Telephone/video case-study interviews with staff in 16 intervention schools (four from each of the four ERO providers). Initial interviews will be scheduled for April-June 2021, with follow-up interviews in April-June 2022. Each case study will involve interviews with one or more ECTs (we have assumed two on average per school), the school's induction lead and in-school mentor(s). In year 1 (2021), case-study schools will be drawn from the sample of schools that agreed to take part in an early induction lead telephone interview. Where additional schools are required, these will be selected to ensure a mixture of primary and secondary schools are included. The achieved sample will also include schools in each of the pilot areas, as well as schools in different Ofsted categories and schools with different proportions of pupils in receipt of free school meals. In year 2 (2022), up to ten of these schools will be interviewed again, with the aim being to track progress and outcomes over time. The remaining schools will be purposively selected based on interviews with providers and data from the induction leads and ECT surveys conducted in June/July 2021. The aim here will be to select examples of what schools consider to be good or innovative practice, or schools where early roll-out appears to have gone particularly well. The case-study selection criteria is likely to include both academies (single and multiple academy trusts) and local authority maintained schools.

Table 10: IPE methods overview

Research methods	Data collection methods	Participants/data sources (type, number)	Data analysis methods	Research questions addressed	Context	Fidelity	Quality of delivery	Reach and responsiveness	Compliance/ dosage	Moderators	Adaptation	Perceived outcomes	Cost data	Usual practice
IDEA workshop	TIDIER framework; logic model completion	Providers, NFER team	Descriptive analysis	Programme docs	✓									
Project manager telephone/ video interviews	Semi-structured telephone/video interviews	Up to two project managers from each provider and the DfE project manager (up to nine in total interviewed twice)	Deductive coding; thematic analysis	RQ8, RQ9, RQ10, RQ11, RQ14, RQ15, RQ16	✓	✓	✓	✓	✓		✓	✓		
Desk review/ analysis	Desk review/ analysis	Explore training materials/ resources (key documentation from each provider e.g., mentor and ECT handbooks)	Descriptive analysis; deductive coding; thematic analysis	N/A	✓	✓	✓		✓		✓			
Early-stage telephone interviews with participating schools	Semi-structured telephone interviews	Interviews with spokesperson in each of 16 schools (induction lead/ senior leader)	Thematic analysis	RQ8, RQ9, RQ10, RQ11, RQ14, RQ15		✓	✓	✓	✓	✓	✓	✓		✓

Main implementation dimensions/logic addressed

Research methods	Data collection methods	Participants/data sources (type, number)	Data analysis methods	Research questions addressed	Main implementation dimensions/logic addressed									
					Context	Fidelity	Quality of delivery	Reach and responsiveness	Compliance/ dosage	Moderators	Adaptation	Perceived outcomes	Cost data	Usual practice
Telephone interviews with schools who decided not to take part in the ERO	Semi-structured telephone interviews	Interviews with spokesperson in each of 16 schools (induction lead/ senior leader)	Thematic analysis	RQ16	✓									
Attend remotely delivered induction events	Remote observation	Attend up to eight induction events for ECTs/mentors	Thematic analysis	RQ8, RQ9, RQ10, RQ11, RQ14, RQ15		✓	✓	✓	✓	✓	✓			
Items on usual practice (incorporated in surveys)	Midpoint online questionnaires of induction leads (15 – 20 mins)	Induction leads in intervention and comparison schools (response rate– 60% for intervention (n=720); 30% for comparison (n=720))	Descriptive analysis to explore variability in schools' approach	RQ12, RQ13	✓					✓				✓
DfE/provider MI and analytics data	Data export from DfE/ providers	Dosage/usage data for intervention schools	Frequency counts; regression	RQ8, RQ9, RQ10, RQ11		✓	✓	✓	✓		✓			
Survey (ECTs)	Midpoint and endpoint online questionnaire (15 – 20 mins)	Intervention and comparison ECTs (response rate per wave – 60% for intervention	Descriptive statistics; thematic analysis	RQ8, RQ9, RQ10, RQ11, RQ13, RQ14		✓	✓	✓	✓	✓	✓	✓	✓	

Research methods	Data collection methods	Participants/data sources (type, number)	Data analysis methods	Research questions addressed	Main implementation dimensions/logic addressed									
					Context	Fidelity	Quality of delivery	Reach and responsiveness	Compliance/ dosage	Moderators	Adaptation	Perceived outcomes	Cost data	Usual practice
		(n=1,140); 30% for comparison (n=1,140))												
Survey (induction leads)	Midpoint and endpoint online questionnaire (15 – 20 mins)	Intervention and comparison ECTs (response rate per wave – 60% for intervention (n=420); 30% for comparison (n=420))	Descriptive statistics; thematic analysis	RQ8, RQ9, RQ10, RQ11, RQ12, RQ13, RQ14, RQ15		✓	✓	✓	✓	✓	✓	✓	✓	✓
Longitudinal case-study interviews	Short semi-structured telephone/video interviews with staff in 16 intervention schools (some interviewed twice to track outcomes)	One or more ECTs (assume two per school on average), the school's induction lead and in-school mentor (sample drawn from MI, interview and survey data)	Thematic analysis	RQ8, RQ9, RQ10, RQ11, RQ11, RQ12, RQ13, RQ14, RQ15		✓	✓	✓	✓	✓	✓	✓	✓	✓

Analysis

We will conduct quantitative analysis of the survey data to measure the type and scope of engagement (fidelity/adherence and compliance) and perceived quality/effectiveness of the provision, including differences by provider model.

We will use descriptive statistical methods to analyse the survey data, including frequencies and cross-tabulations with key variables (such as school phase and provider model).

All of our interviews will be recorded (following participant agreement) and summarised and imported into qualitative data analysis package, MAXQDA, for analysis. We will: gain an overview of the data via a cross-team analytical meeting; build an initial framework drawing on the logic models and research questions; code data according to framework themes; and interpret data in the framework to identify the range and prevalence of themes of response. We will initially use a deductive analysis approach by devising a broad set of codes from the research questions. Sub-codes will, however, arise from the data. We will conduct comparative analysis (e.g., of different phases of school and different models of support) to aid reporting against the research questions and objectives.

We will collate and triangulate all of our MI, qualitative and quantitative data sources to compare and contrast findings, and to provide a comprehensive assessment of the implementation and outcomes of the ERO.

Cost evaluation

Information will be collected from providers and schools that will allow an assessment of the pre-requisite, set-up and ongoing costs to the schools engaging in the early roll-out of the ECF. The methodology for collection will follow similar objectives to those used as part of the EEF-commissioned evaluation of three pilot programmes, which each used a different model for supporting mentoring and the development of ECTs[20]. These were developed in line with the December 2019 'Cost Evaluation – guidance for EEF evaluations' in estimating the costs of the delivery of the intervention, although adaptations were required to fit with the nature of the evaluation. For example, the evaluation will require re-interpretation of the EEF guidance given the focus on ECTs (rather than pupils). Despite the fact that development of mentors may have an indirect benefit to other colleagues and to pupils, we propose to estimate a cost per ECT, and, in line with the pilot evaluation, to consider a model of how development costs might be spread over the first 1,000 ECTs in our analysis. As schools will vary in the number of ECTs, we will investigate the costs associated with ECTs, mentors, and Induction Leads with a view to creating a 'school perspective', rather than generating a composite of the time and financial costs associated with differing numbers of these participants.

We will collect cost data from induction leads and ECTs via the use of online surveys which will be administered at two time-points (June-July 2021 and June-July 2022). We will use the survey instruments to collect data on staff time spent on set up and running the ECT programme, as well as any financial costs incurred (e.g., programme and equipment costs). We will use case studies with a variety of school-based staff to triangulate findings.

---

[20] See https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/early-career-support/ [22/01/21]

Costs associated with business as usual will be collected from comparison schools and ECTs (e.g., usual costs to schools of delivering support to NQTs/NQTs+1) as part of the broader data collection from these schools. This will again be through surveys administered at two time-points.

Costs incurred by providers will also be collected via interviews, together with actual expenditure at the end of the programme, obtained by reviewing payments from DfE. These interviews will again occur at two time-points during the evaluation.

The evaluation will also engage with the DfE's policy paper, 'Early career framework reforms: overview' (DfE, 2020), which outlines the support that will be provided to schools for additional costs associated with the roll-out of the ECF. This financial support, where relevant, will be incorporated into the cost evaluation.

Ethics

The evaluation went through ethical approval at project start up on 13th July 2020. The ethics checklist is a key process within NFER's Code of Practice (CoP), and any issues raised are escalated to the CoP group. All items on the checklist met with approval and did not need to be raised. A copy of the checklist is in Appendix A.

All participants take part in the evaluation activities with informed consent. On joining the ERO, a school representative, usually the headteacher, signs an agreement with their provider. The headteacher confirms that: they have read and understood the information provided about the programme; school participants will be required to take part in evaluation activities; and that they understand what personal data will be collected, how it will be stored and transferred, and data subjects' rights in relation to this data. Individual teachers, such as ECTs and mentors, are not asked to sign anything, but they are provided with a copy of providers' Fair Collection Notices for data collection and sharing purposes.

In addition, all evaluation participants (ECTs, induction leads, mentors and provider staff) are provided with a privacy notice specific to processing their evaluation data. Participants can withdraw from data processing at any time during the evaluation – and instructions are provided in the privacy notice for how to inform the evaluator that they do not want their data to be processed.

All interviewee and survey participants are provided with information about the purpose of the data collection and how their data will be used, prior to taking part in that data collection/giving their views. As above, they can withdraw from data processing at any time.

Data protection

Data protection statement and GDPR compliance

The evaluation will be compliant with the Data Protection Act 2018 (DPA) and General Data Protection Regulation (GDPR). NFER has ISO27001 and Cyber Essentials Plus certifications and registration with the Information Commissioner's Office.

The Evaluator has put in place appropriate measures to prevent evaluation participants' personal information from being accidentally lost, used or accessed in an unauthorised way, altered or disclosed. In addition, the Evaluator will limit access to respondents' personal information to their staff members who have a business need to see it. Any personal data shared between the DfE, the providers, EEF and the

Evaluator will be via NFER's secure portal or the DfE's secure file transfer platform, Egress.

The online surveys with ECTs and induction leads will be administered using Questback. Questback's privacy statement can be found at https://www.questback.com/data-privacy/.

Legal basis

To make the use of evaluation participants' data in the evaluation lawful, the Evaluator has identified specific grounds, known as a legal basis, for its processing.

The legal basis for processing personal data is covered by:

> *GDPR Article 6 (1) (f) which states that 'processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of the personal data'.*

We have carried out a legitimate interest assessment, which demonstrates that the evaluation fulfils one of NFER's core business purposes (undertaking research, evaluation and information activities) and will not cause damage or distress to the data subjects. It has broader societal benefits and will contribute to the evidence base on what works in supporting teachers early in their careers. The evaluation cannot be done without processing personal data but processing does not override the data subject's interests.

A separate legal basis for processing special category personal data has been identified:

> *GDPR Article 9 (2) (j) which states 'Processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.*

This processing should not cause substantial damage or distress to the data subjects. Processing this data will not result in any decisions being made about them.

How personal data will be obtained

NFER will collect personal data regarding school participants via the following methods:

For participating early roll-out schools:

- surveys and interviews with induction leads/senior leaders
- surveys and interviews with early career teachers
- interviews with school-based mentors
- teachers' administrative data obtained from the School Workforce Census (SWC)

- data collected by the four provider organisations and DfE, including contact details for the induction lead, mentors and ECTs

For comparison schools:

- surveys with induction leads/senior leaders
- surveys with early career teachers
- teachers' administrative data obtained from the School Workforce Census (SWC).

For schools that decided not to participate in the early roll-out of the ECF:

- interviews with induction leads/senior leaders

NFER will also collect personal data regarding DfE staff and providers via the following method:

- Interviews with representatives from the four provider organisations and the DfE.

Details of the personal data that will be collected

Personal data for this evaluation will include:

- Personal identifiers, contacts and characteristics of **early career teachers**. This information will be used to invite teachers to participate in evaluation activities, to match ERO participants to their SWC records, and to support analysis of the findings. Personal data will be collected on: name, job title/role, Teacher Reference Number (TRN), date of birth, school name, address and URN, email address, phone number, subject(s) taught and/or phase, degree specialism, working patterns, whether the teacher is still in the teaching profession two years after the programme (i.e. retention in the profession), whether the teacher is still in the same school two years after the programme (i.e. retention in the school in which they began their induction) and notification and reason of withdrawal from the programme (where applicable).

- Personal identifiers, contacts and characteristics of **mentors** and **induction leads**. This information will be used to invite teachers to participate in evaluation activities, and to support analysis of the findings. Personal data will be collected on: name, job title/role, Teacher Reference Number (mentors only), school name, address and URN, email address, phone number, background and experience relevant to their role as a mentor (e.g., subject(s) taught and/or phase, degree specialism, length of time in teaching, prior mentoring experience) working patterns and notification and reason of withdrawal from the programme (where applicable).

- **Early career teachers', induction leads'/senior leaders' and mentors'** views and experiences (including, for example: frequency and engagement in early-career professional development activities, perceptions of the quality and impacts of early-career support, information about the time/costs associated with the activities, how the support programme differs from usual provision for early career teachers in participating schools, 'business as usual' support for early career teachers in comparison schools, reasons for declining to participate in the early roll-out in schools which did not take up the offer).

- Information on participants' programme engagement (including, for example: attendance at provider organisation induction/training events and engagement with the programme materials and activities).
- **Providers' staff names**, contact details, job titles/roles, and views about programme (e.g., implementation, engagement and impacts).
- **DfE's staff names**, staff names, contact details, job titles/roles and views about programme (e.g., implementation, engagement and impacts).

Rights and retention periods

Participants can withdraw from the programme and/or from their data being processed as part of the evaluation at any time. If participants choose to withdraw their data from being processed as part of the evaluation, they can still participate in the programme. Should a participant withdraw from the programme or the evaluation, any personal data they have provided up to that point will still be used unless they indicate otherwise. Contact details are provided in the Privacy Notice for who a participant should contact should they wish to withdraw their data or have errors corrected in it.

Personal data collected and used by NFER for the purposes of this evaluation will not be kept longer than is necessary and will be deleted securely in accordance with NFER's internal policy within one year of report publication, currently expected to be 2024.

Data controller and processing roles

NFER and DfE are joint data controllers for any personal data used as part of the evaluation of the ERO. This means they are jointly responsible for deciding how any personal data is used and for keeping it safe. This study plan only covers data processed for the evaluation.

Personnel

Table 11: Personnel

Name	Institute	Roles and responsibilities
Overall project leadership		
Simon Rutt	NFER	Principal Investigator and Project Director– responsible for directing the work and for quality of delivery
Matt Walker	NFER	Project Leader – responsible for day-to-day management of the evaluation and liaison with EEF
Data management and survey operations		
Kathryn Hurd	NFER	Workstream lead – responsible for overseeing data management, evaluation and comparison school recruitment and contacting schools
Keren Beddow/ Priscilla Antwi/	NFER	Project managers – responsible for overseeing the day-to-day running of the operations of the project

Name	Institute	Roles and responsibilities
Kathryn Hurd/ Waleed Bhatti		43
Impact		
Connie Rennie/ Jose Liht	NFER	Impact workstream leads – responsible for overseeing the impact workstream
Jack Worth	NFER	Impact QA
IPE		
Suzanne Straw	NFER	Research Director - responsible for IPE quality
Jennie Harland/ Rachel Classick	NFER	IPE day-to-day lead

Risks

Table 12: Risks, consequences and mitigating actions

	Risk description	Consequences	Probability	Impact	Mitigating action
1	Covid-19 impacts on ERO activities and timelines	Impact evaluation compromised	H	M	• Work in collaboration with EEF and providers to adapt evaluation as required. • Process evaluation will help capture the extent of Covid-19 effects.
2	Difficulty in securing target response rates and sustained participation	Analysis would be less robust	M	M	• Clear and effective communication with ECTs, selling the benefits of involvement, appropriate reminder activities, and close liaison with other in-school stakeholders, as well as external delivery partners, will all help engagement. • A 'light-touch' IPE. Surveys of ECTs and induction leads conducted online to reduce burden. Telephone/video case-study interviews will be less time intensive and burdensome than face-to-face, and offer greater scheduling flexibility. • Induction leads will be offered tailored feedback reports from the survey; comparison group ECTs will receive vouchers for completion of surveys.
3	Timely access not given to necessary datasets (e.g., SWC 2020 data is not available in April 2021)	Evaluation activities could be delayed. Comparison group sampling methodology might not be feasible	L- M	M	• We will arrange with DfE and the providers to get access to the necessary data as early as we can. If the SWC data is available in early June, then a simplified version of the comparison group selection could be considered to get letters out by the end of the term. Sampling in September 2021 could also be considered, although some ECTs may have already left teaching or their school by then.
4	Intervention and comparison schools and ECTs are not closely matched	Impact would not be measured accurately	L	M	• We will use propensity score matching to match on both teacher, school and local labour market characteristics; an iterative process to achieve the right balance between similarity and sample size. The area-based targeting of ERO means a high likelihood that enough ECTs and schools in similar circumstances will be available for matching.
5	Comparison schools provide similar support to ERO schools	Will affect identification of impact	L	M	• Process evaluation will ascertain the level of support comparison group ECTs receive which will be built into the assessment of impact.
6	Poor quality and consistency of MI data	Will impact on understanding of dosage, reach and outcomes	L	H	• We will agree a detailed specification and process for internal MI data collection up-front and processes for informing providers and schools about what is required and gaining their agreement. • We will collect MI data at an early stage to assess its quality; improvements will be recommended if necessary.
7	Data security compromised	Could result in legal and reputational consequences	L	H	• Steps taken at every stage to ensure personal data is protected, in accordance with NFER's Data Security Policy. • The identity of schools and respondents will be anonymised in reporting. Feedback reports to schools will not allow identification of individual teachers.
8	Changes to key project personnel due to sickness, absence or staff turnover	Delivery quality or timescales could be affected	L	M	• NFER has over 100 researchers and statisticians; staff will be replaced on a skills-matched basis if necessary. • Project documentation centrally stored and kept up-to-date to ensure effective handover.

Timeline

Date	Activity	Staff responsible/ leading
July-December 2020	- Set up meetings (face-to-face) - IDEA/TIDieR workshops/logic models for each of the four providers and overall programme (face-to-face) - Desk research of programme materials - Agree processes for sharing DfE/provider MI/analytics data and what data needs to be collected - Early roll-out of ECF fully funded programme support begins (September) - Write privacy notice - Intervention group confirmed by suppliers (November 2020) - Commence observations of induction events for ECTs, mentors and/or induction leads	SR, MW, KH, KB, PA, SS, JH
January-March 2021	- Commence telephone interviews with participating and non-participating schools - Instrument design (online surveys and interview schedules) - Provider project manager telephone/video interviews - Identify and recruit case-study schools - Findings update 1 (31/03/21) (headline findings from telephone interviews with participating and non-participating schools and observations of induction events for ECTs and mentors)	MW, KH, KB, PA, SS, JH
April-June 2021	- SWC pre-publication data 2021 (to sample comparison group) (expected in April but received in June 2021) - Identify and recruit comparison group - Initial school telephone/video case-study interviews (intervention) (May) - Commence online survey of ECTs (intervention) (June) - Commence online survey of induction leads (intervention) (June)	SR, MW, KH, KB, PA, SS, JH

Date	Activity	Staff responsible/ leading
July-December 2021	• End of Year 1 ECF fully funded programme support • Commence online survey of ECTs (comparison group; delayed) (September) • Commence online survey of induction leads (comparison group; delayed) (September) • SWC data published 2021 (July) • Findings update 2 (28/10/21) (headline findings from first round of school case-studies)	SR, MW, KH, KB, PA, SS, CR, JL, JH
January-March 2022	• Findings update 3 (14/01/22) (headline findings from initial surveys of ECTs and induction leads) • Provider project manager telephone/video interviews (March)	MW, SS, JH
April-June 2022	• Follow-up school telephone/video case-study interviews (intervention) (May) • Commence online survey of ECTs (intervention and comparison) (June) • Commence online survey of induction leads (intervention and comparison) (June)	SR, MW, KH, KB, PA, SS, JH
July-December 2022	• End of Year 2 ECF fully funded programme support • Collect and analyse DfE/provider MI/analytics data • SWC data published 2022 (July) • Findings update 4 (30/09/22) (headline findings from second round of school case-studies, follow-up surveys of ECTs and induction leads	SR, MW, KH, KB, PA, SS, CR, JL, JH
January 2023-March 2024	• SWC data published 2023 (July) • All analysis completed • Emerging findings meeting with EEF (remote) (November 2023) • Draft final report (December 2023) • Presentation of final report findings (February 2024) • Final report (March 2024)	SR, MW, KH, KB, PA, SS, JH, CR, JL, JW

References

Allen, R. and Sims, S., 2017. *Improving science teacher retention: do national STEM learning network professional development courses keep science teachers in the classroom*? [pdf] Available at: <https://wellcome.ac.uk/sites/default/files/science-teacher-retention.pdf> [Accessed 31 March 2021].

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Belfield, C., Brooks Bowden, A. and Levin, H., 2018. Cost estimation in education: the ingredients method.' In: Farrow, S., ed. 2018. *Teaching benefit-cost analysis.* [e-journal], 16, pp.200 – 207. https://doi.org/10.4337/9781786435323.00023.

Buchanan, J., Prescott, A., Schuck, S., Aubusson, P., Burke, P. and Louviere, J., 2013. Teacher retention and attrition: views of early career teachers. *Australian Journal of Teacher Education*, [e-journal] 38(3). http://dx.doi.org/10.14221/ajte.2013v38n3.9.

Burge, Peter, Hui Lu, and William D. Phillips, 2021. Understanding Teaching Retention: Using a discrete choice experiment to measure teacher retention in England. Santa Monica, CA: RAND Corporation. Available at: https://www.rand.org/pubs/research_reports/RRA181-1.html.

Department for Education, 2018. *School workforce in England: November 2017.* [pdf] Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/719772/SWFC_MainText.pdf> [Accessed 31 March 2021].

Department for Education, 2019a. *Early career framework.* [pdf] Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/913646/Early-Career_Framework.pdf> [Accessed 31 March 2021].

Department for Education, 2019b. *School workforce in England: November 2018*. [pdf] Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/811622/SWFC_MainText.pdf> [Accessed 31 March 2021].

Department for Education, 2019c. *Teacher recruitment and retention strategy*. [pdf] Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/786856/DFE_Teacher_Retention_Strategy_Report.pdf> [Accessed 31 March 2021].

Department for Education, 2020. *School workforce in England: reporting year 2019.* [pdf] Available at: <https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england> [Accessed 31 March 2021].

Fletcher-Wood, H. and Zuccollo, J., 2020. *The effects of high-quality professional development on teachers and students: a rapid review and meta-analysis*. [pdf] Available at: <https://epi.org.uk/wp-content/uploads/2020/02/EPI-Wellcome_CPD-Review__2020.pdf> [Accessed 31 March 2021].

Ginnis, S., Pestell, G., Mason, E. and Knibbs, S., 2018. *Newly qualified teachers: annual survey 2017*. [pdf] Available at:

<https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/738037/NQT_2017_survey.pdf> [Accessed 31 March 2021].

Grant, R.L., 2014. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. British Medical Journal, 348:f7450. doi: 10.1136/bmj.f7450.

Hardman, M., Taylor, B., Daly, C., Glegg, P., Stiasny, B., Pillinger, C. and Gandolfi, H. 2020. *Early Career Teacher Support: Pilot Report*. [pdf] Available at: https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/ECT_Support_Pilot_Report_-_final.pdf [Accessed 14 February 2022]

Hobson, A.J. and Ashby, P., 2012. Reality aftershock and how to avert it: second year teachers' experiences of support for their professional development. *Cambridge Journal of Education*, [e-journal] 42(2), pp.177–196. https://doi.org/10.1080/0305764X.2012.676628.

Perryman, J. and Calvert, G., 2019. 'What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *British Journal of Educational Studies,* [e-journal] 68(1), pp.3-23. https://doi.org/10.1080/00071005.2019.1589417.

Sibieta, L., 2018. *The teacher labour market in England: shortages, subject expertise and incentives*. [pdf] Available at: <https://epi.org.uk/wp-content/uploads/2018/08/EPI-Teacher-Labour-Market_2018.pdf> [Accessed 31 March 2021].

Sibieta, L., 2020. *Teacher shortages in England: analysis and pay options*. [pdf] Available at: <https://epi.org.uk/wp-content/uploads/2020/02/Teacher-shortages-and-pay_2020_EPI.pdf> [Accessed 31 March 2021].

Sims, S. and Jerrim, J., 2020. *TALIS 2018: teacher working conditions, turnover and attrition.* [pdf] Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/873922/Teaching_and_Learning_International_Survey_2018_March_2020.pdf> [Accessed 31 March 2021].

Tschannen-Moran, M. and Woolfolk Hoy, A. 2001. Teacher efficacy: Capturing an elusive construct. Teaching and Teacher Education, 17, 783-805.

Walker, M., Straw, S., Worth, J. and Grayson, H., 2018. *Early career CPD: exploratory Research.* [pdf] Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/752572/Early_career_CPD-exploratory_research.pdf> [Accessed 31 March 2021].

Walker, M., Worth, J. and Van den Brande, J., 2019. *Teacher workload survey 2019: research report*. [pdf] Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/855933/teacher_workload_survey_2019_main_report_amended.pdf> [Accessed 31 March 2021].

Wolstenholme, C., Coldwell, M. and Stevens, A., 2012. *The impact of science learning centre continuing professional development on teachers' retention and careers: final report*. [pdf] Available at:

<https://www.stem.org.uk/sites/default/files/pages/downloads/The%20Impact%20of%20Science%20Learning%20Centre.pdf> [Accessed 31 March 2021].

Worth, J. and De Lazzari, G., 2017. *Teacher retention and turnover research: research update 1 – teacher retention by subject.* [pdf] Available at: <https://www.nfer.ac.uk/publications/nufs01/nufs01.pdf> [Accessed 31 March 2021].

Worth, J., De Lazzari, G. and Hillary, J., 2017. *Teacher retention and turnover research: interim report.* [pdf] Available at: <https://www.nfer.ac.uk/teacher-retention-and-turnover-research-interim-report> [Accessed 31 March 2021].

Worth, J., 2018. Latest teacher retention statistics paint a bleak picture for teacher supply in England. *NFER blog*, [blog] 28 June. Available at: <https://www.nfer.ac.uk/news-events/nfer-blogs/latest-teacher-retention-statistics-paint-a-bleak-picture-for-teacher-supply-in-england/> [Accessed 31 March 2021].

Worth, J., Lynch, S., Hillary, J., Rennie, C. and Andrade, J., 2018. *Teacher workforce dynamics in England: nurturing, supporting and valuing teachers.* [pdf] Available at: <https://www.nfer.ac.uk/media/3111/teacher_workforce_dynamics_in_england_final_report.pdf> [Accessed 31 March 2021].

Worth, J., 2019. School workforce in England: is the 'perfect storm' subsiding? *NFER blog*, [blog] 28 June. Available at: <https://www.nfer.ac.uk/news-events/nfer-blogs/school-workforce-in-england-is-the-perfect-storm-subsiding> [Accessed 31 March 2021].

Worth, J., 2020. *Teacher labour market in England: annual report 2020.* [pdf] Available at: <https://www.nfer.ac.uk/media/4063/tlm_annual_report_2020.pdf> [Accessed 31 March 2021].

Worth, J. and Faulkner-Ellis, H. (2021). Teacher Labour Market in England: Annual Report 2021. Slough: NFER.

Zhang, Z., Kim, H. J., Lonjon, G., Zhu, Y., & written on behalf of AME Big-Data Clinical Trial Collaborative Group (2019). Balance diagnostics after propensity score matching. Annals of translational medicine, 7(1), 16. https://doi.org/10.21037/atm.2018.12.10

APPENDIX A – NFER ethics checklist

Section of Code of Practice	Consideration of Code of Practice (CoP)	Yes	No	N/A
Ethics	Level of consent required – does the project allow for the level of consent required?	✓		
	Will research participants be provided with all the required information to enable them to make an informed choice?	✓		
	Have you looked at and do you intend to follow the guidance on selecting children/young people for interview?			✓
	Will you follow the protection and safety guidelines?	✓		
	If the project involves children/young people have all those involved undergone disclosures/child protection training?			✓
Data protection	Will the project follow the 8 principles of the data protection act?	✓		
	Will the project follow the rules for the processing of sensitive personal data?	✓		
Data security	Will the project allow for safe transfer of data into and out of our systems?	✓		
	Will the project include a secure coding system for recording participants' names?	✓		
	Have data transfer issues / protocols been discussed / confirmed with the client?	✓		
Caring for research participants	Will the project take into account designing research questions that make sense to children/young people?			✓
	Will the project follow the guiding principles for the development of assessment instruments, methods and systems?	✓		
	Will the project involve taking, producing and using visual images? (Please refer to points to consider when taking photographs or video images, storing images, producing illustrations and using visual images)			✓

APPENDIX B – Results from the process of matching ECTs to a comparison sample using SWC records

The theory of change assumes that as a result of the intervention there will be improvements in secondary outcomes like teacher self-efficacy and career satisfaction, which will lead to better retention (primary outcome). In order to assess this, it was necessary to identify a group of ECTs as similar as possible to the ECTs participating in the ERO intervention, and to invite them to respond to a survey. We used the SWC in order to identify ECTs working in non-ERO geographic areas. Importantly, the version of the SWC we obtained included the names of teachers. This made it possible for the research team to write to them and invite them to participate in the survey. This section documents how we proceeded to select this comparison sample for secondary outcome assessment.

Datasets to be matched

The full sample of ECTs participating in the ERO of the ECF contained 1,857 records. Once only the agreed upon school types were selected, 1,760 records remained. This was considered the ECT intervention set to be identified in the school workforce census (SWC). The research team used the full release of the 2020 census, which contained 430,237 classroom teachers and 76,173 school staff. It was agreed with the EEF that not all establishment types would be included in the impact assessment as the variation in types would make it impossible to design a survey general enough to assess their conditions. Consequently, the school types that were included were:

- Academy converter
- Academy sponsor led
- Community school
- Foundation school
- Free schools
- Local authority nursery school
- Studio schools
- Voluntary aided school
- Voluntary controlled school

Excluded establishment types were:

- Academy 16 to 19 sponsor led
- Academy 16/19 converter
- Academy alternative provision converter
- Academy alternative provision sponsor led
- Academy special converter
- Academy special sponsor led
- British schools overseas
- City technology college
- Community special school
- Foundation special school

- Free schools 16 to 19
- Free schools alternative provision
- Free schools special
- Further education
- Higher education institutions
- Institution funded by other government department
- Miscellaneous
- Non/maintained special school
- Offshore schools
- Other independent school
- Other independent special school
- Pupil referral unit
- Secure units
- Service children's education
- Sixth form centres
- Special post 16 institution
- University technical college
- Welsh establishment.

The first step we took was to identify the ERO records in the SWC. This would then allow us to find a sample of matched ECTs from the non-participating geographic areas. The teacher records in the school workforce census in the ERO geographic areas contained 114,436 individuals. We attempted identification by several methods (e.g. TRN, URN plus names, URN plus phonetic names), but it was only possible to identify 1,533 ERO ECTs (out of 1,760) to their school workforce census data. These 1,533 records were used to find a comparison match through propensity scores. Nevertheless, all 1,760 ERO ECTs were invited to respond to the survey, as it was assumed that they would not be significantly different to those for whom a SWC record was available, and that their inclusion would not affect the final balance. The covariate balance reported for the survey response sample includes the responses for ECTs which were part of the 1,760 ERO ECTs even if their SWC record was not identified.

The school workforce census for the non-ERO geographic regions contained 297,355 teacher records but only 38,377 were listed as teachers who had arrived at their current school, started their current job and had qualified within a year of the census date. These were the records used to find a PSM based comparison sample as there is no census indicator for who is an ECT. Although this does not guarantee that the teachers selected can be considered ECTs, the instructions that were sent to those selected to participate in the survey explained that the questionnaire should only be completed by newly qualified teachers who started their induction in the 2020/21 school year.

According to the census data of the ERO ECTs for whom it was possible to locate their SWC record, 92% had qualified, 95% started working and 86% arrived in their current school within one year or less of the census date. Those that did exceed a year were not excluded from the intervention. Consequently, a minority of teachers in the

intervention group might have been in their school for a longer period than those in the comparison group. These were retained in the sample as it was assumed that they would not be different from those that had only one year of experience. This will be verified through a sensitivity analysis that includes only those ECTs who have been in their current school for a year or less.

PSM procedure

As explained in the methods section, matching was performed using logistic regression fitted scores (see table below for the odds ratio estimates) as a distance measure for matching, without replacement, together with the 'nearest neighbour' algorithm. A caliper of 0.2 and a ratio of 3:1 was used as this resulted in the closest to the targeted sample of 3,800 comparison records. The variables used to match on were age, gender (exact match), phase, school size, school proportion of Ever 6 FSM pupils, Ofsted rating and local labour market conditions (using local area weekly median wages as a proxy[21]). The proportion of teachers with less than two years' experience (a proxy for turnover) was not used as it would limit the number of possible matches. The result was a match of 4,233 control ECTs to 1,533 intervention ERO ECTs. As the matched sample resulted in good enough balance and no unmatched intervention cases, no alternative specifications were needed for the secondary outcome match.

Table B1: Logistic regression for treatment membership

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	7.86	0.48	16.48	0.000
Age	-0.02	0	-4.79	0.000
Gender	0	0.07	0	1.000
Phase Primary	0.11	0.22	0.49	0.620
Phase Secondary	0.46	0.19	2.46	0.010
Ever 6 FSM	0.04	0	16.97	0.000
School size	0	0	1.59	0.110
Wages (weekly median)	-0.02	0	-29.28	0.000
Ofsted Rating Inadequate	0.45	0.12	3.71	0.000
Ofsted Rating Needs improvement	0.17	0.08	2.08	0.040
Ofsted Rating Outstanding	0.15	0.09	1.71	0.090
Ofsted Rating Too new	0.85	0.21	4.13	0.000

In order to be able to match all intervention and school workforce census records, missing values in the matching covariates were imputed with the multiple imputations by chained equation (MICE) procedure[22]. Nevertheless, only one imputation set was used for the matching. Moreover, only the matching covariate variables were entered in MICE for imputation purposes[23]. The percentage of missing values for both

---

[21] Although mean wages were the original variable in the plan, median wages were considered a better choice due to the skewness in their distribution.
[22] Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work?. International journal of methods in psychiatric research, 20(1), 40–49. https://doi.org/10.1002/mpr.329
[23] Missing Ofsted rating for recently opened schools was set to an extra 'too new' rating category in order not to impute a score for recently opened schools.

intervention and census data is presented in Table B2 below. Missing data was within the expected range (no more than 5.03% for the whole sample) for imputation without an elevated risk of bias being introduced.

Table B2: Missing data for matching covariates

	All (%)	Control (%)	Treatment (%)
Age	0.00	0.00	0.00
Gender	0.05	0.05	0.00
Phase	0.00	0.00	0.00
Ever 6 FSM	5.03	4.90	8.35
School size	1.81	1.79	2.28
Wages (weekly median)	0.01	0.01	0.00
Ofsted Rating	2.01	1.98	2.87

As can be seen in Table B3 below, wages, phase (primary and secondary), age, Ever 6 FSM, school size and all Ofsted ratings except for the 'too new' category presented elevated standardised mean differences (SMD) before matching. Standardised mean differences should ideally be below 0.1 but 0.25 is acceptable if statistically controlled for. Moreover, the variance ratio should be between 0.8 and 1.2. However, a less strict criterion considers that between 0.5 and 2.0 is also acceptable (variance ratios only apply to continuous variables)[24].

Table B3: Balance for full sample (pre-matching)

	Means Treated	Means Control	Std. Mean Diff.	Var. Ratio
Distance	-1.2	-3.14	2.03	0.3
Age	26.77	27.99	-0.2	0.62
Gender	1.73	1.76	-0.06	1.07
Phase All-through	0.03	0.03	-0.02	NA
Phase Primary	0.31	0.45	-0.3	NA
Phase Secondary	0.66	0.52	0.3	NA
School Ever 6 FSM %	34.95	25.12	0.65	1.06
School size	873.11	811.16	0.13	0.85
Wages (weekly median)	445.39	508.97	-2.6	0.12
Ofsted Rating Good	0.54	0.61	-0.13	NA
Ofsted Rating Inadequate	0.09	0.03	0.19	NA
Ofsted Rating Needs improvement	0.2	0.11	0.22	NA
Ofsted Rating Outstanding	0.15	0.23	-0.24	NA
Ofsted Rating Too new	0.03	0.02	0.07	NA

After matching, SMDs were all within the strict threshold of <= |0.1| (see Table B4 and B5 below) except for school FSM percentage which was within the more lenient <=|0.25| statistical control criterion. Wages seemed very differently distributed before matching (see Figure: Dist. Balance for Wages, Table B3 above and Table B4 below) but within the 0.5 to 2.0 criterion for acceptable variance ratio after matching. This indicated that, notwithstanding the somewhat arbitrary cut-off nature of the

---

[24] Leite, W. (2017). Practical propensity scoring methods using R. London: Sage.

balance thresholds cited in the literature, after PSM, balance was satisfactorily achieved for all covariates.

Table B4: Balance for matched sample (post-matching)

	Means Treated	Means Control	Std. Mean Diff.	Var. Ratio
Distance	-1.21	-1.29	0.08	1.05
Age	26.77	26.91	-0.02	0.88
Gender	1.73	1.74	-0.01	1.01
Phase All-through	0.03	0.03	0	NA
Phase Primary	0.31	0.31	0	NA
Phase Secondary	0.66	0.66	0	NA
School Ever 6 FSM %	34.87	32.28	0.17	1.02
School size	873.65	855.31	0.04	0.98
Wages (weekly median)	445.49	444.18	0.05	0.54
Ofsted Rating Good	0.54	0.57	-0.05	NA
Ofsted Rating Inadequate	0.09	0.07	0.05	NA
Ofsted Rating Needs improvement	0.2	0.19	0.02	NA
Ofsted Rating Outstanding	0.15	0.15	-0.02	NA
Ofsted Rating Too new	0.03	0.02	0.04	NA

Note: Age and gender are individual characteristics and all other variables are school level variables.

Dist. balance for Distance
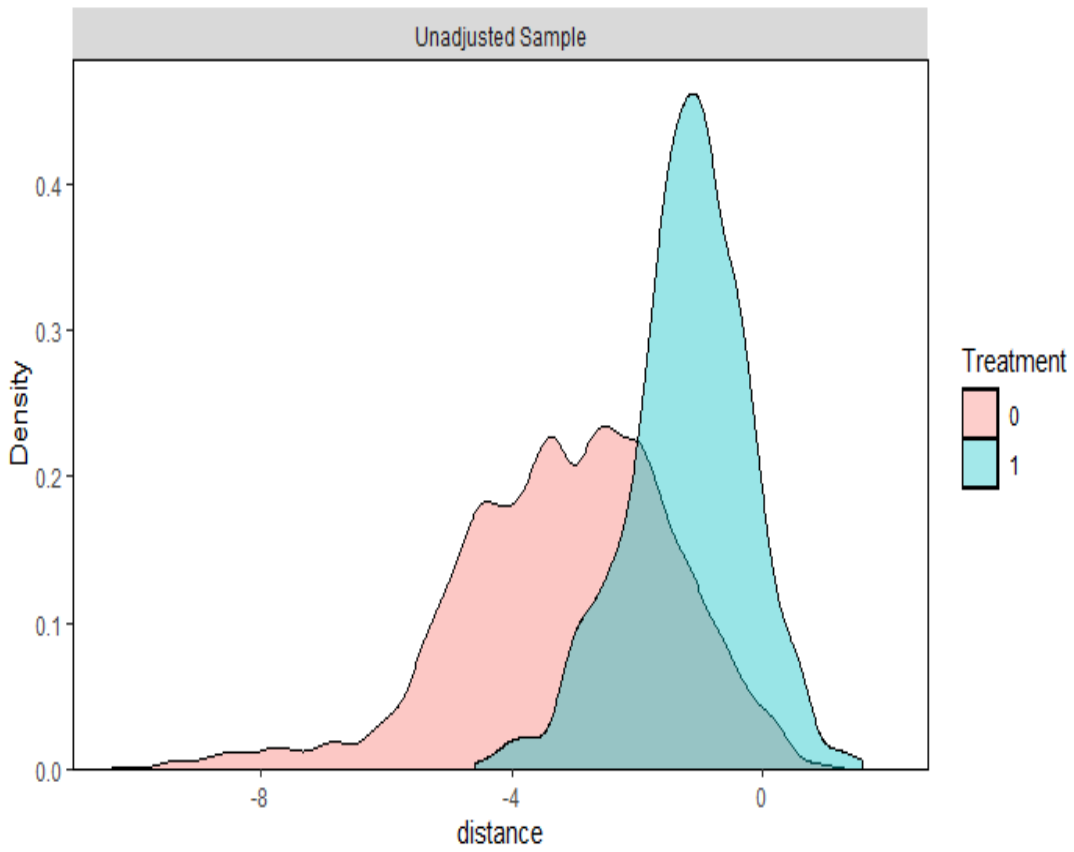
Unadjusted Sample

Dist. balance for Distance
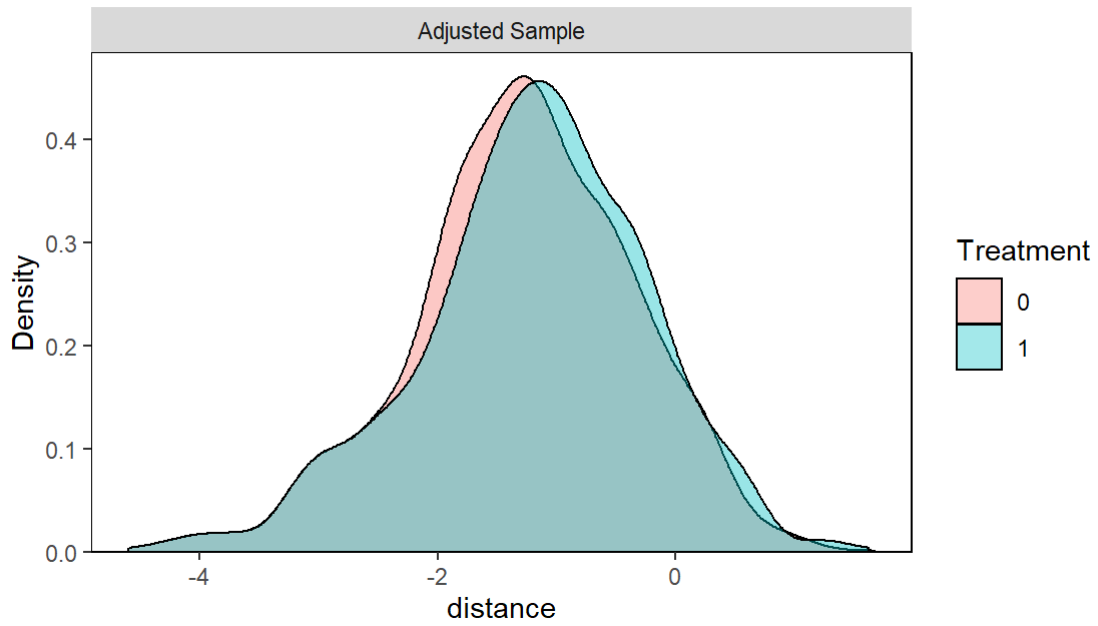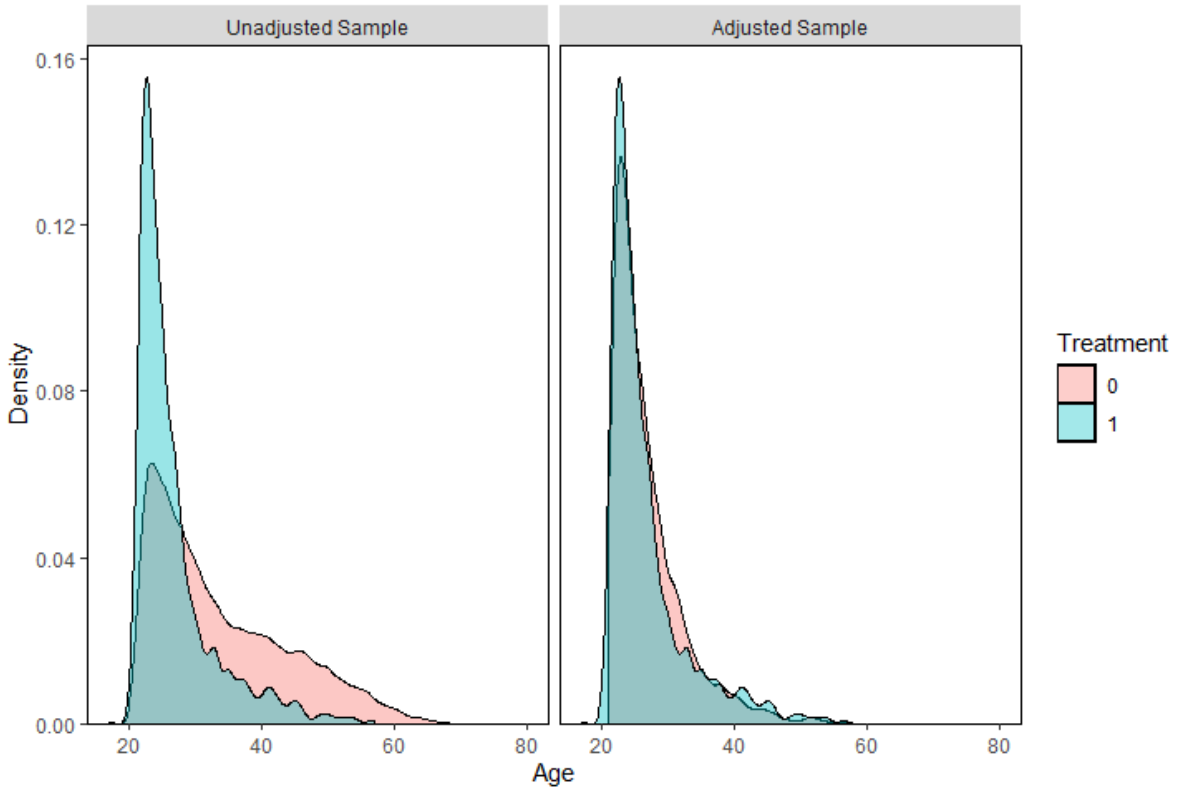
Adjusted Sample

Dist. balance for Age

Dist. balance for Gender

Dist. balance for Wages

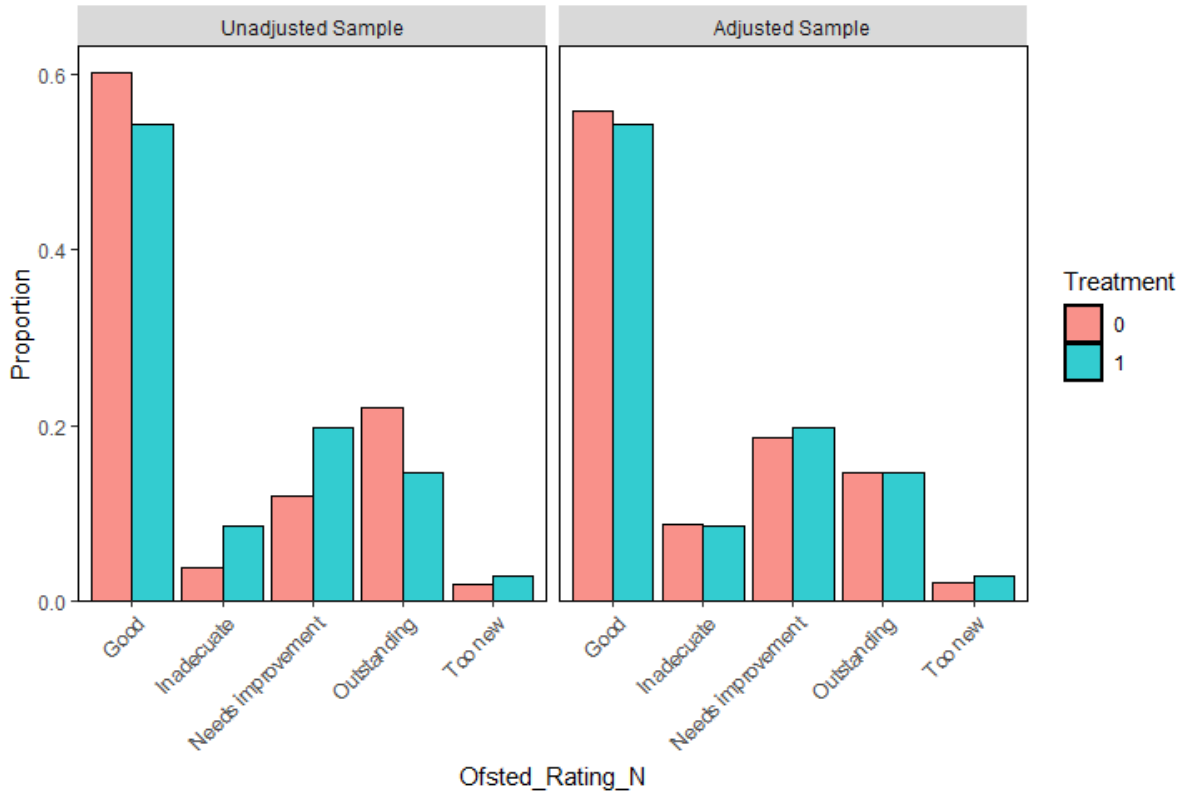Dist. balance for FSME Ever

58
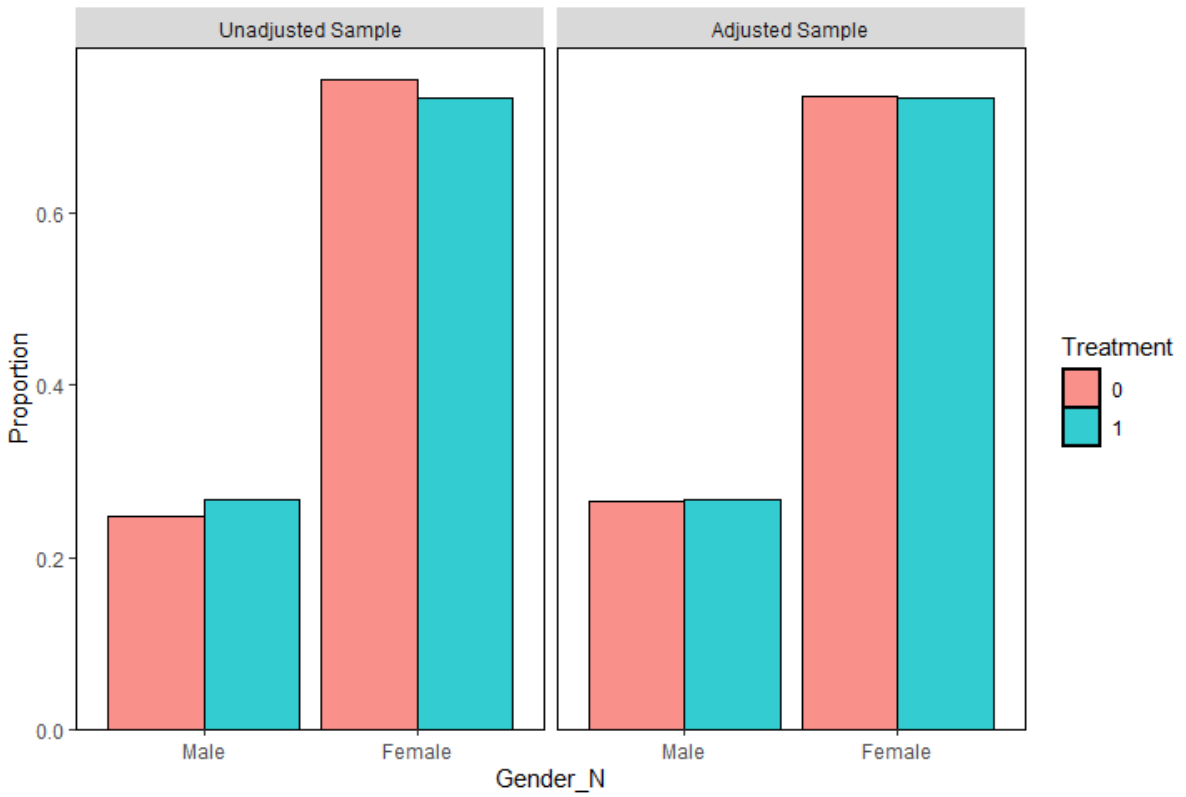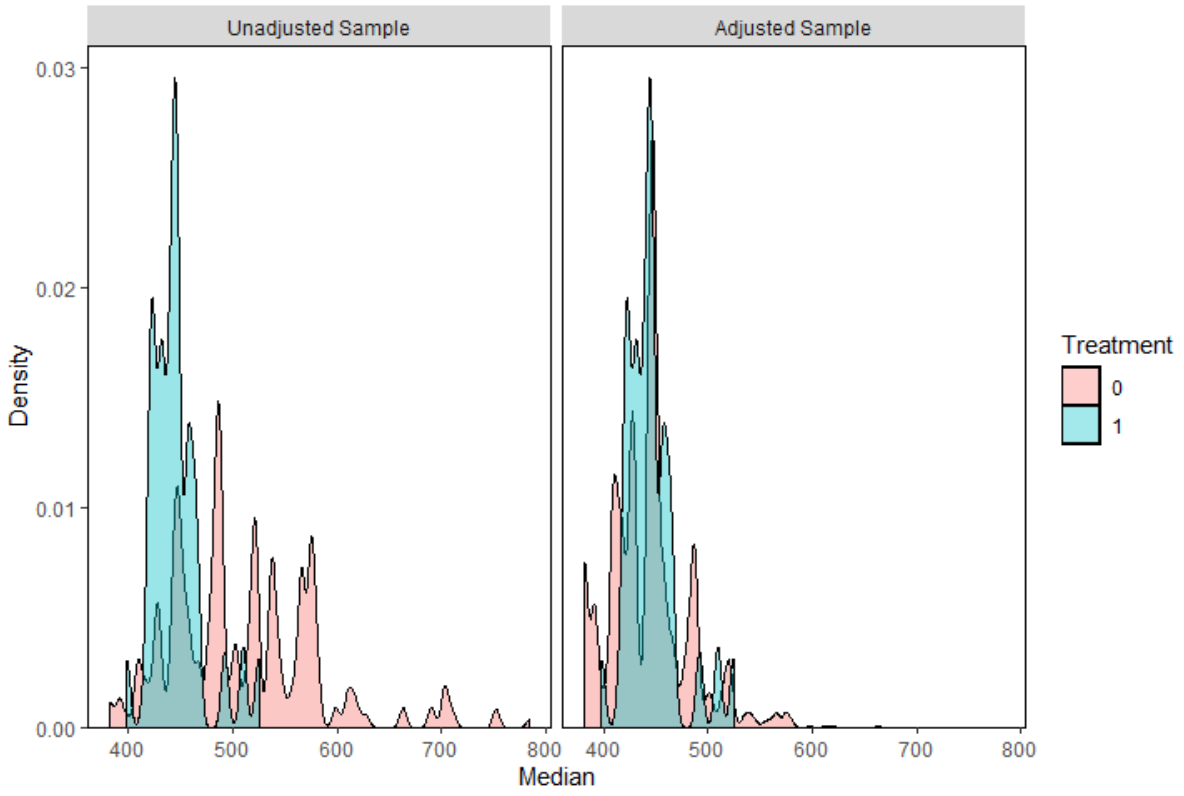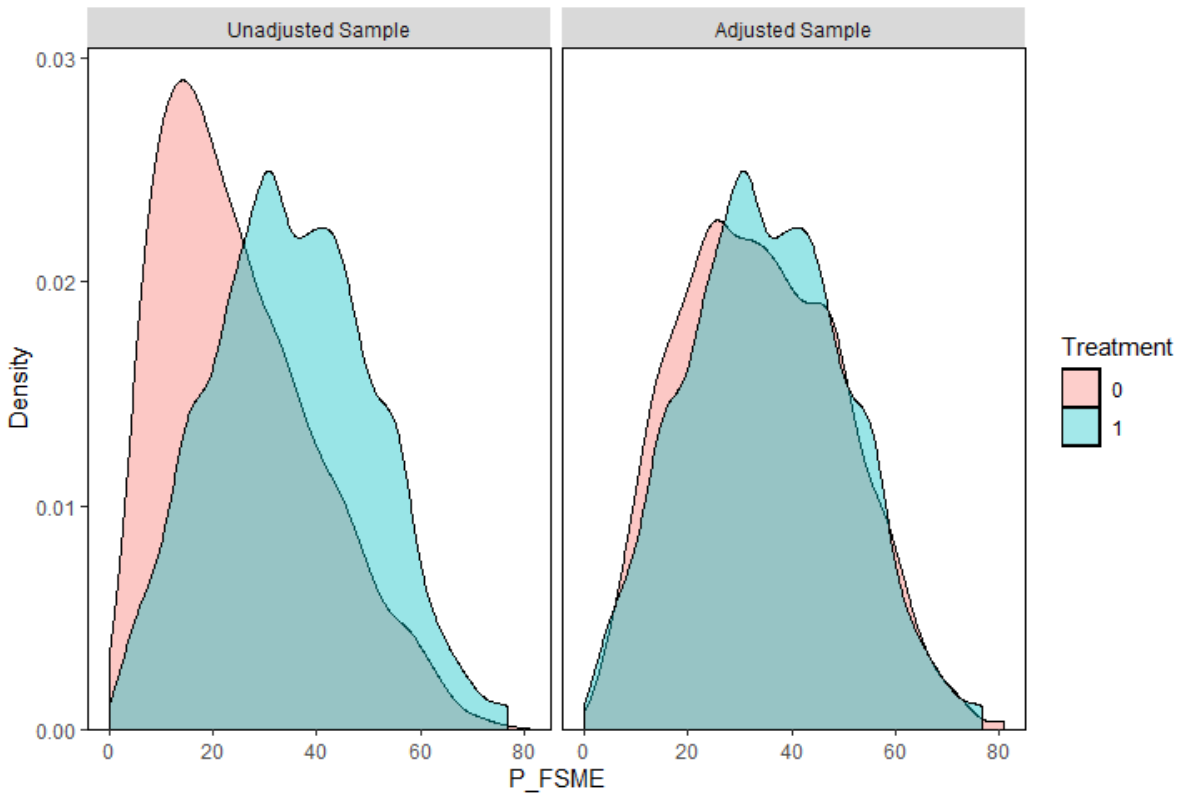
Dist. balance for Phase
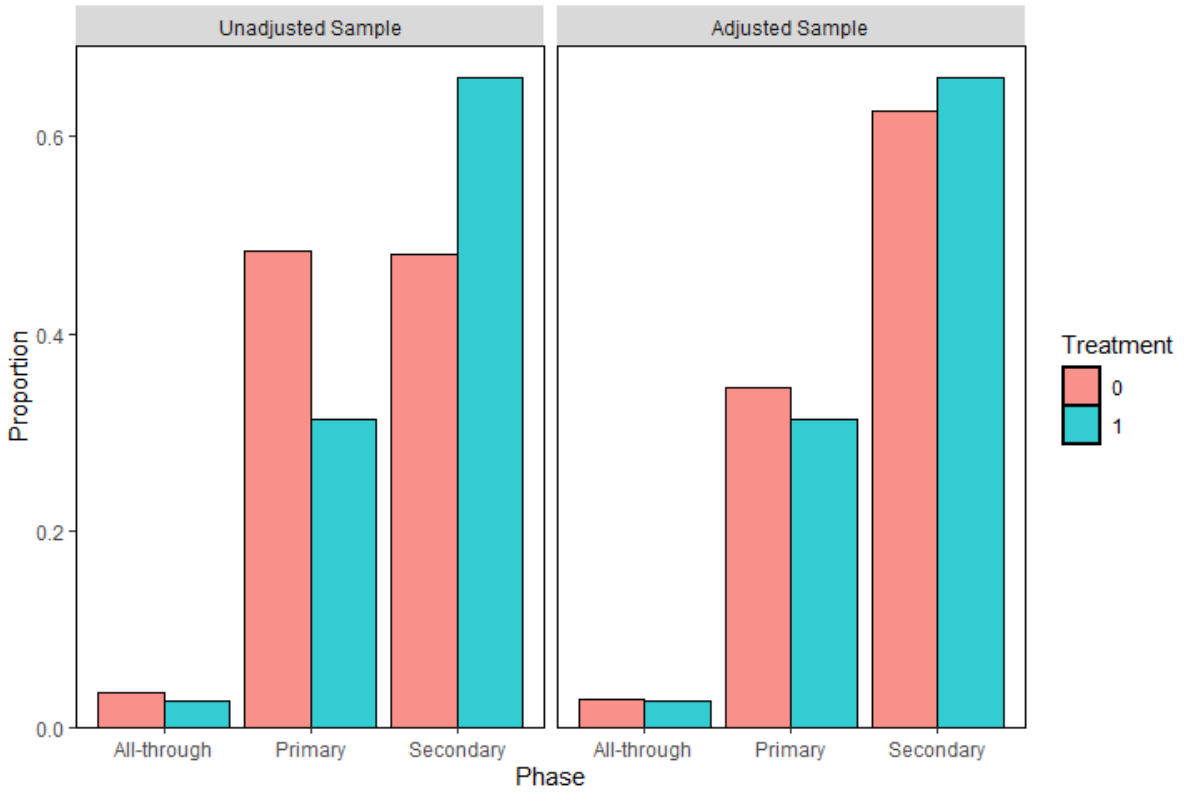
Dist. balance for School size
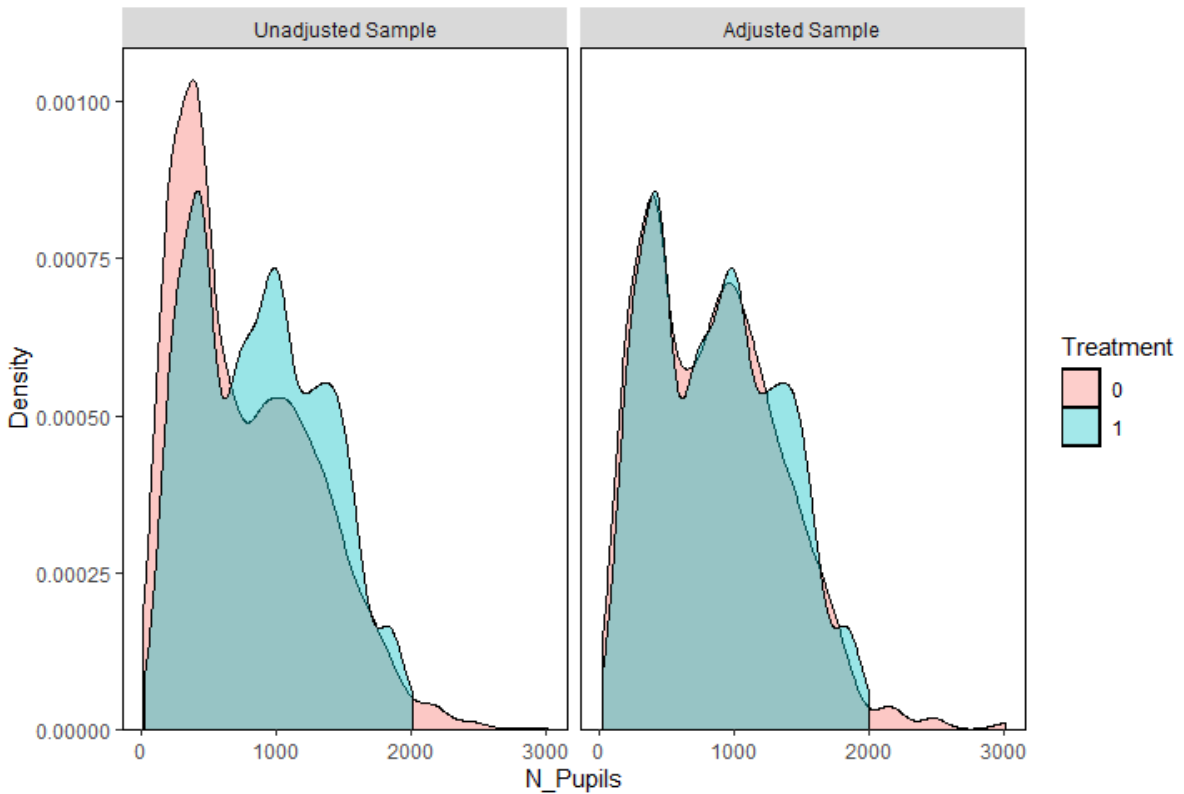
Dist. balance for Ofsted Rating

Balance in achieved sample

The survey was answered by 733 (out of 1,760) treatment and 692 (out of 4,233) control ECTs. All 1,760 participating ECTs (those in selected school types) were invited to respond regardless of whether or not they were matched with their SWC data.

Table B5: Achieved survey responses

	Eligible	Located in SWC/Matched	Responded	% Responded
Intervention	1,760	1,533	733	42
Comparison	38,377	4,233	692	16

Note: Percentage responded for the intervention is calculated out of 1,760 (since all eligible ECTs were sent an invitation).

In order to examine balance across intervention and comparison samples, without omitting cases, missing values for the matching covariates were imputed with the MICE procedure and the first imputed dataset used for analysis. Only the covariates used for matching were entered in MICE for imputation purposes. Percentage of missing data is presented in Table B6 below. As can be seen, the percentage of missing data for School Ever 6 FSM% was somewhat elevated, exceeding 5% and higher in the treatment than in the comparison condition.

Table B6: Missing data for matching covariates

|	All (%)	Control (%)	Treatment (%)
Age	0.0	0.0	0.0
Gender	0.3	0.1	0.4
Phase	0.0	0.0	0.0
School Ever 6 FSM %	6.5	3.3	9.4
School size	2.3	0.6	4.0
Wages (weekly median)	0.8	0.0	1.5
Ofsted Rating	2.6	0.9	4.2

The covariate balance for the achieved sample of respondents is presented in Table B7 below[25]. As can be seen in the table, only school FSM presented adjusted standardised mean differences above a threshold of 0.25, which indicates a lack of balance. In regards to the criterion of a variance ratio between 0.5 and 2.0 for continuous variables, none of the continuous variables presented extreme distribution ratios (see Table B7 below).

It is our recommendation that statistical control should be used to offset the lack of balance in particular for percentage of Ever 6 FSM. Table B7: Balance for achieved sample (after first survey wave)

---

[25] For gender and age, balance was computed on self-reported responses and not on the administrative data as was previously done for calculating differences in the previous section. We considered that respondent provided information would be more accurate and updated than census information. Consequently, age is now an ordinal instead of a continuous variable. Moreover, as the all-through category was small, phase now appears dichotomised into primary and secondary.

|	Means Treated	Means Control	Std. Mean Diff.	Var. Ratio
Distance	-0.27	0.13	0.71	0.49
Age: 24 or under	0.41	0.47	0.13	
Age: 25-34	0.44	0.43	-0.01	
Age: 35-44	0.11	0.07	-0.15	
Age: 45-54	0.04	0.02	-0.12	
Male	0.25	0.28	0.07	
Female	0.75	0.72	-0.07	
Phase Primary	0.33	0.40	0.13	
Phase Secondary	0.67	0.60	-0.13	
School Ever 6 FSM %	35.42	27.58	-0.56	0.88
School size	883.09	840.52	-0.08	1.15
Wages (weekly median)	445.01	448.57	0.11	1.62
Ofsted Rating Inadequate	0.10	0.05	-0.22	
Ofsted Rating Needs improvement	0.19	0.13	-0.18	
Ofsted Rating Good	0.55	0.63	0.16	
Ofsted Rating Outstanding	0.16	0.19	0.08	
Ofsted Rating Too new	0.01	0.02	0.13	

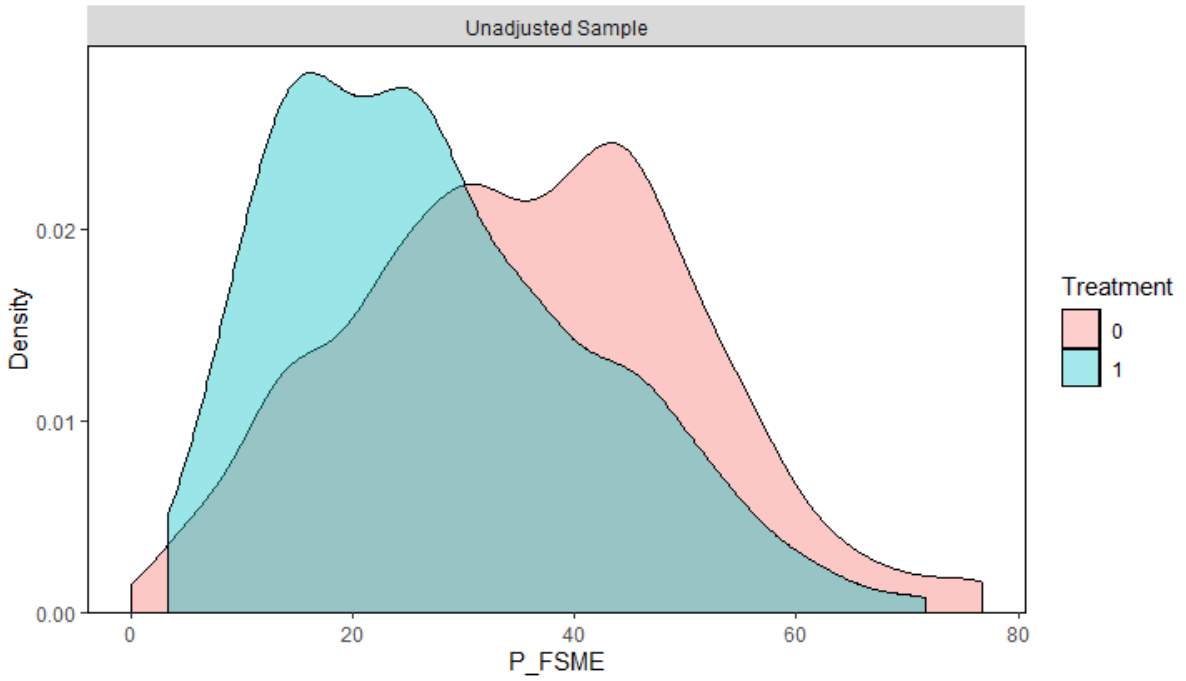## Dist. balance for Distance
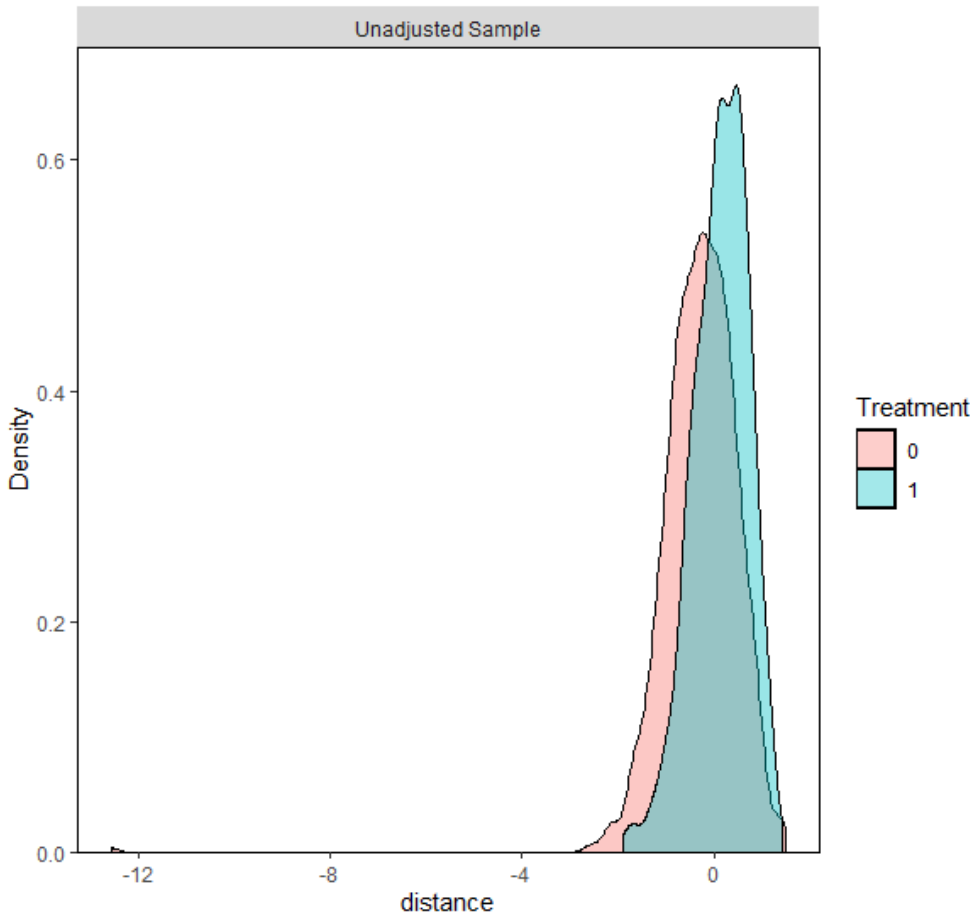
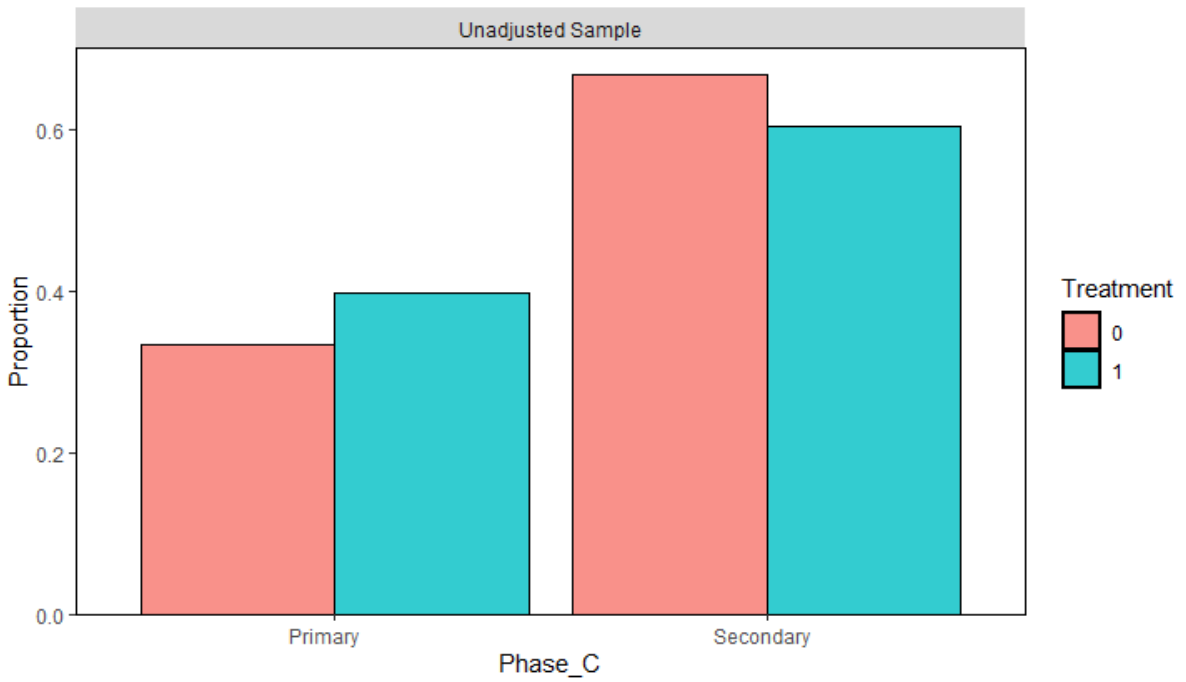

Dist. balance for Age
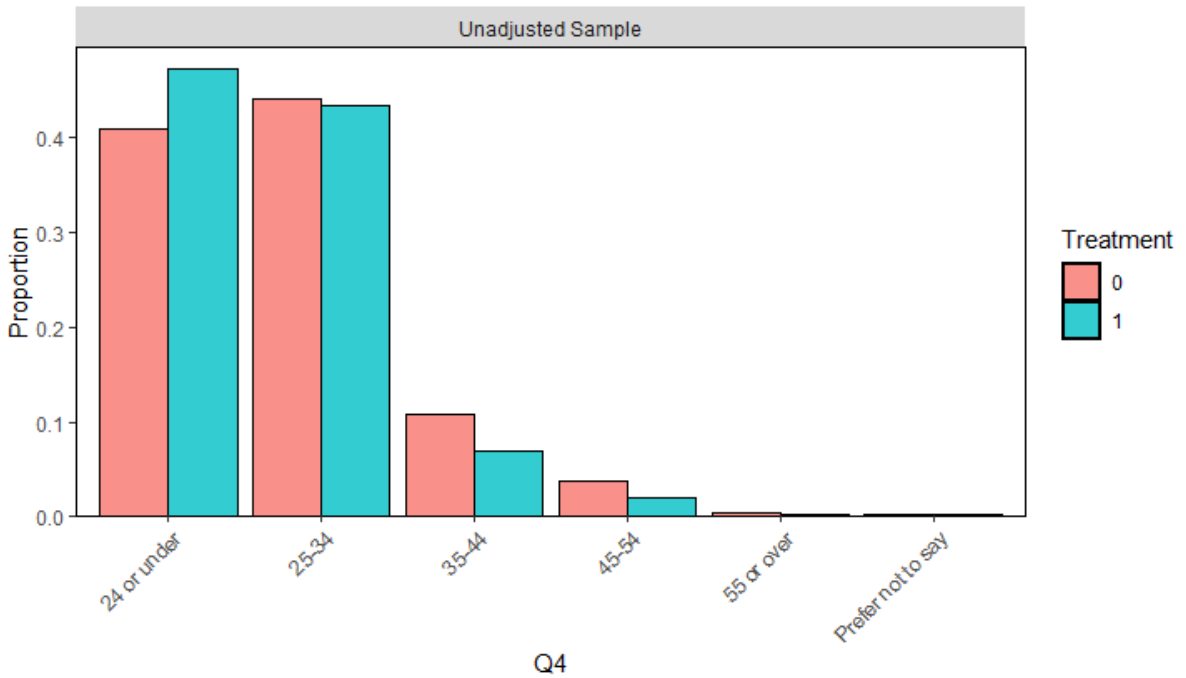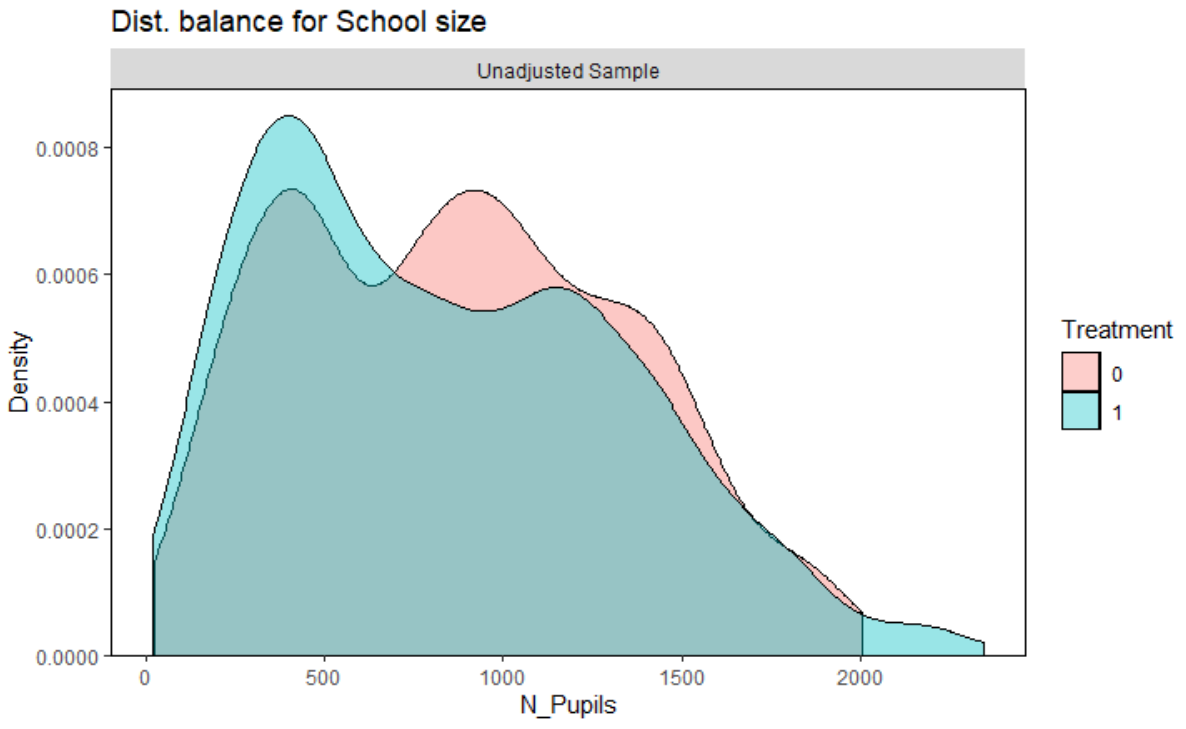
Dist. balance for Gender
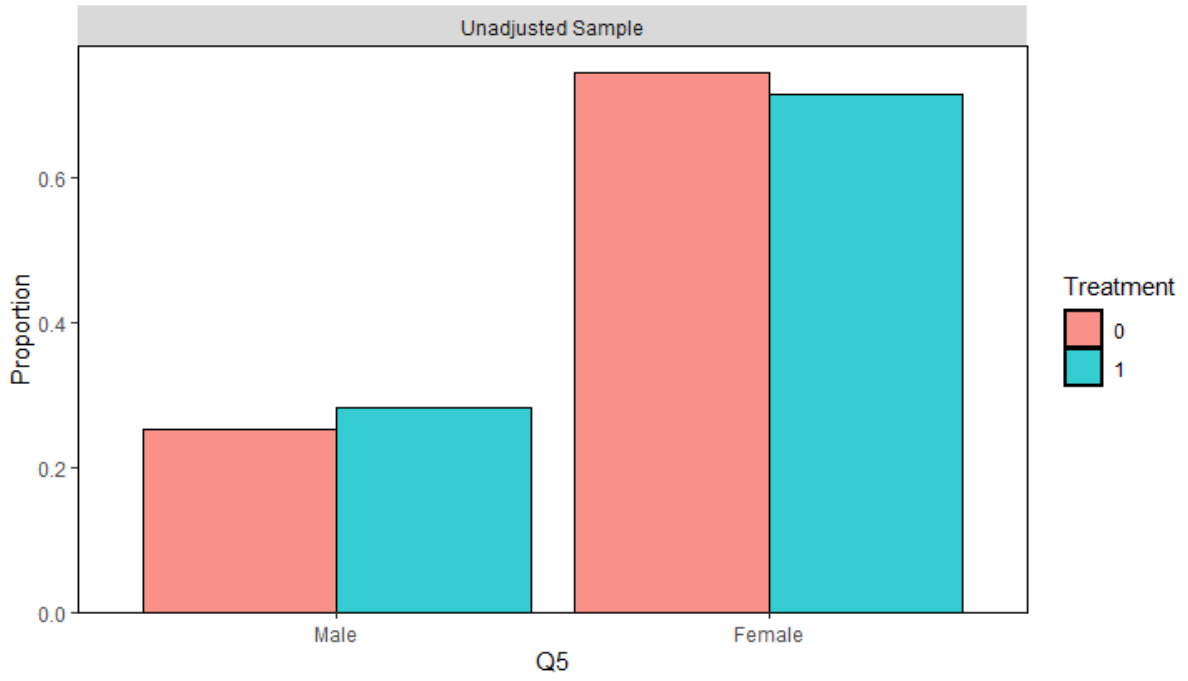
Dist. balance for Wages
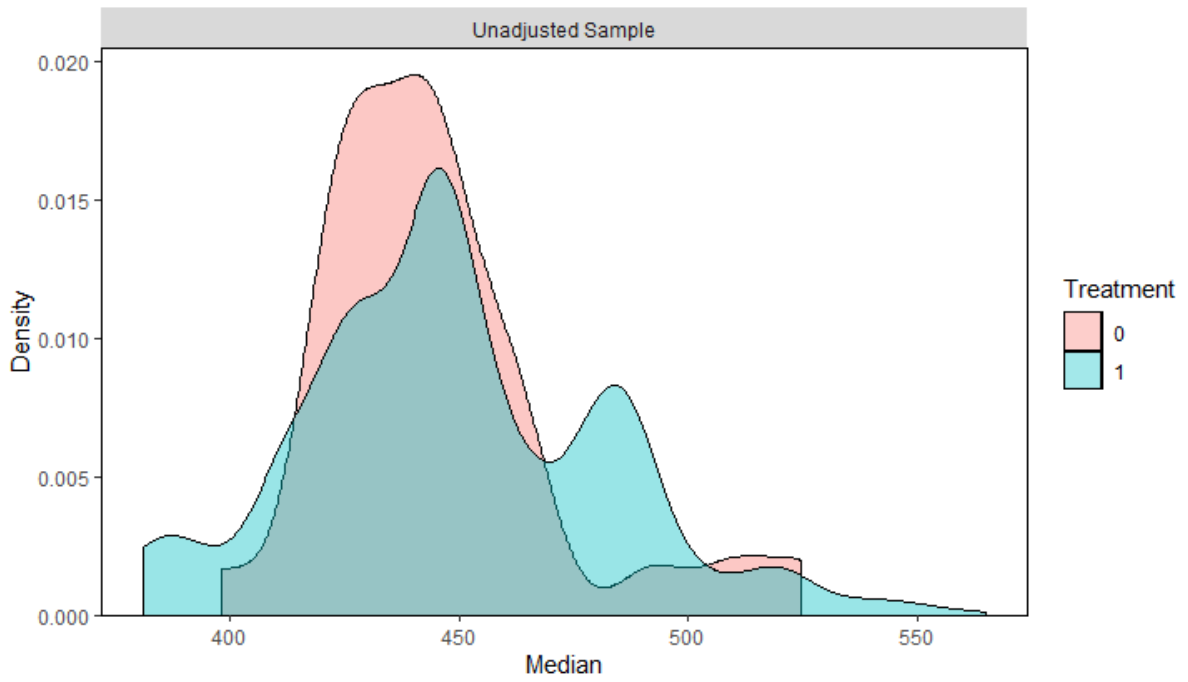
Dist. balance for FSME Ever

Dist. balance for Phase

65

Dist. balance for School size

Dist. balance for Ofsted Rating