Report

The impact of early career retention payments on teacher retention

An evaluation of impact and value for money

National Foundation for Educational Research (NFER)

The impact of early career retention payments on teacher retention: an evaluation of impact and value for money

Jack Worth and Dawson McLean

Published in June 2025

By the National Foundation for Educational Research,

The Mere, Upton Park, Slough, Berkshire SL1 2DQ

https://www.nfer.ac.uk/

ISBN: **978-1-916567-32-0**

How to cite this publication:

Worth, J. Mclean, D. (2025). The impact of early career retention payments on teacher retention: an evaluation of impact and value for money. Slough: NFER

Contents

This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners and does not imply the endorsement of the ONS or other data owners.

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation.

Bluesky: @nuffieldfoundation.org

LinkedIn: Nuffield Foundation

Website: nuffieldfoundation.org

NFER
National Foundation for Educational Research

Executive Summary

The Government faces a teacher recruitment and retention challenge, driven by under-recruitment to initial teacher training and high leaving rates (McLean and Worth, 2025). The challenge has been particularly intense for secondary schools, and under-supply has been especially marked for physics, computing, maths and chemistry teachers. Further under-supply of the specialist teachers required for a high-quality science, technology, engineering and maths (STEM) education in schools in England is a significant risk to education quality.

The Government is committed to recruiting '6,500 new expert teachers in key subjects', including plans to 'get more teachers into shortage subjects, support areas that face recruitment challenges and tackle retention issues' (Labour Party, 2024). The Government has also pledged to review 'the way bursaries are allocated and the structure of retention payments'. Strained public finances mean there has been limited resource available for improving teacher recruitment and retention. The Government has aimed to use targeted measures as part of its strategy to focus scarce resources on where they are most needed to improve teacher recruitment and retention.

Teachers in the first few years of their careers, particularly those teaching shortage subjects such as physics, chemistry and maths, have high rates of leaving the profession (Worth, Lazzari and Hillary, 2017). Early career retention payments (ECRPs) have been a key policy lever for improving retention among this group. Several different variants of the policy approach of making additional payments to early career teachers have been implemented in England since 2018 and some have been evaluated for their effectiveness at improving retention.

Previous research evidence from the UK and United States has indicated that early career retention payments are effective at improving retention (Feng and Sass, 2017; Bueno and Sass, 2019; Sims and Benhenda, 2022; CFE Research and FFT Education Datalab, 2023). This evaluation, funded by the Nuffield Foundation, presents new analysis of the impact on teacher retention of five major ECRPs that have been piloted in England since 2018. We use a combination of difference-in-differences and triple-differences methodologies to estimate the impact of the retention payments on teacher leaving rates, isolating the impacts from other influences such as changing leaving rates over time (including dramatically during the Covid-19 pandemic) and underlying differences in leaving rates by teacher experience, subject, area and school type. The study expands the evidence base on how effective and cost-effective retention payments are and draws out the implications for future policy design.

Key findings and conclusions

- Overall, eligibility for the five ECRPs that have been piloted in England since 2018 is associated with teacher leaving rates that are 5.1 per cent per year lower than they otherwise might have been. However, the impact estimate is not statistically significant at the five per cent level, having a p value of 0.06. The implied payment elasticity of attrition[1] is -0.7, meaning that

[1] An 'elasticity' is a concept from economics relating to how responsive an individual's decision or behaviour is to a change in a contextual factor. In this case the 'payment elasticity of attrition' captures how responsive

a one per cent increase in overall remuneration is associated with a 0.7 per cent decrease in the leaving rate. This suggests that while our findings are partially supportive of the conclusions from previous research that retention payments are effective at improving retention, they are not conclusive.

- Eligibility for the levelling up premium (LUP) in 2022/23 is associated with a 6.3 per cent reduction in the leaving rate and an elasticity of -1.0, although neither are statistically significant. This suggests that we cannot conclude with a high degree of confidence that LUP is associated with higher teacher retention and indicates that the retention impact may be due to chance.
- Eligibility for the maths phased bursary (MPB) is associated with a 10.9 per cent per year reduction in leaving rate and an elasticity of -2.1. Both estimates are statistically significant, implying with confidence that MPB led to lower teacher leaving rates. This is consistent with the estimates from the previous evaluation (CFE Research and FFT Education Datalab, 2023).
- Eligibility for the maths and physics retention payment (MPRP) is associated with an 8.1 per cent per year reduction in the leaving rate and an elasticity of -1.2, although neither are statistically significant. This is lower than estimates from a previous evaluation (Sims and Benhenda, 2022). Our replication analysis suggests this difference is due to differences between the respective analyses in terms of which pre-intervention cohorts were included in the analysis and how eligibility was defined. We believe that both the previous study and ours take defensible approaches to undertaking the same analysis, suggesting that estimated impacts on retention are sensitive to the particular approach taken to defining eligibility and estimating the impact. Any one study, including this one, therefore needs to be interpreted cautiously and within the context of the wider literature.
- Both the teacher student loan reimbursement (TSLR) and early career payment (ECP) schemes are associated with teacher leaving rates that are higher than they otherwise might have been, although neither is statistically significant. This is counterintuitive to the hypothesis that making additional payments to teachers would be likely to improve retention, but the results could be driven by unobserved factors or simply be down to chance.
- We use these findings and a simulation modelling technique to understand the difference in long-run teacher supply that such changes in remuneration might result in over the career span for a cohort of teachers. An implied elasticity of -0.7 suggests that the cost per additional teacher-year gained is around £43,000[2]. This compares to an estimated cost per additional

retention decisions are to changes in overall remuneration. Specifically, it is measured as the percentage change in the leaving rate that is associated with a one per cent increase in overall remuneration.

[2] By 'additional teacher-year' we mean one more teacher working in the state-funded sector for one more year than would otherwise. We estimate this for a hypothetical cohort of teachers over an entire career span, taking into account currently available information about how likely teachers are to leave and to subsequently return. The estimates that range from £9,000 to £43,000 are higher than the payments themselves, which reflects the fact that there is some degree of 'deadweight' cost, i.e. payments to teachers whose recruitment or retention behaviour is unaffected by the change in payment. In other words, while the retention decisions of some teachers are affected by the payment, which leads to the effects and additionality, the decisions of many teachers are unaffected, which thereby contributes to the cost per additional teacher-year gained.

teacher-year for training bursaries of £9,000 - £13,000. Assuming instead an elasticity of -1.5 (a consensus estimate for the teacher pay elasticity of attrition derived from a review of the literature by the Department for Education) implies that the cost per additional teacher-year for retention payments is around £20,000, which is still higher than for bursaries. These scenarios suggest, with low certainty given the somewhat inconclusive findings from our estimates, that ECRPs may offer positive value for money so long as it is infeasible or unreasonable to raise the bursary any further (as is arguably the case for perennial shortage subjects such as maths, physics, chemistry and computing). ECRPs could therefore provide additional scope for improving retention and teacher supply as part of a wider strategy to improve teacher supply and where they are focussed on subjects with bursaries that are at a clear maximum.

- There is limited evidence of variation in responsiveness to retention payments by characteristics, although some heterogeneity is evident on some characteristics. There is an indication that career changers, early career teachers that trained through School Direct and teachers in Outer London may be more responsive to payments than other teachers. However, these are not strong enough findings to confidently inform future policy design, nor are they factors that clearly link to wider policy goals that might also be emphasised in future policy design.

Recommendations

We recommend that:

- The Government should maintain a policy of teacher retention payments focussed on shortage subjects. Our evidence suggests that while their marginal cost is likely to be high, the impact of eliminating retention payments could be to worsen the teacher supply in these important subjects. Removing the payments could lead retention to worsen and sustaining them long-term as an offer for future early career teachers could prompt additional recruitment benefits.
- The Government should retain a policy of raising bursaries for subjects experiencing teacher supply challenges where bursaries are low and maintain high bursaries for maths, physics, chemistry and computing, raising them over time with the level of the teaching starting salary.
- The Government should continue to monitor and evaluate the impacts of new retention payment policies, such as the Targeted Retention Incentive. This report gives a comprehensive overview of the impact of the policies that it is possible to assess using SWC data up to 2023, but further evaluation opportunities will continue to become available. Evaluating the impacts of current and future policies will be a priority, alongside assessing the longer-term and post-eligibility impacts of previous policies.
- The Government should invest in deepening the evidence base of interventions that can improve teacher recruitment and retention. While the quality of research evidence around ECRPs and bursaries is high and growing, there is less high-quality and quantifiable evidence about the impacts of, for example, workload reduction, flexible working and professional development. Evidence on the impacts and costs of a wider range of policy measures would enable better comparative assessments of the relative costs and impacts, informing overall strategy development that is focussed on maximising cost effectiveness.

1. Introduction

Teachers in the first few years of their careers, particularly those teaching shortage subjects such as physics, chemistry and maths, have high rates of leaving the profession (Worth, Lazzari and Hillary, 2017). Early career retention payments (ECRPs) have been a key policy lever for improving retention among this group. Several different variants of the policy approach of making additional payments to early career teachers have been implemented in England since 2018 and some have been evaluated for their effectiveness at improving retention.

This evaluation, funded by the Nuffield Foundation, presents new analysis of the impact on teacher retention of five major ECRPs that have been piloted in England since 2018. The study expands the evidence base on how effective and cost-effective retention payments are and draws out the implications for future policy design.

1.1. Policy context

The Government faces a teacher recruitment and retention challenge, driven by under-recruitment to initial teacher training and high leaving rates (McLean and Worth, 2025). The challenge has been particularly intense in secondary subjects, and under-supply has been especially marked for physics, computing, maths and chemistry teachers. Further under-supply of the specialist teachers required for a high-quality science, technology, engineering and maths (STEM) education in schools in England is a significant risk to education quality. There has been chronic under-recruitment and higher-than-average leaving rates for maths and science for many years, primarily due to STEM graduates having relatively attractive career options outside of teaching, compared to teachers of other subjects (Worth and Van den Brande, 2019; *Migration Advisory Committee (MAC) report: teacher shortages in the UK*, no date).

The Government is committed to recruiting ‘6,500 new expert teachers in key subjects’, including plans to ‘get more teachers into shortage subjects, support areas that face recruitment challenges and tackle retention issues’ (Labour Party, 2024). The Government has also pledged to review ‘the way bursaries are allocated and the structure of retention payments’.

Strained public finances mean there has been limited resource available for improving teacher recruitment and retention. The Government has aimed to use targeted measures as part of its strategy to focus scarce resources on where they are most needed to improve recruitment and retention. Targeted measures also potentially offer good value for money as they avoid the deadweight cost of spending money making teaching more financially attractive in subjects that already have healthy supply. Training bursaries have been used for many years and become highly differentiated by subject, with shortage subjects attracting bursaries of up to £30,000, while others attract no bursary. As bursaries for shortage subjects have approached the level of the teacher starting salary, there has been focus on further improving the financial attractiveness of teaching through retention payments. Such payments supplement the pay of early career teachers during their employment and in addition to their basic salary.

1.2. About the retention payments

Several payment schemes aimed at improving early career teacher retention have been introduced in England since 2018, all of which have varied in their scale, scope and duration. The five key payments we evaluate the impact of in this report are summarised in Table 1 below, along with the Targeted Retention Incentive, which was introduced in 2024 and is therefore too recent to be included in this evaluation. The table shows the subjects that were eligible (although how subjects were defined differed between the schemes) and the years that payments were made. Table 10 in Appendix A gives a more detailed summary of the eligibility criteria and payment amounts for each scheme.

Table 1 Summary of teacher retention payment schemes implemented in England since 2018

Retention payment	Subjects eligible	Payment years
Maths and physics retention payment (MPRP)	Maths, physics	2019/20 – 2020/21
Teacher student loan reimbursement (TSLR)	Physics, chemistry, biology, computer science, languages	2018/19 – 2031/32
Maths phased bursary (MPB)	Maths	2021/22 – 2024/25
Early career payment (ECP)	Physics, chemistry and languages, maths	2022/23 – 2024/25
Levelling up premium (LUP)	Maths, physics, chemistry, computer science	2022/23 – 2023/24
Targeted retention incentive (TRI)	Maths, physics, chemistry, computer science (also in FE)	2024/25 – 2025/26

Note: TRI was initially launched as an extension of the levelling up premium but renamed by the new Government in summer 2024.

Sources: (Department for Education, 2017, 2018, 2019, 2022)

1.3. Existing evidence

Several evaluations have assessed the impact of some of the above retention payment schemes on teacher retention. Sims and Benhenda (2022) evaluated the first year of the MPRP scheme, finding that eligible teachers were 23 per cent less likely to leave teaching in state-funded schools in years they were eligible for payments. Sims and Benhenda estimate that the payment elasticity of attrition is -3, meaning that a one per cent increase in overall remuneration is associated with a three per cent fall in the leaving rate.

The Department for Education analysed the impact of the TSLR scheme on teacher retention in 2019 and 2020, finding that the scheme was associated with lower teacher leaving rates of approximately 5 to 20 per cent (CFE Research, 2023). However, none of the estimates was statistically significant, due to small sample sizes. A payment elasticity of retention was not calculated, but our analysis (see Table 4) finds that the TSLR during this period was equivalent to around a one per cent increase in annual remuneration, suggesting that the payment elasticity of attrition could be in the region of -5 to -20 (but still subject to a high degree of imprecision).

CFE Research and FFT Education Datalab (2023) analysed the impact of the MPB scheme, finding that the lower-value £5,000 payment was associated with a 37 per cent reduction in the probability of leaving, implying a payment elasticity of attrition of -2.2. The study also found that the higher-value £7,500 payment was associated with a 58 per cent reduction in the probability of leaving, implying a payment elasticity of attrition of -2.3. However, the study also found that the bursary reduction that also formed a part of the overall policy design was associated with a 10 to 15 per cent reduction in recruitment to ITT, more than offsetting the retention gain in terms of overall teacher numbers.

A wider literature exists on how responsive teachers tend to be to retention payments, pay and other financial incentives. A study from Georgia, USA found that bonuses paid to early career maths and science teachers were associated with a 25-28 per cent lower leaving rate (Bueno and Sass, 2019). According to Sims (2017), this is equivalent to a payment elasticity of attrition elasticity of -3.4. Another study from Florida, USA found that bonuses paid to early career teachers for specific subjects were associated with a 32 per cent lower leaving rate (Feng and Sass, 2017). Other studies have found null effects overall of payments on retention, but impacts on other outcomes such as recruitment into hard-to-fill roles and retention within sub-groups (Springer, Swain and Rodriguez, 2016; Theobald *et al.*, 2024).

A study on training bursaries in England found that bursary increases are associated with increases in recruitment into initial teacher training and the additional teachers induced to enter training by a bursary increase tended to complete their training, enter teaching and be retained in teaching at the same rate as other teachers in their cohort (McLean, Tang and Worth, 2023). It found that bursaries offer good cost effectiveness compared to other targeted policy measures such as early career retention payments, especially where the existing bursary for a subject is low.

An evidence review by the Department for Education established that estimates of pay elasticities of attrition vary in the literature depending on the study designs, location of the study and types of teachers included (DfE, 2020). It concluded that a reasonable assumption for an overall pay elasticity for all teachers is around -1.5, with an acknowledgement that the elasticity may be higher among early career teachers.

In summary, the existing evidence finds that retention payments are likely to be associated with improvements in retention, although there is a wide range of estimates that depend on the circumstances of the policy intervention and the evaluation design. We conclude this report by situating our findings within this wider literature.

1.4. Aims of this research

The aim of this evaluation is to identify and estimate the impact that early career retention payments have had on the retention of early career teachers.

The financial attractiveness of teaching, and specifically the contribution made by retention payments, is just one of the many factors that influences teachers' career decision making. Other factors include personal factors (e.g. age and experience, subject taught), workplace factors (e.g. workload, levels of support from senior leaders and the availability of professional development and flexible working opportunities) and wider economic factors (e.g. pay relative to outside earnings, the state of job opportunities in the wider economy) (Hutchings, 2011; Worth *et al.*, 2018; Adams *et al.*, 2023; Harland, Bradley and Worth, 2023; Martin *et al.*, 2023). Nonetheless, the research literature suggests that retention payments can be an important tool for improving teacher retention.

Our research questions for this evaluation are:

1. Are early career retention payments (ECRPs) effective at increasing teacher retention?
a. Does the response hold even after teachers stop being eligible for payments?
2. How does the effectiveness of ECRPs differ by:
    a. type and size of payment?
    b. teacher characteristics and subject?
    c. geographical region, school type and local levels of disadvantage?

The impact of ECRPs on recruitment of teachers to ITT, through anticipation of future payments, was not in scope for this evaluation. However, as retention payments become more established and may in future be supported by longer-term policy guarantees and/or integration within the pay framework, this should be a focus for future research.

1.5. Structure of this report

Section 2 outlines methodology we used for this evaluation and the data we analysed. Section 3 summarises the findings on the impact of ECRP eligibility on retention, while section 4 incorporated information about the costs of the schemes to analyse value for money. Section 5 explores heterogeneity of impact, assessing the extent to which this varied according to teacher-, school- and area-level characteristics. Section 6 concludes and makes some recommendations.

NFER
National Foundation for Educational Research

2. Methodology and limitations

2.1. Methodology

2.1.1. Evaluation approach

To evaluate the impact of financial incentives in isolation from other changes and factors, we need to compare the retention rates of groups of teachers who were otherwise similar apart from one group being eligible for financial incentives and another group not. The comparison group that are not eligible then act as a counterfactual, allowing us to estimate what the retention rate might have been expected for the group of teachers who received the payments, if they had not received them.

Our general approach is to use a difference-in-differences evaluation framework. Eligibility for the payments varied across the different schemes (see Table 1 below and Table 10 in Appendix A for more details) but was typically based on eligibility criteria that related to factors including subjects, time periods, training cohorts, regions and school characteristics. The time period criterion (i.e. payments were only made in specific years) means we can compare the retention rates of teachers who would have been eligible before the policy was introduced to teachers who were eligible. However, we also need to account for the fact that retention rates can change over time for a range of reaons, as change may have occurred even in the absence of the financial incentive. Under the assumption that the trends over time (e.g. changes to pay, economic environment) affect the retention rate of all teachers similarly, we can compare the before-and-after change in retention rates in the comparison group with the before-and-after change in retention rates in the treated group to derive a robust 'difference-in-differences' estimate of the policy impact.

Following Sims and Benhenda (2022), we also go a step further by deploying a 'triple differences' approach. Eligibility for many of the payments varied across multiple eligibility criteria. For example, MPRP eligibility was based on subject (only maths or physics specialist teachers were eligible), geography (only teachers in some areas were eligible) and time period (payments were only made to eligible teachers in the 2019/20 and 2020/21 academic years). We exploit these features to make multiple comparisons between different types of teacher to further improve robustness. We can compare the retention rates of teachers of eligible subjects in eligible and ineligible areas to control separately for subject-specific area effects. We can also compare the retention rates of teachers in eligible areas in eligible and ineligible subjects to control separately for area-specific subject effects. Making these comparisons over time also allows us to relax the assumption that the trends over time need to affect all teachers similarly, since they can vary by area and subject and still be isolated from the estimated impact of the financial incentive via the triple difference controls.

As we do not have data on which teachers applied for and received retention payments, we infer eligibility from the teacher characteristics captured in the data (see below). This means our evaluation captures the estimated impact of being eligible for the payments, but not necessarily the impact of having received it if not all eligible teachers applied and received payments. Formally, this means our estimate is an 'intention to treat' estimate. While an 'intention to treat' estimate can be diluted if take-up is below 100 per cent, it has the advantages of reflecting real-world policy implementation and not be affected by ‘selection bias’ (i.e. the types of teacher receiving the payment systematically differing from those who did not in ways that are associated with a higher or lower retention rate). We understand from DfE that take-up of the payments was generally high, meaning that our ‘intention to treat’ estimate may not suffer much from dilution.

A key assumption for identifying causal effects from difference-in-differences approaches is that retention rate trends in the treated and comparison groups would have been parallel in the absence of the retention payments (the ‘parallel trends’ assumption). This cannot formally be tested, but it is conventional to, where there is enough pre-intervention data available to make it feasible, assess whether trends in the outcome variable were parallel in the two groups prior to intervention. However, we are unable to conduct such tests. First, this is because we use a triple-differences approach with no single comparison group to compare the trends with. The comparison group is a hybrid combination of three groups: teachers of eligible subjects in ineligible areas, teachers of ineligible subjects in eligible areas and teachers of ineligible subjects in ineligible areas. Second, as explored further below, the timing of the retention payments overlapped with one another to some extent and the ‘before’ and ‘after’ periods for our respective difference-in-differences approach overlapped to a large extent. This introduces complexity when attempting to test parallel pre-trends as it is challenging to identify a ‘clean’ comparison. Finally, triple differences is often used in the literature as a robustness check *in case of* parallel trends, or as a remedy where non-parallel pre-trends are identified, arguably negating the need for formal testing of parallel pre-trends.

2.1.2. Data

The main data source we use is teacher data from the School Workforce Census (SWC). The SWC is an individual-level data collection of all teachers working in state-funded schools in England. The SWC has been collected since 2010/11 and we use data from 2015/16 to the most recently available data at the time of the analysis from 2023/24. We also use data about teachers’ initial teacher training (ITT) from linked ITT performance profiles data.

The key outcome measure is retention in the state-funded sector. To measure this, we identify teachers who were present in a particular census and code them according to whether they were present in the following census. If a teacher moves to a different state-sector school but remains present in the data then they are counted as retained. If they are not present, then they are very likely to have left teaching in the state-funded sector. There is a small possibility that they did not leave, but their record was not collected. This is known to be true in a small number of cases, which DfE corrects by comparing census records with pension records. However, this is not likely to contribute to any bias in our estimates. Also, teachers who leave the state-sector may still be teaching in the independent sector, further or higher education or another country.

The SWC captures information about teachers’ personal characteristics (e.g. age, sex, ethnicity), employment (e.g. role, contract type, working pattern), career (e.g. ITT route, years since qualification), subjects taught and qualifications (e.g. ITT subject, degrees). Retention payments were only available for secondary teachers, so we only retained data for teachers working in secondary schools in our sample.

Generally, the data has a very high coverage of the workforce and a low degree of data missingness. However, data on which subject a secondary teacher teaches was not collected from around a third of secondary schools due to the automated way that element of the data is collected from schools by DfE not aligning with some schools' systems. As subject taught was a key variable required for identifying eligibility, all teachers with missing subject data were dropped from the sample. The linked ITT data was available for cohorts 2009/10 to 2022/23. As ITT subject was also a key variable required for identifying eligibility, all teachers with missing ITT data were dropped from the sample. However, the ITT data covers all of the cohorts who were eligible for retention payments, as well as many of earlier cohorts who form part of the before-treatment comparison in our difference-in-differences approach.

The data also identifies the teachers' school ID, which we use to link school information such as school type, proportion of pupils eligible for free school meals and Ofsted rating from various sources including the DfE's Get Information About Schools database, pupil statistics and Ofsted management information.

To identify eligibility for retention payments and payment amounts, we incorporated data capturing key policy information gathered from DfE policy documents (see Department for Education (2017, 2018, 2019, 2022)). We relied on variables in the SWC or school data to code these into information for our estimation:

- Where a retention payment was conditional on a teacher being a specific number of years into their career, we used years since qualification as a proxy.
- Where a retention payment was conditional on subject taught, we used data collected from the SWC curriculum module. As noted above, this meant dropping some data as it does not have complete coverage. As it is a snapshot data collection, the subject data does not capture any in-year changes to class allocations. It may therefore misallocate some teachers' eligibility by not capturing these changes, but we believe these cases are likely to be rare and unlikely to introduce any bias.
- Where a retention payment was conditional on qualification subject, we used subject codes to identify relevant school subjects each degree linked to. These were mapped to school subjects using a qualification mapping framework used by DfE.
- Where a retention payment was conditional on local authority (LA) area, we used the LA code attached to the employing school available in the data.
- Eligibility for LUP was based on the teacher's school's decile of proportion of pupils eligible for pupil premium and whether the school was in an eligible local area. The list of eligible schools was published by DfE, so we matched this to our data using school ID to directly identify corresponding eligibility and payment amount.

We conducted extensive data cleaning to generate our sample for the analysis. This involved identifying a set of 'in-scope' secondary schools and teachers who worked at those schools. Further details on the data cleaning steps we used in the research can be found in Appendix A.

2.1.3. Regression model estimation

We implemented our difference-in-differences approach by estimating a logistic regression model. Retention is a binary dependent variable (taking values of 0 or 1), so the logistic link function reflects this. Implementing a difference-in-differences model using a binary dependent variable has been questioned by some academics, with a preference for a linear probability model expressed by Wooldridge (2010). We test our model for this specification and find it yields very similar results (see Appendix A).

Other researchers exploring the effectiveness of retention payments have used Cox regression models instead of logistic models (e.g. see Sims and Benhenda (2022); CFE Research and FFT Education Datalab (2023)), but the literature suggests the differences in estimates from these respective model types tend to be minimal (Annesi, Moreau and Lellouch, 1989; Ingram and Kleinman, 1989).

As the timing of the retention payments overlapped with one another to some extent – and the 'before' and 'after' periods for our respective difference-in-differences approach overlapped to a large extent – we estimated the effects for each payment within a single model, rather than individual models. This allowed us to tease out and isolate the impacts of individual payments over and above the impact of others.

However, this model set up introduced complexity and necessitated careful construction to ensure it avoided pitfalls of difference-in-differences with multiple time periods (Callaway and Sant'Anna, 2021). To avoid inadvertently and erroneously using not-yet-treated teachers who could anticipate being treated in the future in our comparison group, we coded such teachers as being eligible (with the payment amount coded as zero). Similarly, to avoid inadvertently and erroneously using post-treated teachers in our comparison group, we separately estimated post-treatment effects. In some cases these were meaningful parameters of post-treatment effects on retention, which we report. However, in other cases they reflected particular circumstances (e.g. picking up where teachers had moved school from an eligible to an ineligible local area) so were included in the estimation as incidental parameters to ensure these teachers were not included as part of the comparison group, but are not reported.

We estimated our difference-in-differences and triple differences approach by including eligibility control variables and an impact estimation variable. Taking MPRP as an example, we included:

- an indicator of whether a teacher taught and had relevant qualifications in an eligible subject, to account for the general tendency for maths and physics teachers to have above-average leaving rates
- an indicator of area eligibility, to account for underlying differences in leaving rates between eligible and ineligible local authority areas
- year fixed effects, which accounted for general changes in retention rates over time in ways that affected all teachers similarly, such as the Covid-19 pandemic reducing leaving rates in 2020 and 2021
- interaction terms between subject eligibility and years, to account for any subject-specific variation over time in leaving rates

- interaction terms between area eligibility and year, to account for any area-specific variation over time in leaving rates
- interaction terms between subject and area eligibility, to account for any subject-specific differences in leaving rates between eligible and ineligible areas.

Finally, an interaction term that identified teachers who were eligible for the payment according to their subject, area and the year was entered into the model to estimate the impact of the payment on retention.

We took a similar approach for all the payments, including eligibility criteria and all relevant interactions to estimate difference-in-differences or triple differences estimates of their impact. We began by estimating a fully disaggregated model in which each retention payment was estimated separately, including treating payment schemes with more than one payment amount as separate payments. Each payment had a separate set of controls, as described above.

We then proceeded to estimate more aggregated models in which the indicators used to estimate the impact of each payment were combined regardless of payment level. We first estimated an aggregated model in which the impact of the five different payment schemes were estimated separately. Finally, we estimated a model in which we aggregated all retention payments to estimate the overall impact of ECRPs. However, when we aggregated the variables used to estimate the impacts, we retained the disaggregated set of underlying control variables and interaction terms that were based on treating each payment level as a separate payment. This approach retained the granularity of variation that the control variables were accounting for, while increasing the associated sample size and precision for estimating the payment impacts. For the model estimating the overall impact of ECRPs across schemes, we calculated an aggregated average payment amount. Where a teacher was eligible for more than one payment in a year, we assumed that they received the higher of the payments, except for TSLR, which we assumed teachers received in addition to other payments. This is reflected in Table 11 in Appendix A.

We also included fixed effect controls for ITT cohort and years of experience, to account for differences in retention rates between cohorts and different experience levels. In our preferred model specification, we also included a range of teacher and school characteristics known to be associated with retention rates, including age, sex, ethnicity, working pattern, ITT subject, ITT route, region, school type, trust status, pupil deprivation level and Ofsted rating. The model includes data for many of the same teachers in different years, so we cluster the standard errors at the individual teacher level.

2.1.4. Sensitivity analysis

We estimated a range of specifications of our baseline regression model to check how sensitive the model was to the definition of the sample, exploring the inclusion of different sets of years, subjects within the comparison group and sets of covariates. We also examined whether different definitions of our key eligibility criteria could have been driving our results. We did this to try and understand some of the differences between our estimates and similar estimates from the literature. We show the full set of sensitivity analyses we conducted in Appendix A.

2.1.5. Estimating parameters

The main impact parameter estimated from our logistic regression model is the difference in log odds of leaving state-sector teaching associated with eligibility for a retention payment. This is not a parameter that is easy to understand the meaning of, so we estimate a range of parameters to illustrate what the estimates mean for retention. We first estimate odds ratios, which are commonly used in research studies with binary outcomes. If the odds ratio is below one it means teachers who received retention payments are generally less likely to leave than otherwise similar teachers who did not, and vice versa if the odds ratio is above one. We calculate standard errors and 95 per cent confidence intervals around the odds ratio to assess statistical significance.

However, the meaning of odds ratios is also not easy to understand. We therefore also estimate the implied leaving rates among treated and comparison teachers to illustrate what impact the retention payment the model is suggesting it had on retention. We first calculate the leaving rate of teachers who received the payments in the years they were eligible. We then used the estimated odds ratio to calculate an implied leaving rate among otherwise similar teachers in the comparison group. This enables us to establish estimates of the extent to which retention payments are associated with changes in leaving rates, in both percentage point differences (i.e. the leaving rate reduced from X per cent to Y per cent) and proportional terms (i.e. the leaving rate fell by Z per cent). We estimate standard errors and confidence intervals using the delta method.

We also use our estimates to calculate elasticities, a key value for money parameter used to assess how responsive teachers are to changes in the financial attractiveness of a particular choice. The payment elasticity of attrition measures the percentage change in leaving rate associated with a one per cent change in overall remuneration (including salary). We therefore combine estimates of the proportional change in retention with an estimate of how much the retention payment is adding to total pay.

To calculate the latter we use data on gross teacher pay (including additional pay, such as teaching and learning responsibility (TLR) payments) for teachers eligible for each ECRP from the SWC. A limitation with SWC pay data is that, due to the timing of pay increases being confirmed to schools, it can often be reported with a lag, which means our measure of average pay could have understated a teacher’s actual earnings that year. We do not account for the fact that the DfE pays the tax and national insurance contributions (NIC) on ECRPs, while teachers pay tax and NIC on their gross salary. This was primarily due to the additional complexity of estimating tax and NIC that would be payable on gross salary. We estimate standard errors and confidence intervals for the elasticities by using those estimated for the proportional change in leaving rates, described above.

2.1.6. Heterogeneity analysis

To explore research question 2, we undertook heterogeneity analysis to examine whether the impact of ECRPs differed by characteristics. Based on the aggregate model that estimated the overall impact of ECRPs on retention described above, we estimated a set of separate regression models that included interaction terms between the impact estimate and teacher-, school- and area-level characteristics. We examined heterogeneity in sex, years of experience, age, ethnicity, working pattern (full time or part time), subject, degree class, ITT route, school FSM quintile, education investment area (EIA) status of the school, Ofsted rating and region.

We assessed statistical significance of the heterogeneity by determining whether the interactions of the triple-difference term and the set of characteristics in question led to statistically significantly different results (using a joint test where there was more than one category of the characteristic). We then estimated predicted probabilities of leaving separately for eligible and non-eligible teachers within each category of the characteristic (e.g. separately for eligible and non-eligible female teachers and again for male teachers). We also estimated average pay and ECRP payment amount for eligible teachers across each heterogeneity characteristic category, using this to estimate elasticities for each group. We summarise the heterogeneity findings in section 5 and show the full set of results in Appendix A.

2.2. Limitations

As noted above, the key limitations of our analysis are that:

- teachers we believe were eligible for an ECRP might not actually have been and teachers who we believe were in the comparison group might actually be eligible. We closely followed the published eligibility criteria to minimise this risk, but this could dilute the sample and reduce our estimated impact.
- teachers may have been eligible (and identified in our analysis as eligible) but not claimed the payment. We understand from DfE that take-up rates were generally high so this is unlikely to be a major issue, but it could serve to dilute the sample and our estimated impact.
- missing data means we lose some teachers from the sample where we cannot observe their teaching subject, ITT course or qualification. This is unlikely to introduce any bias but reduces the sample size and thereby the precision of our estimates. Less precision means larger standard errors, wider confidence intervals and the results of a particular size of impact being less likely to be statistically significant.
- teacher salary data may be lower than actual because of lags to schools updating salary information in time for the SWC data collection. This may impact our elasticity estimates but is unlikely to introduce bias.
- we do not account for the fact that ECRPs are tax-free payments made to teachers while their salary is taxable, nor model complexities relating to pensionable income. This could affect our elasticity estimates but is very difficult to model robustly.

2.3. Descriptive analysis of the sample

Table 2 summarises the number of teachers in our estimation sample that were eligible for each payment scheme. MPRP has the smallest number of eligible teachers, as it was limited to maths and physics teachers, a small geographical area and was active for two years. TSLR has more eligible teachers as it was applied to a range of subjects, several cohorts of trainees and a larger geographical area. LUP has the largest number of eligible teachers, since its eligibility covered four subjects and a large number of schools.

Table 2 Number of eligible teachers in estimation sample, by scheme

Group	Number of teachers
MPRP	1,314
TSLR	4,420
MPB	2,680
ECP	2,770
LUP	9,244

Note: includes counting teachers who were eligible for more than one payment scheme against each one.

Source: NFER analysis of ITT-PP and SWC data.

Table 3 shows the extent to which teachers experienced being eligible for one scheme or more than one scheme within the estimation sample, whether that was in the same year or in different years. Around three in ten (29 per cent) eligible teachers were eligible for two schemes, while small number (333 or 2 per cent) were eligible for three. This emphasises the challenge experienced in designing an evaluation approach that teased out the respective impacts of these schemes from one another, described above. The size of the comparison group was large, with around 105,000 secondary teachers available to estimate the counterfactual, although most taught subjects that were never eligible for any of the ECRP schemes. Our evaluation approach is carefully designed to take account of underlying differences between subjects that might be associated with retention rates but unrelated to ECRPs. However, our sensitivity analysis also tested the sensitivity of the model to restricting this comparison group in various ways (see Appendix A).

Table 3 Number of teachers in estimation sample, by ECRP eligibility status

Group	Number of teachers
Eligible for one scheme	9,515
Eligible for two schemes	3,957
Eligible for three schemes	333
Total eligible for at least one scheme	13,805
Comparison group	104,293
Total sample	118,098

Source: NFER analysis of ITT-PP and SWC data.

Table 4 shows the average payment amount received by teachers who were eligible for each scheme and the proportion of their concurrent gross salary that the payment amount represented. This is the payment amount we assume teaches received on the basis of the stated eligibility criteria, but this has not been verified against management information on payments made. As noted above in the limitations, this may not reflect who actually received payments, potentially diluting our estimates.

It is important to note that MPB and ECP average payment amounts have been calculated across the entire period of teachers' eligibility, including during years when teachers were eligible for a future payment (if they remained working at an eligible school) but did not receive a payment in that year. TSLR payments were the smallest, because they were linked to student loan repayment amounts. Overall, ECRPs represented around seven per cent of eligible teachers' concurrent salary, representing a substantial uplift to teachers' overall remuneration.

Table 4 Average payment amount and proportion of salary, by ECRP scheme

ECRP scheme	Average payment amount when eligible (£)	Proportion of concurrent salary (%)	Number of eligible teacher-years in sample
MPRP	2,000	7.0	1,941
TSLR	325	1.0	5,359
MPB (£5,000)	1,382	4.7	5,723
MPB (£7,500)	1,979	7.3	1,258
MPB (combined)	1,490	5.1	6,981
ECP (£2,000)	1,042	3.7	3,567
ECP (£3,000)	1,479	5.5	781
ECP (combined)	1,121	4.0	4,348
LUP (£1,500)	1,500	4.4	1,205
LUP (£2,000)	2,000	6.2	1,865
LUP (£2,500)	2,500	8.1	649
LUP (£3,000)	3,000	10.0	575
LUP (combined)	2,069	6.4	4,294
ECRPs (combined)	2,257	7.1	23,346

Note: MPB and ECP average payment amounts have been calculated across the entire period of their eligibility, including during years when teachers were eligible for a future payment (if they remained working at an eligible school) but did not receive a payment in that year. 'Teacher years' counts the number of eligible teachers per year, so counts teachers who were eligible for payments in multiple years more than once.

Source: NFER analysis of ITT-PP and SWC data.

3. Impact of ECRPs on retention

This section presents our findings on the impact of ECRPs on retention, based on the estimates from our difference-in-differences and triple differences approach described in section 2. The estimates are presented in terms of odds ratios and leaving rates. Odds ratios represent the change in the odds of leaving associated with being eligible for a retention payment, relative to an otherwise similar teacher not being eligible. An odds ratio of less than one suggests that payment eligibility was associated with a lower-than-otherwise leaving rate, while an odds ratio of greater than one suggests that payment eligibility was associated with a higher-than-otherwise leaving rate. Odds ratios are displayed on a logarithmic scale to account for them representing proportions. Confidence intervals (95 per cent) are shown in the charts. Where the confidence interval overlaps the ‘1’ line, this suggests that the estimated difference is not statistically significant at the five per cent level.

3.1. Impact of retention payment schemes at different payment levels

Table 5 shows the estimated impact on retention of each ECRP, treating different payment levels as distinct payments. As described in the methodology, each impact estimate is supported by sets of incidental parameters that control for between-group differences and change over time, enabling the estimate to be interpreted as a difference-in-differences or triple differences estimate of the impact on retention. Figure 1 shows the same odds ratios visually, with accompanying 95 per cent confidence intervals.

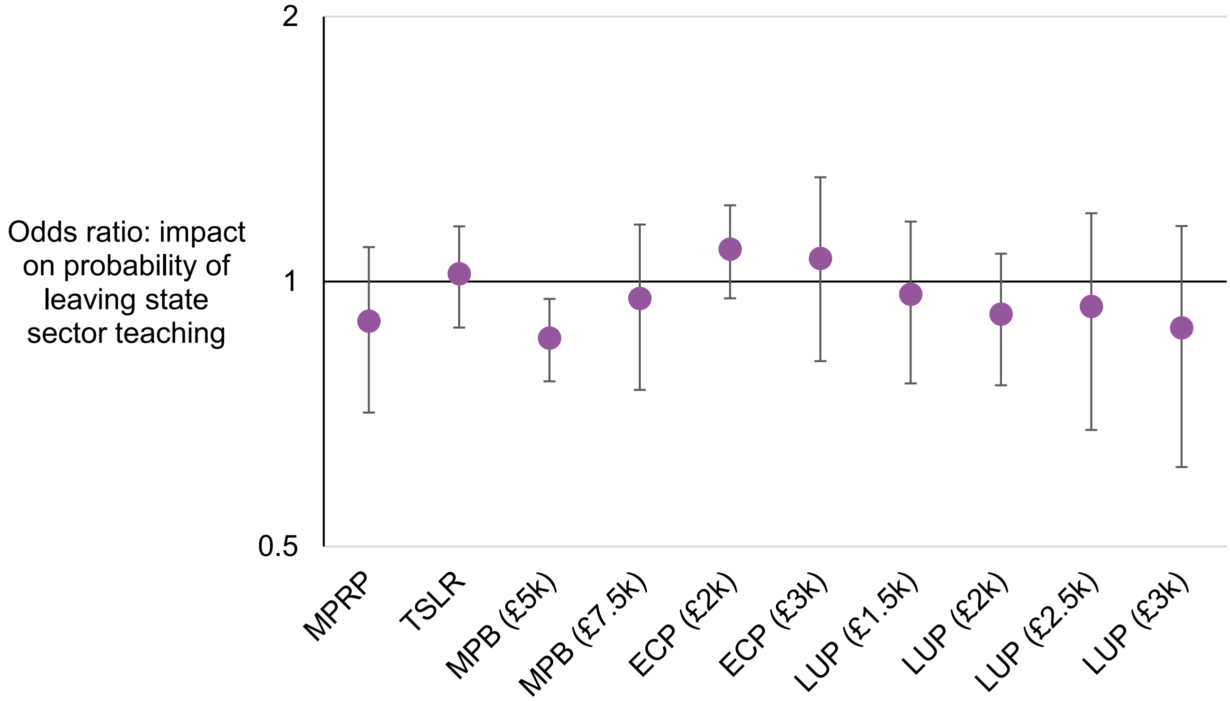
Table 5 Most ECRPs were associated with slightly lower leaving rates, but only the lower-value MPB payment was statistically significant

Retention payment scheme	Odds ratio	Standard error	Significant?
MPRP	0.90	0.10	No
TSLR	1.02	0.07	No
MPB (£5k)	0.86	0.05	Yes
MPB (£7.5k)	0.96	0.10	No
ECP (£2k)	1.09	0.07	No
ECP (£3k)	1.06	0.13	No
LUP (£1.5k)	0.97	0.10	No
LUP (£2k)	0.92	0.08	No
LUP (£2.5k)	0.94	0.13	No
LUP (£3k)	0.89	0.14	No
Difference-in-differences controls?	Yes		
Teacher and school characteristics?	Yes		
Post-treatment effect controls?	Yes		

Note: dependent variable is whether or not the teacher left the state-funded sector in the following year. Statistical significance of the difference of the odds ratio from one is assessed at the five per cent level.

Source: NFER analysis of ITT-PP and SWC data.

Figure 1 Most ECRPs were associated with slightly lower leaving rates, but only the lower-value MPB payment was statistically significant



Note: The confidence intervals shown are 95 per cent confidence intervals. Statistical significance of the difference of the odds ratio from one is assessed at the five per cent level.

Source: NFER analysis of ITT-PP and SWC data.

Overall, seven out of ten of the odds ratios are estimated to be below one, suggesting that they are more likely to have been associated with lower teacher leaving rates than higher leaving rates. Three of the estimates (TSLR and both payment levels for the ECP scheme) are estimated to be above one, suggesting they were more likely to have been associated with higher teacher leaving rates. However, only one of the estimates is statistically significant at the five per cent level. This suggests that there is a high degree of imprecision associated with the estimates and we have low confidence that any individual payment was associated with either an improvement or deterioration in retention.

The estimated odds ratio of the impact of MPRP is 0.90, but the difference from one is not statistically significant. This differs from the estimate of an evaluation of the same scheme by Sims and Benhenda (2022), which found that MPRP was statistically significantly associated with a lower teacher leaving rate. We explore the reasons why our respective estimates differ in Appendix A. In short, differences between the respective analyses in terms of how eligibility was defined, which years were included and which pre-intervention cohorts were included in the analysis explain these substantial differences, with the latter making the most substantial difference. While Sims and Benhenda used a Cox regression model whereas we use a logistic regression model, our ability to replicate the results when using similar definitions suggests the difference in model type is unlikely to be a factor. Both of the approaches to cohort inclusion used by the respective research teams are defensible approaches to undertaking the same analysis. This suggests that any estimate of the impact on retention is likely to be subject to uncertainty that depends on how the analysis is undertaken, in addition to a high degree of estimation imprecision relative to the size of effect due to the limited available sample size.

The estimated odds ratio for TSLR is 1.02, just above one but not statistically significantly different from one. This is outside the range of estimates in the DfE's analysis of the impact of TSLR on teacher retention, which varied between 0.95 to 0.82 (CFE Research, 2023). However, the wide confidence interval associated with both our estimates and those in DfE's analysis suggests they are unlikely to be significantly different from one another. While we haven't conducted a similar replication analysis for TSLRs, it seems likely that uncertainty depending on how the analysis is undertaken may also play a role in explaining the variation (as further demonstrated by the wide range of estimates in DfE's analysis).

The estimated odds ratio for the lower-value payment for MPB is 0.86 and the difference from one is statistically significant. This suggests that the lower-value MPB is associated with a lower teacher leaving rate than otherwise similar teachers, in line with the findings from CFE Research and FFT Education Datalab (2023). The estimated odds ratio for MPB is larger (relative to one) for the lower-value payment (0.86) than for the higher-value payment (0.96). However, the underlying sample size for estimating the higher-value payment is considerably lower, meaning that the level of precision is also lower. The difference between these two impact estimates is not statistically significant.

The estimated odds ratios for the ECP payments are both above one, suggesting that these payments are associated with a higher leaving rate than otherwise similar teachers. However, neither estimated odds ratio is statistically significantly different from one, suggesting we cannot be confident that it had an effect of worsening retention. Nevertheless, an estimate of greater than one is counterintuitive to the hypothesis that making additional payments to teachers would be likely to improve retention. There is no *a priori* reason to expect that ECP would have a different effect on retention to any other payment type. The payments were deferred for certain cohorts (i.e. ITT cohorts could anticipate being eligible once they got to their second, third and fourth years of teaching), but this was also the case for MPB, which was associated with a significant improvement in retention. ECP and TSLR were the only payments for which languages teachers were eligible, but our heterogeneity analysis did not suggest that languages teachers had significantly different behaviour from teachers of other subjects (see heterogeneity analysis in section 5). This result could be driven by unobserved factors or simply down to chance.

The estimated odds ratios for the four levels of the LUP are all below one, although none of the estimated differences are statistically significant. The odds ratios for different payment levels tend to be slightly further below one for higher payment values but are not statistically significantly different from one another.

3.2. Impact of retention payment schemes at scheme level

The findings summarised in section 3.1 hint that retention payments may be associated with lower retention, but with a low level of confidence given the relatively low estimation precision associated with each estimate. Each payment scheme's estimate was based on a limited sample size, varying between 575 to 5,000 teacher-years. To gain further sample size and precision, we combine the estimation of impact to overall scheme level, combining the estimates at different payment levels into one. Further, we estimate an overall combined model that estimates the odds ratio of the combined impact of all ECRPs, across all schemes and payment levels.

Table 6 shows the estimated odds ratios and standard errors of these combined estimates, while Figure 2 summarises them visually along with 95 per cent confidence intervals.

Table 6 ECRPs overall are associated with lower leaving rates, but the difference is not statistically significant at the five per cent level

	Odds ratio	Standard error	Significant?
Scheme-by-scheme combined model			
MPRP	0.91	0.10	No
TSLR	1.02	0.07	No
MPB	0.88	0.04	Yes
ECP	1.08	0.06	No
LUP	0.93	0.06	No
Overall combined model			
ECRPs	0.94	0.03	No

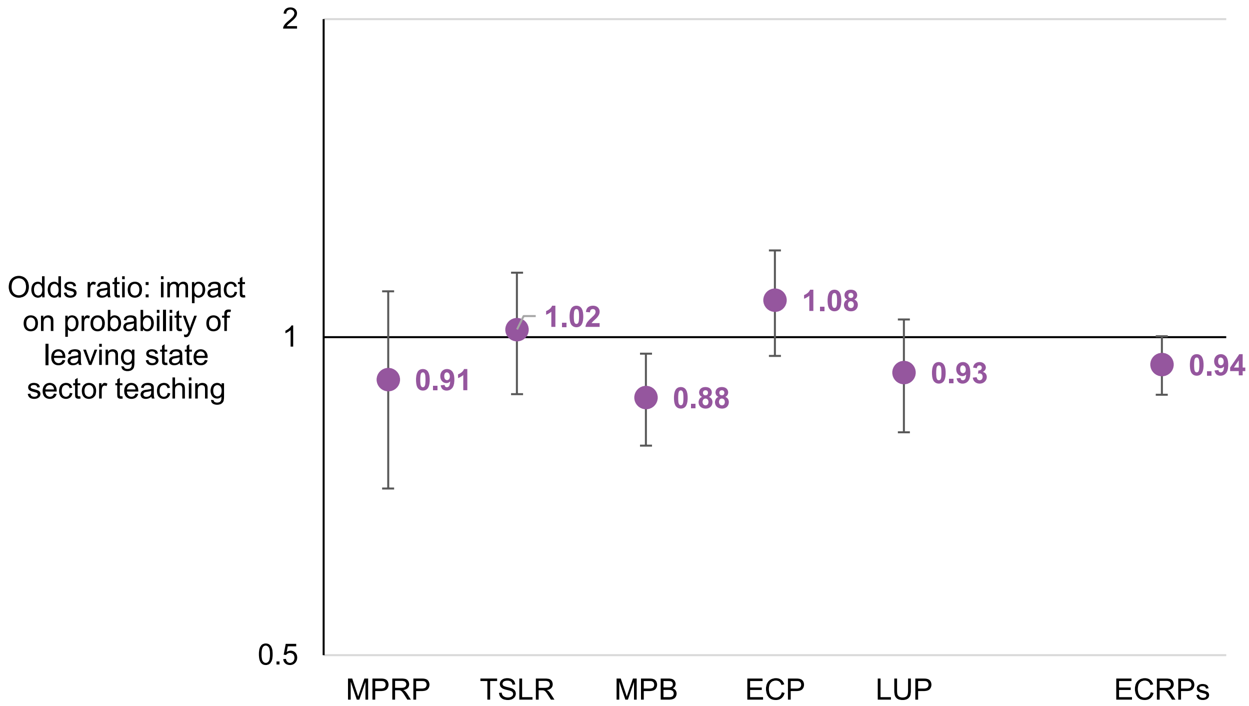
Note: Statistical significance of the difference of the odds ratio from one is assessed at the five per cent level.

Source: NFER analysis of ITT-PP and SWC data.

The estimates for MPRP and TSLR from the combined model are very similar to the disaggregated model, which is unsurprising given that they did not have differential payment levels. The estimated odds ratios of the combined impacts of MPB, ECP and LUP are similar to the respective underlying estimates separated out by payment level but estimated with more precision as the sample sizes are larger. The combined impact of MPB across payment levels has an estimated odds ratio of 0.88 and the difference from one is statistically significant. The combined impact of ECP across payment levels has an estimated odds ratio of 1.08, which is greater than one (i.e. associated with a higher teacher leaving rate) but is not statistically significant. The combined impact of LUP across payment levels has an estimated odds ratio of 0.93 but the difference from one is not statistically significant.

The overall estimated odds ratio for all ECRPs combined is 0.94. The difference is not statistically significantly different from one at the five per cent level, having a p value of 0.06. This suggests that ECRPs overall are associated with a slightly lower teacher leaving rate compared to otherwise similar teachers who did not receive payments, but this may be down to chance as it does not meet the conventional threshold for assessing statistical significance.

Figure 2 ECRPs overall are associated with lower leaving rates, but the difference is not statistically significant at the five per cent level



Note: The confidence intervals shown are 95 per cent confidence intervals. Statistical significance of the difference of the odds ratio from one is assessed at the five per cent level.

Source: NFER analysis of ITT-PP and SWC data.

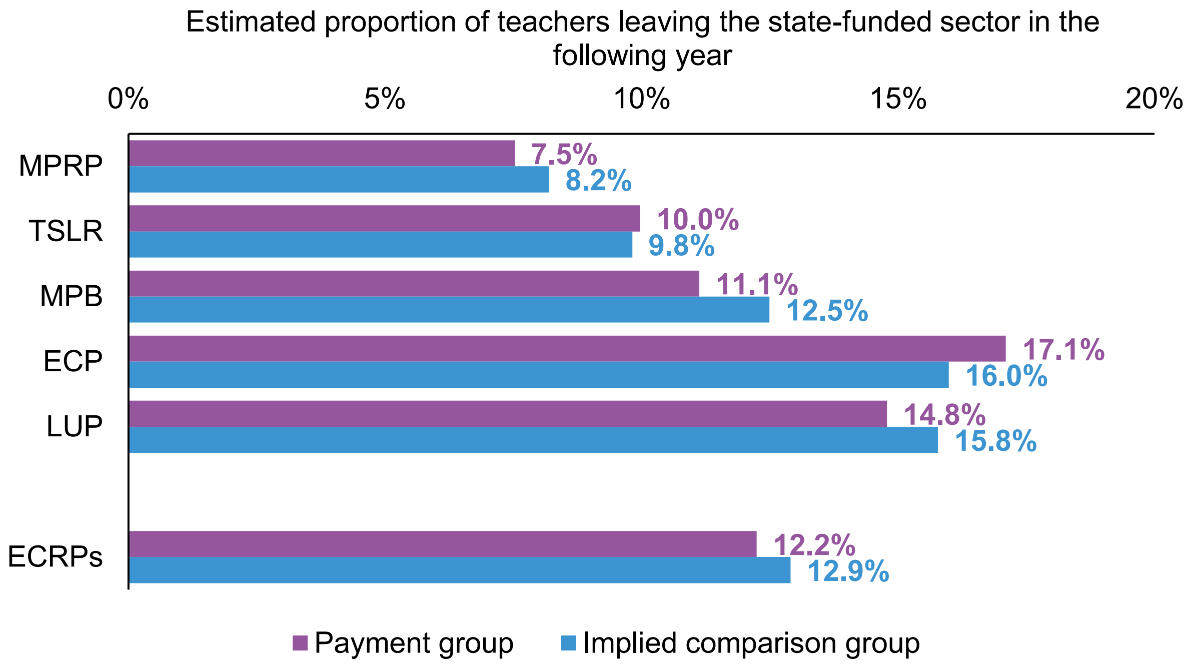
3.3. Impact on retention rates

Figure 3 shows the annual rate of teachers leaving the state-funded sector for the group of teachers who were eligible for each payment and its respective comparison group. The leaving rates for the payment group are the rate among eligible teachers during the time of eligibility, while the comparison group leaving rate is estimated from the odds ratios summarised above. This is because of the multi-faceted nature of the triple-differences comparison group, where there is no one distinct group of comparison teachers.

The annual leaving rates differ considerably between the payment schemes, in part because of different impacts of the payments, but mostly because of background and contextual factors. For example, the leaving rates for the MPRP teachers and associated comparison group are lower than the others, because it was implemented during the Covid-19 pandemic, when leaving rates among all teachers were lower. Likewise, some of the TSLR and MPB samples were eligible during the Covid-19 pandemic, as well as after. LUP eligibility in our sample occurred only in 2022/23, when the effects of the pandemic on retention were less and leaving rates were higher than they had been during the pandemic.

The data echoes the findings on odds ratios, indicating that the change in retention rates associated with ECRPs was variable, from a 1.1 percentage point increase in leaving rates for ECP (although not statistically significant) to a 1.4 percentage point decrease for MPB. LUP was associated with a one percentage point decrease in the leaving rate, although was not statistically significant. The estimated overall impact of ECRPs translated into a 0.7 percentage point reduction in the leaving rate, from 12.9 to 12.2 per cent, although this was not statistically significant.

Figure 3 Annual leaving rates differ considerably between payment schemes, mostly because of background and contextual factors



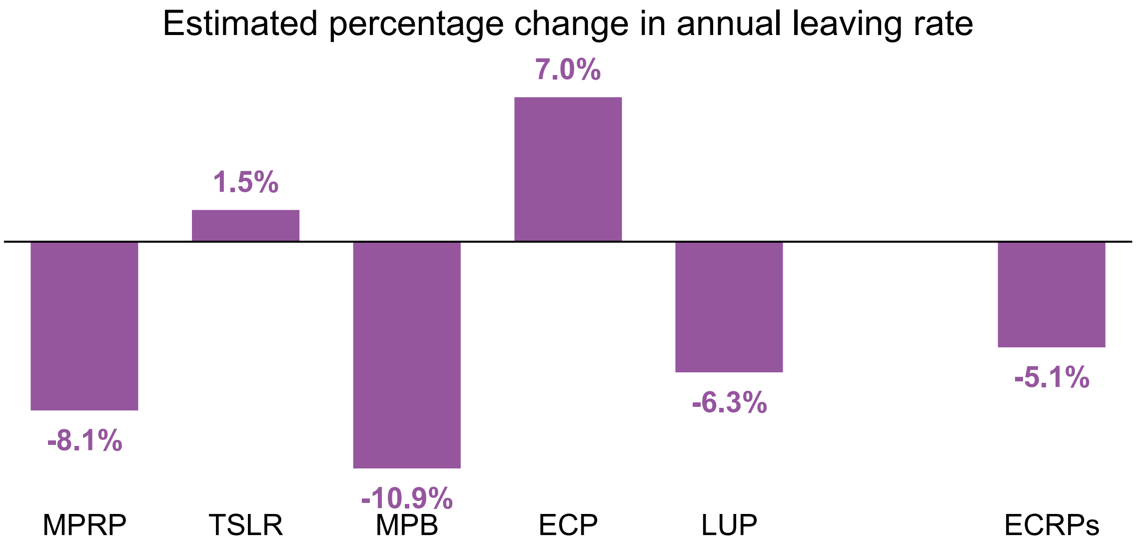
Source: NFER analysis of ITT-PP and SWC data.

Figure 4 expresses these changes in leaving rate in proportional terms, showing that the changes in leaving rate varied from a seven per cent per year increase for ECP to a 10.9 per cent per year decrease for MPB.

Accounting for the fact that the average MPB-eligible teacher was eligible for 2.6 years in our sample, this equates to MPB being associated with an overall 29 per cent reduction in attrition per teacher, similar to the finding of CFE Research and FFT Education Datalab (2023) of a range of 31-38 per cent reduction. The average MPRP-eligible teacher was eligible for 1.5 years in our sample, meaning the scheme was associated with an overall 12 per cent reduction in attrition per teacher, although this was not statistically significant.

The estimated overall impact of ECRPs translated into a 5.1 per cent per year reduction in the leaving rate, although this was not statistically significant.

Figure 4 ECRPs were associated with a 5.1 per cent per year reduction in the leaving rate, although the difference was not statistically significant



Source: NFER analysis of ITT-PP and SWC data.

3.4. Post-eligibility effects

A concern about financial incentives that is often expressed is that while payments may have an impact while recipients are eligible, that effect may be reversed once recipients are no longer eligible, limiting the long-term impact on supply and affecting the longer-term value for money (See *et al.*, 2020). However, if financial incentives have a positive immediate impact but little reversal effect then they can be associated with long-term impacts on supply and enhanced value for money (e.g. see the evidence on training bursaries in McLean, Tang and Worth, (2023)).

We explore post-eligibility effects by primarily looking at them for MPRP, since it is the only payment scheme to end eligibility during our analysis period. We include indicators in the regression model that identify teachers who received a payment but are no longer eligible for that payment. As noted in the methodology section, including these terms also ensures that the post-eligible teachers are not erroneously included in the comparison group. Payments for MPRP were made in 2019/20 and 2020/21, meaning that a group of teachers in their fifth year of teaching in 2019/20 received a payment but were subsequently ineligible. Likewise, all original MPRP recipients did not receive a payment through MPRP in 2021/22 or 2022/23.

Table 7 summarises the findings from this post-eligibility analysis. The combined impact of post-eligibility is estimated to have an odds ratio of 1.11 but is not statistically significant. This slight increase in the leaving rate is similar in magnitude to the main estimated impact of MPRP but in the opposite direction. However, a breakdown of the post-eligibility impacts by year shows no consistent pattern. Overall, the sample sizes mean that the level of confidence associated with these findings is very low, so we cannot conclude anything definitive about the extent of reversal effects. Future research should explore this question further, by analysing the extent of post-eligibility effects for larger-scale schemes such as LUP.

Table 7 There is mixed and inconclusive evidence on the extent of reversal effects on retention after teachers stopped being eligible for MPRP

	Odds ratio	Standard error	Significant?
MPRP post-eligibility – combined	1.11	0.10	No
MPRP post-eligibility – one year since last payment	1.18	0.16	No
MPRP post-eligibility – two years since last payment	0.92	0.12	No
MPRP post-eligibility – three years since last payment	1.30	0.22	No

Note: Statistical significance of the difference of the odds ratio from one is assessed at the five per cent level.

Source: NFER analysis of ITT-PP and SWC data.

3.5. Sensitivity analysis

We estimated a range of specifications of our baseline regression model to check how sensitive the model was to the definition of the sample, exploring the inclusion of different sets of years, subjects within the comparison group and sets of covariates. We also examined whether different definitions of our key eligibility criteria could have been driving our results. We did this to try and understand some of the differences between our estimates and similar estimates from the literature. We show the full set of sensitivity analyses we conducted in Appendix A.

4. Value for money of retention payments

The findings in section 3 relate to the impact that different retention payments, payment schemes and retention payments overall had on teacher retention. However, the size of the payments eligible teachers received differed considerably by scheme, as shown in Table 4. In this section, we combine the findings on impact with those on cost to analyse how responsive teachers are to financial incentives as a proportion of their salaried income and explore the policy implications for value for money.

4.1. Elasticities

An elasticity is a concept from economics, which measures how responsive an individual's decision or behaviour is due to a change in a factor. In this case, we explore the payment elasticity of retention: the percentage change in the teacher leaving rate associated with a one per cent change in teachers' overall income. A negative elasticity implies the payment is associated with a reduction in the leaving rate, while a positive elasticity implies the payment is associated with a higher leaving rate compared to otherwise similar teachers. Elasticities are a standard value for money metric in the economics literature and can be readily compared to estimates of other interventions and from other evaluations. They are also useful in simulation modelling and policy analysis when considering the relative value for money of competing policy actions.

Payment amounts differing between retention payment schemes makes it particularly important to compare value for money on a comparable scale, as it might be expected that larger payments would lead to larger effects and therefore that comparing the impact of eligibility on retention does not compare like with like. For example, the TSLR payments are considerably smaller than other retention payments, which is important context when comparing estimates of impact.

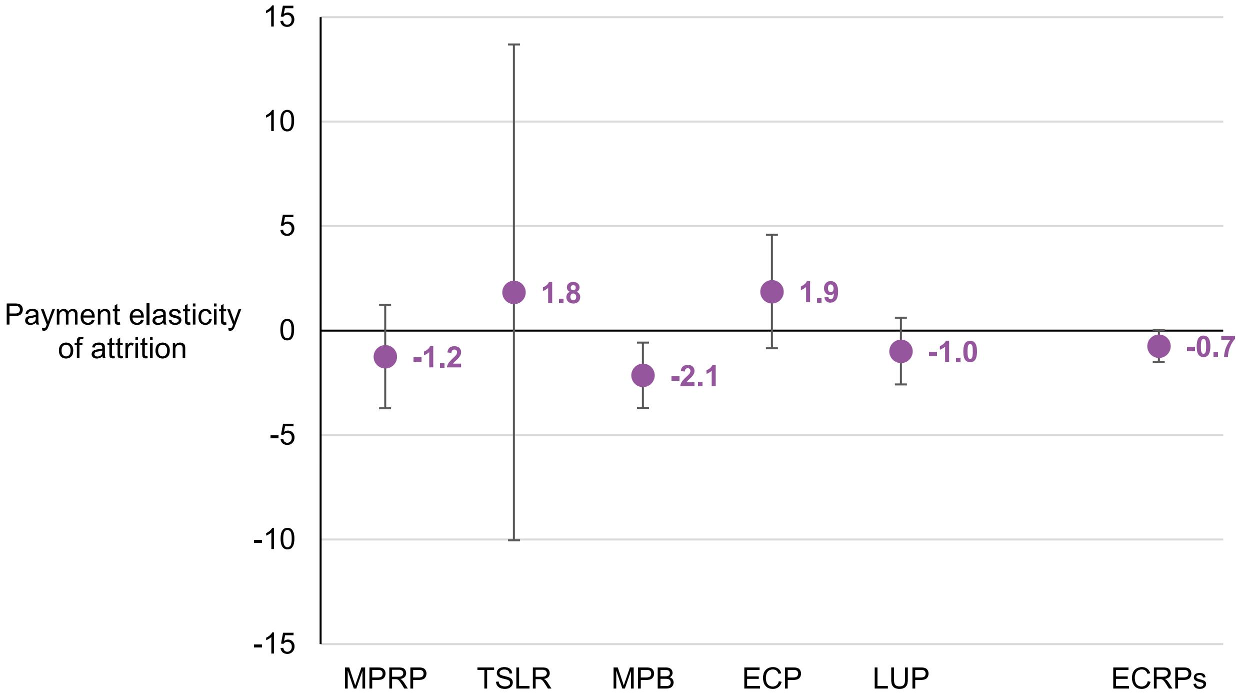
Figure 5 shows estimated elasticities for each payment scheme and a combined elasticity estimate for ECRPs overall. The patterns are somewhat similar to the estimated impacts on retention but differ due to differences in the size of the payments as a proportion of teachers' salary.

The estimated elasticity for MPRP is -1.2 and not statistically significantly different from zero. This is lower than the elasticity estimated by Sims and Benhenda (2022) of -3, which is related to the differences in impact estimate, as explored in Appendix A. The estimated elasticity for TSLR is slightly greater than zero because the odds ratio of impact on retention was above one. However, as with the impact estimate, the confidence interval around this estimate is large, due to the large degree of uncertainty. The estimated elasticity for ECP is also greater than zero because the odds ratio of impact on retention was above one but is not statistically significant.

The estimated elasticity for MPB is -2.1 and statistically significant. This is in line with the estimated elasticities of -2.2 (lower payment) and -2.3 (higher payment) from CFE Research and FFT Education Datalab (2023). The estimated elasticity for LUP is -1 but is not statistically significant.

The overall estimated elasticity for retention payment across schemes is -0.7. As with the impact estimate, the elasticity estimate is not statistically significant from zero at the five per cent level, with a p value of 0.06. However, it is also not statistically significantly different from a DfE estimate of overall teacher pay elasticity of retention, derived from a review of the literature, of -1.5 (DfE, 2020).

Figure 5 The estimated payment elasticity of attrition for ECRPs overall is -0.7, but is not statistically significant at the five per cent level



Note: The confidence intervals shown are 95 per cent confidence intervals. Statistical significance of the difference of the odds ratio from one is assessed at the five per cent level.

Source: NFER analysis of ITT-PP and SWC data.

4.2. Implied long-term cost effectiveness of retention payments

While the estimated difference from zero is not statistically significant, our central estimate of a negative payment elasticity of retention for ECRPs overall implies that retention payments may lead to improved retention. This in turn means that there could be more teachers from a given cohort of teachers still present in the state-sector over the long term than otherwise. In the context of the wider literature, the evidence implies that this is likely to be the case, even though our evidence is, on its own, somewhat inconclusive and suggests that previous estimates of the payment elasticity of retention may be over-estimates.

However, a policy of introducing or increasing early career payments comes with a cost. Given resources are scarce, it is therefore important to assess the cost effectiveness of such a policy action compared to other actions that could otherwise be taken.

In this sub-section we explore the implications of these findings further, illustrating the implied long-term impact of early career payments on teacher numbers and teacher supply. We also compare the teacher supply implications and cost effectiveness of ECRPs to bursaries, another policy measure that has a known impact on teacher supply. Using the same framework as Tang and Worth (2024), we model the longer-term impact of an ECRP scheme using a hypothetical cohort of 100 teachers. We estimate the impact of a retention payment scheme on teacher supply over the period of an entire teaching career. We measure impact through estimating the additional number of teachers per year (or teacher-years) over the career of one cohort compared to the baseline cohort. The baseline is determined using average entry and leaving rates derived from analysis of SWC data. See Tang and Worth (2024) for more details on the methodology.

To evaluate the relative value for money of an ECRP scheme, we compare the impact of the ECRP scheme to that of other financial incentive policies using the same cost envelope. While retention payments are paid to in-service teachers, training bursaries are payments made to trainees while they are completing their course of initial teacher training. We estimate the total lifetime cost (including additional teacher training costs) of increasing a training bursary by £5,000 for a cohort of 100 teachers that have no existing bursary (using estimates on the recruitment and retention impact from (McLean, Tang and Worth, 2023)). We divide the total lifetime cost by the number of additional teacher-years over the career span of the hypothetical cohort to obtain an estimate of the cost per additional teacher-year the policy gains.

We then found what level of ECRP, paid to all early career teachers in the cohort of 100 in each year of their first five years, would yield the same total additional cost for a cohort of 100 teachers entering the profession. We derive this ECRP amount using two different elasticity estimates:

- the first elasticity is the overall elasticity estimated for ECRPs from section 3, which is -0.7.
- the second elasticity is DfE's estimate of pay elasticity of retention from the literature, which is -1.5. Additional support for this elasticity as a reasonable one to assume is that it is situated within the range of estimates from our -0.7 to those of Sims and Benhenda (2022) and CFE Research and FFT Education Datalab (2023) (a range of -2.2. to -3).

Equalising the total cost across scenarios yields an ECRP amount of £2,940 per year under the -0.7 elasticity and £2,888 under the -1.5 elasticity.

We also ran two other bursary scenarios: one with a £30,000 prevailing bursary (as is approximately the case for shortage subjects) and one with a £9,800 prevailing bursary. The value of £9,800 was used as it is an estimate for the weighted average bursary paid in 2023/24. It is therefore a proxy for simulating the effect of increasing the bursary for all subjects (even those with no existing bursary) by the same amount. As for ECRPs, we found the level of training bursary that would total the same additional cost for a cohort of 100 teachers entering the profession. Under these scenarios, the values of the bursary increases offered are £5,000 for no prevailing bursary, £4,381 for a £9,800 prevailing bursary, and £3,473 for a £30,000 prevailing bursary.

Table 8 shows the estimated total lifetime cost per additional teacher from these different scenarios. The lowest cost per additional teacher-year is from increasing bursaries where there is currently no bursary, at around £9,000 per additional teacher-year. This suggests that where a subject has no existing bursary, increasing the bursary is a very cost-effective way to improve teacher supply over the long term. The cost per additional teacher-year is higher where the prevailing bursary is higher, although the cost only reaches around £13,000 at a prevailing bursary of £30,000. In contrast, under the payment elasticity assumptions made in our scenarios, the cost per additional teacher-year for ECRPs is considerably higher.

Under an elasticity of -1.5, the cost per additional teacher-year is around £20,000, while under an elasticity of -0.7 the cost per additional teacher-year is around £43,000. This implies that increasing bursaries should always be considered first since it is highly cost effective, even at high levels of the prevailing bursary.

Table 8 The implied lifetime costs and impact of ECRPs on a cohort of teachers suggest that ECRPs have a higher cost per additional teacher-year than training bursaries

	Estimated lifetime cost per additional teacher-year teaching in the state sector (£)
Bursary increase (prevailing bursary = £0)	8,946
Bursary increase (prevailing bursary = £9,800)	10,209
Bursary increase (prevailing bursary = £30,000)	12,881
ECP (assumed elasticity = -0.7)	43,420
ECP (assumed elasticity = -1.5)	20,406

Source: NFER simulation analysis of ITT-PP and SWC data. See Tang and Worth (2024) for more details on the underlying methodology.

These scenarios suggest that ECRPs offer positive value for money so long as it is infeasible or unreasonable to raise the bursary any further. This is arguably the case for perennial shortage subjects such as maths, physics, chemistry and computing, where the bursary has consistently been just below the level of the starting salary for a teacher and increasing bursaries beyond the level of the starting salary may introduce distorted paths of total teacher remuneration. ECRPs can therefore provide additional scope for improving retention and teacher supply as part of a wider strategy to improve teacher supply and where they are focussed on subjects that already have maximum bursaries.

Evidence on the impact and cost effectiveness of other strategies that are likely to improve recruitment and retention, such as workload reduction, flexible working and high-quality professional development, is unfortunately lower quality, precluding a wider set of cost effectiveness comparisons with other policy measures that could inform a wider teacher recruitment and retention strategy.

5. Heterogeneity analysis

We undertook heterogeneity analysis to examine whether the impact of ECRPs differed according to teacher's personal, role and school characteristics. Based on the aggregate model that estimated the overall impact of ECRPs on retention, described in section 3, we estimated a set of separate regression models that included interaction terms between the impact estimate and teacher-, school- and area-level characteristics. We examined heterogeneity in: sex, years of experience, age, ethnicity, working pattern (full time or part time), ITT subject, undergraduate degree class, ITT route, school FSM quintile, education investment area (EIA) status of the school, Ofsted rating and region.

For each characteristic we conducted a joint test of statistical significance to assess whether the heterogeneity was greater than one might expect purely due to chance. Table 9 shows the p values from these joint significance tests. Most of the characteristics did not display significant heterogeneity, implying that the differences in impact did not meaningfully differ between the characteristics. However, several characteristics did show significant variation, including age, ITT route and region. We summarise the findings below.

Table 9 Statistical tests of heterogeneous impacts by characteristics suggests that not many are associated with significant variation in impact

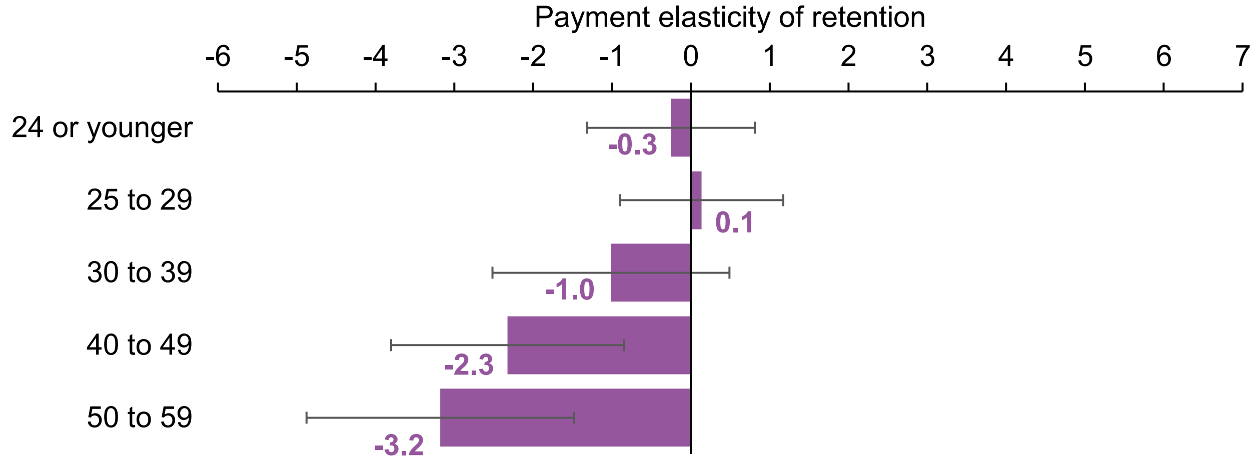
Characteristic	P value from joint significance test of interactions between ECRP impact and characteristics
Sex	0.99
Years of experience in state sector	0.87
Age	0.00 *
Ethnicity	0.08
Working pattern (full time/ part time)	0.08
ITT subject	0.34
Undergraduate degree class	0.51
ITT route	0.00 *
School quintile of FSM eligibility	0.35
Education investment rea	0.22
School Ofsted rating	0.60
Region	0.02 *

Note: * indicates a p value below the conventional five per cent threshold.

Source: NFER analysis of ITT-PP and SWC data.

Figure 6 shows the estimated payment elasticity for early career teachers of different ages. As shown in Table 9, there was significant variation across age. The pattern indicates that while early career teachers in their twenties showed low responsiveness to payments, early career teachers in older age groups showed much higher responsiveness. This suggests that the retention decisions of career changers who enter teaching later in their career are more responsive to retention payments. This echoes previous findings that career changer trainees' decisions about entering teaching are more responsive to training bursaries than for younger trainees (McLean, Tang and Worth, 2023). It's important to note that, as this analysis is based on payments to early career teachers, it only applies to career changers during the early years of their teaching career and does not indicate that the pay elasticity for all teachers in older age brackets is higher than for younger teachers.

Figure 6 The retention decisions of career changers appear to be more responsive to ECRPs than younger early career teachers

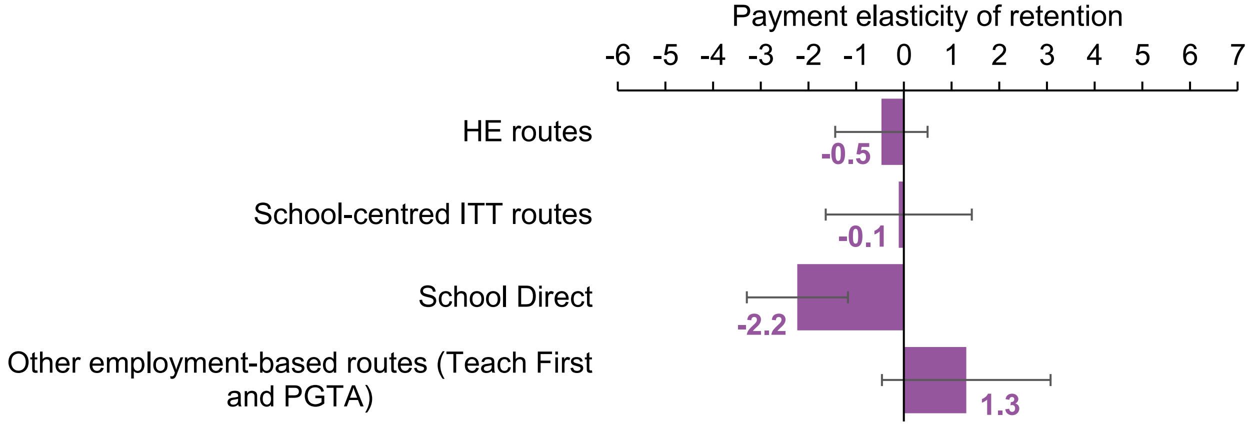


Note: The confidence intervals shown are 95 per cent confidence intervals.

Source: NFER analysis of ITT-PP and SWC data.

Figure 7 shows the variation in impact by ITT route. We explore variation by higher education, school-centred, School Direct and other employment-based routes (including Teach First and postgraduate teaching apprenticeship (PGTA)). Due to limitations in the data, we were not able to separate out School Direct salaried and fee-based routes for all cohorts in our analysis, so have analysed them together. The majority of School Direct trainees during this period were on fee-based rather than salaried courses. The analysis shows that teachers who had trained through School Direct showed higher responsiveness to payments than teachers who had trained through other routes.

Figure 7 The retention decisions of teachers who trained through School Direct appear to be more responsive to ECRPs than teachers who trained through other routes

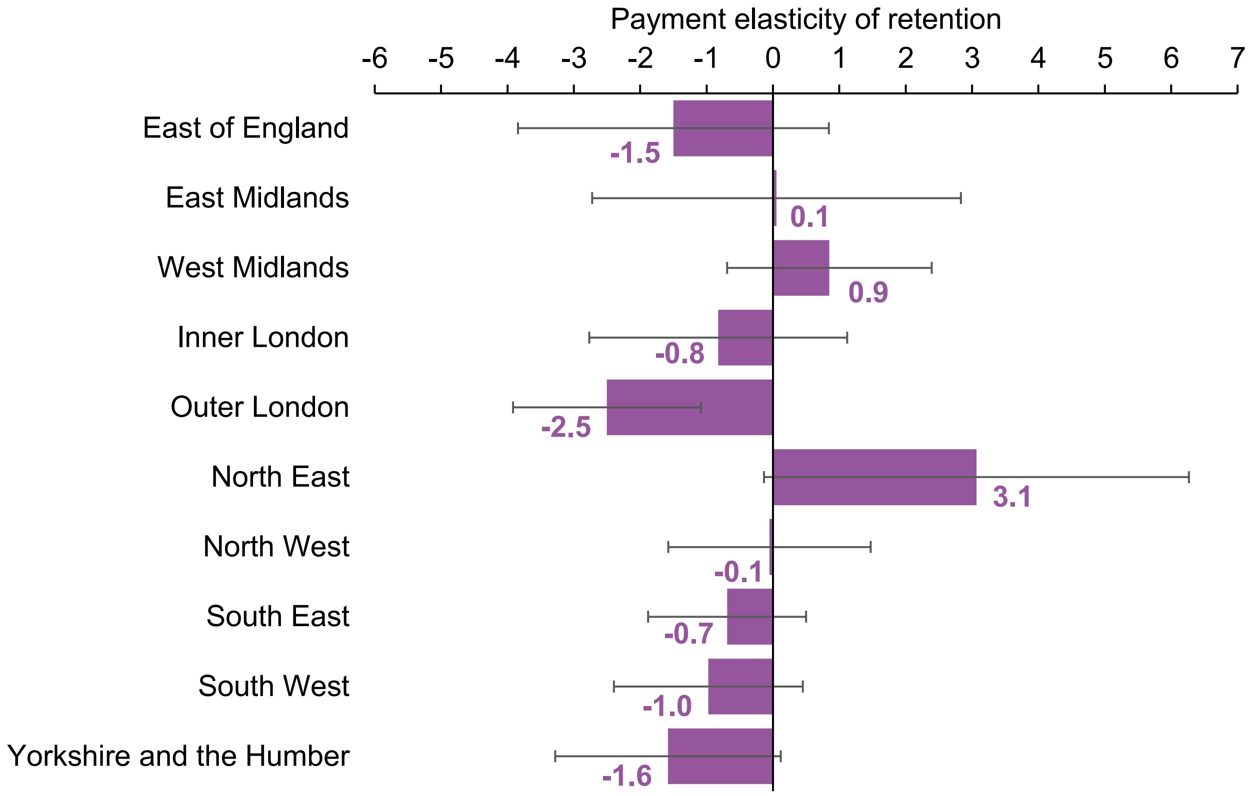


Note: The confidence intervals shown are 95 per cent confidence intervals.

Source: NFER analysis of ITT-PP and SWC data.

Figure 8 shows the variation in estimated elasticity by region. There is considerable variation between regions, although the estimate for each region has a relatively high degree of uncertainty due to sample size, meaning the confidence intervals are wide and differences should be interpreted with caution. The analysis suggests that teachers based in Outer London are more responsive to payments than teachers in other regions, while teachers in the North East have an estimated elasticity that is greater than one (although not statistically significant) implying a higher leaving rate than otherwise similar teachers.

Figure 8 The retention decisions of teachers based in some regions, such as Outer London, appear to be more responsive to ECRPs than teachers in other regions



Note: The confidence intervals shown are 95 per cent confidence intervals.

Source: NFER analysis of ITT-PP and SWC data.

The full set of elasticities estimated from the heterogeneity analysis are shown in Appendix A.

Overall, the findings suggest there is limited variation in responsiveness by characteristics, although some heterogeneity is evident on some characteristics. There is little significant heterogeneity in characteristics that have previously been used to target the retention payments, such as subject, levels of experience, quintiles of eligibility for free school meals and education investment areas.

6. Conclusions and recommendations

6.1. Conclusions

Previous research evidence from the UK and United States has indicated that early career retention payments are effective at improving retention. Our evaluation findings are partially supportive of this conclusion, as our estimates imply that retention payments are generally associated with slight improvements in retention. However, our findings are far from conclusive. Most of our estimates are not statistically significant at the five per cent level, indicating that the retention impact may be due to chance. This includes our overall estimate of the combined impact of all five schemes that have been piloted in the UK since 2018. Some of our impact estimates – most notably for the ECP scheme – indicate that the payments may in some cases even be associated with a higher leaving rate among eligible teachers than among otherwise similar teachers.

The studies that have evaluated ECRP schemes in England so far have mostly suggested that the retention decisions of early career teachers are highly responsive to retention payments, with estimated elasticities of -2 or greater. However, our evidence suggests that previous estimates of the payment elasticity of retention may be over-estimates. For example, while our estimated elasticity for MPB (-2.1) is similar to that estimated by CFE Research and FFT Education Datalab (2023), our estimated elasticity for MPRP (-1.2) is lower than that estimated by Sims and Benhenda (2022). Our estimated elasticity for LUP (-1.0) also has a lower magnitude than the previous literature on retention payments.

More generally, our analysis suggests that estimated impacts on retention are sensitive to the particular approach taken to defining eligibility and estimating the impact. This implies that there is uncertainty associated with any particular estimate in addition to the level of imprecision inherent in estimating the impact from a sample of eligible teachers (which is accounted for in the standard errors and confidence intervals). Any one study therefore needs to be interpreted cautiously and within the context of the wider literature. A review of the evidence by DfE estimated a consensus estimate of the overall pay elasticity of attrition among all teachers to be -1.5. Taking all the evidence on ECRPs into account, it may be reasonable to expect that such an elasticity holds for early career teachers and experienced teachers alike. However, more evidence is needed in this area, particularly on how responsive experienced teachers are to pay variation.

The implications of our findings for future policy are that retention payments may be a valuable additional tool for improving recruitment and retention when targeted at shortage subjects and may offer good value for money. However, the marginal cost per additional teacher-year appears to be considerably higher than for training bursaries, suggesting that retention payments should only be considered as part of a wider strategy for teacher recruitment and retention for subjects with bursaries that are at a clear maximum.

Our analysis suggests there is no conclusive evidence of any one scheme or any one model for retention payment schemes being more effective than any other. There is some indication, but with low confidence, that the MPRP and LUP model of making annual payments to teachers in their first five years may be consistently effective. There is less consistent evidence on the effectiveness for retention of advertising future retention payments to trainees, such as in MPB (which we find is significantly associated with lower retention rates) and ECP (which we find is associated with higher retention rates, although it is not statistically significant).

We also find no evidence of significant heterogeneity of impact by characteristics that is likely to be relevant for the design of future retention payment schemes to maximise their effectiveness. There is an indication that career changers, early career teachers that trained through School Direct and teachers in Outer London may be more responsive to payments than other teachers. However, these are not strong enough findings to confidently inform future policy design, nor are they factors that clearly link to wider policy goals that might also be emphasised in future policy design.

The impact of ECRPs on recruitment of teachers to ITT was not in scope for this evaluation. It is reasonable to hypothesise that anticipation of future payments could provide a boost to recruitment in addition to any impact on retention. As retention payments become more established and may in future be supported by longer-term policy guarantees and/or integration within the pay framework, this should be a focus for future research.

6.2. Recommendations

We recommend that:

- The Government should maintain a policy of teacher retention payments focussed on shortage subjects. Our evidence suggests that while their marginal cost is likely to be high, the impact of eliminating retention payments could be to worsen the teacher supply in these important subjects. Removing the payments could lead retention to worsen and sustaining them long-term as an offer for future early career teachers could prompt additional recruitment benefits.
- The Government should retain a policy of raising bursaries for subjects experiencing teacher supply challenges where bursaries are low and maintain high bursaries for maths, physics, chemistry and computing, raising them over time with the level of the teaching starting salary.
- The Government should continue to monitor and evaluate the impacts of new retention payment policies, such as the Targeted Retention Incentive. This report gives a comprehensive overview of the impact of the policies that it is possible to assess using SWC data up to 2023, but further evaluation opportunities will continue to become available. Evaluating the impacts of current and future policies will be a priority, alongside assessing the longer-term and post-eligibility impacts of previous policies.
- The Government should invest in deepening the evidence base of interventions that can improve teacher recruitment and retention. While the quality of research evidence around ECRPs and bursaries is high and growing, there is less high-quality and quantifiable evidence about the impacts of, for example, workload reduction, flexible working and professional development. Evidence on the impacts and costs of a wider range of policy measures would enable better comparative assessments of the relative costs and impacts, informing overall strategy development that is focussed on maximising cost effectiveness.

References

Adams, L., Coburn-Crane, S., Sanders-Earley, A., Keeble, R., Harris, H., Taylor, J. and Taylor, B. (2023) *Working lives of teachers and leaders - wave 1*. Available at: https://www.gov.uk/government/publications/working-lives-of-teachers-and-leaders-wave-1 (Accessed: 12 June 2025).

Annesi, I., Moreau, T. and Lellouch, J. (1989) 'Efficiency of the logistic regression and cox proportional hazards models in longitudinal studies', *Statistics in Medicine*, 8(12), pp. 1515–1521. Available at: https://doi.org/10.1002/sim.4780081211.

Bueno, C. and Sass, T.R. (2019) *The Effects of Differential Pay on Teacher Recruitment and Retention. Working Paper No. 219-0519*. Available at: https://files.eric.ed.gov/fulltext/ED595205.pdf (Accessed: 13 June 2025).

Callaway, B. and Sant'Anna, P.H.C. (2021) 'Difference-in-Differences with multiple time periods', *Journal of Econometrics*, 225(2), pp. 200–230. Available at: https://doi.org/10.1016/j.jeconom.2020.12.001.

CFE Research (2023) *Teacher student loan reimbursement scheme: final evaluation report*. Available at: https://assets.publishing.service.gov.uk/media/63ceb003d3bf7f3c4a199574/Teacher_student_loan_reimbursement_scheme_final_evaluation_report.pdf (Accessed: 13 June 2025).

CFE Research and FFT Education Datalab (2023) *Evaluation of the phased maths bursaries pilot: final report*. Available at: https://assets.publishing.service.gov.uk/media/656761b65936bb000d3166ea/Evaluation_of_the_phased_maths_bursaries_pilot_-_final_report_November-2023.pdf (Accessed: 13 June 2025).

Department for Education (2017) 'Teachers' student loan reimbursement: guidance for teachers and schools'. Available at: https://dera.ioe.ac.uk/id/eprint/30399/1/Teachers%27%20student%20loan%20reimbursement_%20guidance%20for%20teachers%20and%20schools%20-%20GOV.UK.pdf (Accessed: 13 June 2025).

Department for Education (2018) 'Early-career payments: guidance for teachers and schools'. Available at: https://dera.ioe.ac.uk/id/eprint/34343/1/guidance%20for%20teachers%20and%20schools%20-%20GOV.pdf (Accessed: 4 June 2025).

Department for Education (2019) 'Apply for mathematics and physics teacher retention payments'. Available at: https://dera.ioe.ac.uk/id/eprint/33441/1/Apply%20for%20mathematics%20and%20physics%20teacher%20retention%20payments%20-%20GOV.pdf (Accessed: 4 June 2025).

Department for Education (2022) 'Levelling up premium payments for teachers'. Available at: https://dera.ioe.ac.uk/id/eprint/39242/1/Levelling%20up%20premium%20payments%20for%20teachers%20-%20GOV.pdf (Accessed: 4 June 2025).

Department for Education (2020) *Government evidence to the STRB The 2020 pay award*. Available at: https://assets.publishing.service.gov.uk/media/5e25af7d40f0b62c46060d82/STRB_Written_Evidence_2020.pdf (Accessed: 3 March 2025).

Feng, L. and Sass, T.R. (2017) 'The impact of incentives to recruit and retain teachers in "Hard-to-Staff" subjects', *Journal of Policy Analysis and Management*, 37(1), pp. 112–135. Available at: https://doi.org/10.1002/pam.22037.

Harland, J., Bradley, E. and Worth, J. (2023) *Understanding the factors that support the recruitment and retention of teachers: review of flexible working approaches*. Available at: https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Review-of-flexible-working-approaches.pdf (Accessed: 13 June 2025).

Hutchings, M. (2011) *What impact does the wider economic situation have on teachers' career decisions? A literature review*. Research Report DFE-RR136. London: DfE. Available at: https://assets.publishing.service.gov.uk/media/5a7b14e140f0b66a2fc05055/DFE-RR136.pdf (Accessed: 4 June 2025).

Ingram, D.D. and Kleinman, J.C. (1989) 'Empirical comparisons of proportional hazards and logistic regression models', *Statistics in Medicine*, 8(5), pp. 525–538. Available at: https://doi.org/10.1002/sim.4780080502.

Labour Party (2024) 'Change: Labour Party Manifesto 2024'. Available at: Change Labour Party Manifesto 2024 (Accessed: 13 June 2025).

Martin, K. *et al.* (2023) *Supporting the recruitment and retention of teachers in schools with high proportions of disadvantaged pupils: understanding current practice around managing teacher workload*. Available at: https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Review-of-teacher-workload-management-approaches.pdf (Accessed: 5 December 2023).

Martin, K., Classick, R., Sharp, C. and Faulkner-Ellis, H. (2023) *Supporting the recruitment and retention of teachers in schools with high proportions of disadvantaged pupils: understanding current practice around managing teacher workload*. Available at: https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Review-of-teacher-workload-management-approaches.pdf (Accessed: 13 June 2025).

McLean, D., Tang, S. and Worth, J. (2023) *The impact of training bursaries on teacher recruitment and retention: an evaluation of impact and value for money*. Available at: https://www.nfer.ac.uk/media/bycg5uzk/the_impact_of_training_bursaries_on_teacher_recruitment_and_retention.pdf (Accessed: 13 June 2025).

McLean, D. and Worth, J. (2025) *Teacher labour market in England: annual report 2025*. Available at: https://www.nfer.ac.uk/media/afsn0rmb/teacher_labour_market_in_england_annual_report_2025.pdf (Accessed: 3 June 2024).

Migration Advisory Committee (MAC) report: teacher shortages in the UK (no date) *GOV.UK*. Available at: https://www.gov.uk/government/publications/migration-advisory-committee-mac-report-teacher-shortages-in-the-uk (Accessed: 13 June 2025).

See, B.H. *et al.* (2020) 'Teacher Recruitment and Retention: A Critical Review of International Evidence of Most Promising Interventions', *Education Sciences*, 10(10), p. 262. Available at: https://doi.org/10.3390/educsci10100262.

Sims, S. (2017) *What happens when you pay shortage subject teachers more money? Simulating the effect of early-career salary supplements on teacher supply in England*. Available at: https://www.gatsby.org.uk/uploads/education/datalab-simulating-the-effect-of-early-career-salary-supplements-on-teacher-supply-in-england.pdf (Accessed: 13 June 2025).

Sims, S. and Benhenda, A. (2022) *The effect of financial incentives on the retention of shortage-subject teachers: evidence from England*. CEPEO Working Paper No. 22-04. Available at: https://www.gatsby.org.uk/uploads/education/reports/pdf/the-effect-of-financial-incentives-on-the-retention-of-shortage-subject-teachers-evidence-from-england.pdf (Accessed: 13 June 2025).

Springer, M.G., Swain, W.A. and Rodriguez, L.A. (2016) 'Effective Teacher Retention Bonuses: Evidence From Tennessee', *Educational Evaluation and Policy Analysis*, 38(2), pp. 199–221. Available at: https://doi.org/10.3102/0162373715609687.

Tang, S. and Worth, J. (2024) *Policy analysis of student loan reimbursements for improving teacher retention*. Available at: https://www.nfer.ac.uk/media/sl0abjje/policy_analysis_of_student_loan_reimbursements_for_improving_teacher_retention.pdf (Accessed: 13 June 2025).

Theobald, R. *et al.* (2024) *The Impact of a $10,000 Bonus on Special Education Teacher Shortages in Hawai'i*. Available at: https://caldercenter.org/sites/default/files/2024-11/CALDER%20WP%20290-0823.pdf (Accessed: 3 March 2025).

Wooldridge, J.M. (2010) *Econometric Analysis of Cross Section and Panel Data*. The MIT Press. Available at: https://www.jstor.org/stable/j.ctt5hhcfr (Accessed: 3 June 2025).

Worth, J., Lynch, S., Hillary, J., Rennie, C. and Andrade, J. (2018) *Teacher workforce dynamics in England: nurturing, supporting and valuing teachers. Research overview*. Available at: https://www.nfer.ac.uk/media/kw2bvpum/teacher_workforce_dynamics_in_england_research_overview.pdf (Accessed: 13 June 2025).

Worth, J., Lazzari, G.D. and Hillary, J. (2017) *Teacher retention and turnover research: interim report*. Available at: https://www.nfer.ac.uk/media/3rgbtix0/teacher_retention_and_turnover_research_interim_report.pdf (Accessed: 13 June 2025).

Worth, J. and Van den Brande, J. (2019) *Retaining science, mathematics and computing teachers*, p. 30. Available at: https://www.nfer.ac.uk/media/d0tfx04x/retaining_science_mathematics_and_computing_teachers.pdf (Accessed: 13 June 2025).

NFER
National Foundation for Educational Research

7. Appendix A

7.1. Further methodological details

7.1.1. SWC data cleaning

Our working dataset combined information from several sources. We began by defining a ‘spine’ of secondary schools, based on the schools listed in the LUP eligibility list (Department for Education, 2022). We linked the schools to longitudinal information about them which was available from Get Information About Schools. We therefore removed primary, special and alternative provision schools and others that were clearly out of scope for the evaluation (e.g. schools in Wales, independent schools). We kept schools that were open continuously from 2015/16 to 2023/24, after accounting for changes in school ID and school mergers and splits.

We then merged this spine of schools to the SWC. We dropped the few schools that never merged to any SWC records. We also only kept teachers that had one or more SWC record from 2015/16 to 2023/24 in one of our in-scope schools. If a teacher was in an out-of-scope school for e.g. one year, we just dropped that one year’s record, not all the records for that teacher. We also only kept teachers who had ITT records, as ITT subject was one of our key control variables. This reduced the sample size considerably but suited the evaluation as it contained all eligible early-career teachers and a large group of comparison teachers.

We then matched the resulting teacher data to records in the curriculum and qualifications modules of the SWC that capture what subjects teachers teach and what qualifications they have. We dropped teachers that did not match to the curriculum/ qualifications modules because both subject taught and qualification were key definition variables for some of our ECRPs. However, we kept other records for that teacher if they matched in other years. We defined a teacher to be teaching a particular subject if they spent one or more hour teaching that subject in a school. We defined a teacher to be holding a degree in a particular subject if their qualifications records indicated they hold one or more degree and the subject code associated with that qualification maps to that subject.

7.1.2. Defining eligibility

We defined eligibility for each ECRP based on the published policy information, applying the rules shown in Table 10. Table 11 summarises potential eligibility for each payment scheme according to subject, ITT cohort and year. We merged all the cohort, year, subject, qualification and school area criteria from the SWC onto the eligibility look-up file to determine whether a teacher met each of the criteria for eligibility or not. While the original policy documentation for the MPB and ECP indicated that teachers must have been continuously employed since their first year in teaching to be eligible, this guidance was subsequently updated to relax that, so we therefore relaxed this criteria for our analysis.

We generated the post-eligibility flags by observing whether a teacher was not eligible for a particular ECRP in one year but had been eligible in a previous year. For example, if a teacher had been eligible for MPRP in 2020/21 but was no longer in 2021/22 we recorded them as being post-eligible.

Most schemes were all announced with very little notice given to teachers, so we assume that any anticipation effects on retention for these programmes were negligible. For MPB and ECP, teachers were eligible for the programme from their first year of teaching but only actually received the payment in later years and this was advertised in advance. Therefore, as long as they were in an eligible subject/ school we recorded teachers as being 'eligible' for MPB/ ECP in those early pre-payment years. We included a zero payment amount for these years when estimating the average payment amount.

Given the eligibility criteria summarised in Table 10, we applied the following methodologies to evaluate each payment scheme:

- MPRP: triple differences, comparing: (1) eligible and ineligible subjects, (2) eligible and ineligible areas and (3) years in which the policy was active and inactive.
- TSLR: triple differences, comparing: (1) eligible and ineligible subjects, (2) eligible and ineligible areas and (3) years in which the policy was active and inactive.
- MPB: difference-in-differences, comparing: (1) eligible and ineligible subjects, (2) years in which the policy was active and inactive.
- ECP: difference-in-differences, comparing: (1) eligible and ineligible subjects, (2) years in which the policy was active and inactive.
- LUP: triple differences, comparing: (1) eligible and ineligible subjects, (2) eligible and ineligible schools and (3) years in which the policy was active and inactive.

Table 10 Definitions of eligibility for each ECRP scheme

ECRP	Subject eligibility	ITT cohort eligibility	Teaching years eligible	Other criteria
MPRP	- Degree or ITT qualification in maths/ physics - Teaches some hours in maths/ physics	2014/15 – 2019/20	2019/20 – 2020/21	Only eligible if teaching in certain local authorities
TSLR	- Teaches some hours in languages, physics, chemistry, biology, computer science	2013/14 – 2018/19	2018/19 – 2023/24	Only eligible if teaching in certain local authorities
MPB	- ITT qualification in maths - Teaches at least 50% of hours in maths	2018/19 – 2019/20	2021/22 – 2023/24	Uplift available for teachers in certain local authorities
ECP	- ITT qualification in maths/ physics/ chemistry/ languages and teaches at least 50% of hours in these subjects (including general science)	2018/19 – 2020/21 (depending on subject)	2021/22 – 2023/24 (depending on subject)	Uplift available for teachers in certain local authorities
LUP	- Degree or ITT qualification in physics/ maths/ computing/ chemistry	2017/18 – 2023/24	2022/23 – 2023/24	Only eligible if teaching in specific schools; payment amount varies by school group

Table 11 Summary of potential ECRP eligibility by subject, ITT cohort and year (other factors such as school type and location also determine eligibility)

Subject	Year ITT cohort	2018/19	2019/20	2020/21	2021/22	2022/23
Maths	2013/14					
	2014/15		MPRP			
	2015/16		MPRP	MPRP		
	2016/17		MPRP	MPRP		
	2017/18		MPRP	MPRP		LUP
	2018/19		MPRP + MPB(elig)	MPRP + MPB(elig)	MPB	LUP
	2019/20			MPRP + MPB(elig)	MPB (elig)	LUP/ MPB
	2020/21				ECP(elig)	LUP/ ECP
	2021/22					LUP
	2022/23					
	2023/24					
	2024/25					
Physics	2013/14	TSLR	TSLR	TSLR	TSLR	TSLR
	2014/15	TSLR	MPRP + TSLR	TSLR	TSLR	TSLR
	2015/16	TSLR	MPRP + TSLR	MPRP + TSLR	TSLR	TSLR
	2016/17	TSLR	MPRP + TSLR	MPRP + TSLR	TSLR	TSLR
	2017/18	TSLR	MPRP + TSLR	MPRP + TSLR	TSLR	LUP + TSLR
	2018/19	TSLR	MPRP + TSLR	MPRP + TSLR	TSLR	LUP + TSLR
	2019/20			MPRP		LUP
	2020/21				ECP(elig)	LUP/ ECP
	2021/22					LUP
	2022/23					
	2023/24					
	2024/25					
Chemistry	2013/14	TSLR	TSLR	TSLR	TSLR	TSLR
	2014/15	TSLR	TSLR	TSLR	TSLR	TSLR
	2015/16	TSLR	TSLR	TSLR	TSLR	TSLR
	2016/17	TSLR	TSLR	TSLR	TSLR	TSLR
	2017/18	TSLR	TSLR	TSLR	TSLR	LUP + TSLR
	2018/19	TSLR	TSLR	TSLR	TSLR	LUP + TSLR
	2019/20					LUP
	2020/21				ECP(elig)	LUP/ ECP
	2021/22					LUP
	2022/23					
	2023/24					
	2024/25					

Subject	Year ITT cohort	2018/19	2019/20	2020/21	2021/22	2022/23
Languages	2013/14	TSLR	TSLR	TSLR	TSLR	TSLR
	2014/15	TSLR	TSLR	TSLR	TSLR	TSLR
	2015/16	TSLR	TSLR	TSLR	TSLR	TSLR
	2016/17	TSLR	TSLR	TSLR	TSLR	TSLR
	2017/18	TSLR	TSLR	TSLR	TSLR	TSLR
	2018/19	TSLR	TSLR	TSLR	TSLR	TSLR
	2019/20					
	2020/21				ECP(elig)	ECP
	2021/22					
	2022/23					
	2023/24					
	2024/25					
Computer science	2013/14	TSLR	TSLR	TSLR	TSLR	TSLR
	2014/15	TSLR	TSLR	TSLR	TSLR	TSLR
	2015/16	TSLR	TSLR	TSLR	TSLR	TSLR
	2016/17	TSLR	TSLR	TSLR	TSLR	TSLR
	2017/18	TSLR	TSLR	TSLR	TSLR	LUP + TSLR
	2018/19	TSLR	TSLR	TSLR	TSLR	LUP + TSLR
	2019/20					LUP
	2020/21				ECP(elig)	LUP/ ECP
	2021/22					LUP
	2022/23					
	2023/24					
	2024/25					
Biology	2013/14	TSLR	TSLR	TSLR	TSLR	TSLR
	2014/15	TSLR	TSLR	TSLR	TSLR	TSLR
	2015/16	TSLR	TSLR	TSLR	TSLR	TSLR
	2016/17	TSLR	TSLR	TSLR	TSLR	TSLR
	2017/18	TSLR	TSLR	TSLR	TSLR	TSLR
	2018/19	TSLR	TSLR	TSLR	TSLR	TSLR
	2019/20					
	2020/21					
	2021/22					
	2022/23					
	2023/24					
	2024/25					

Note: "(elig)" indicates where a teacher could be eligible for a future payment but does not receive one in that year. '+' indicates that eligible teachers are assumed to receive both payments, while '/' indicates that eligible teachers would receive the higher of the payments.

7.2. Sensitivity checks

We estimated a range of specifications of our baseline regression model to check how sensitive the model was to the definition of the sample. We also examined whether different definitions of our key eligibility criteria could have been driving our results. We did this to try and understand some of the differences between our estimates and similar estimates from the literature.

7.2.1. Checks on sensitivity to subject and year

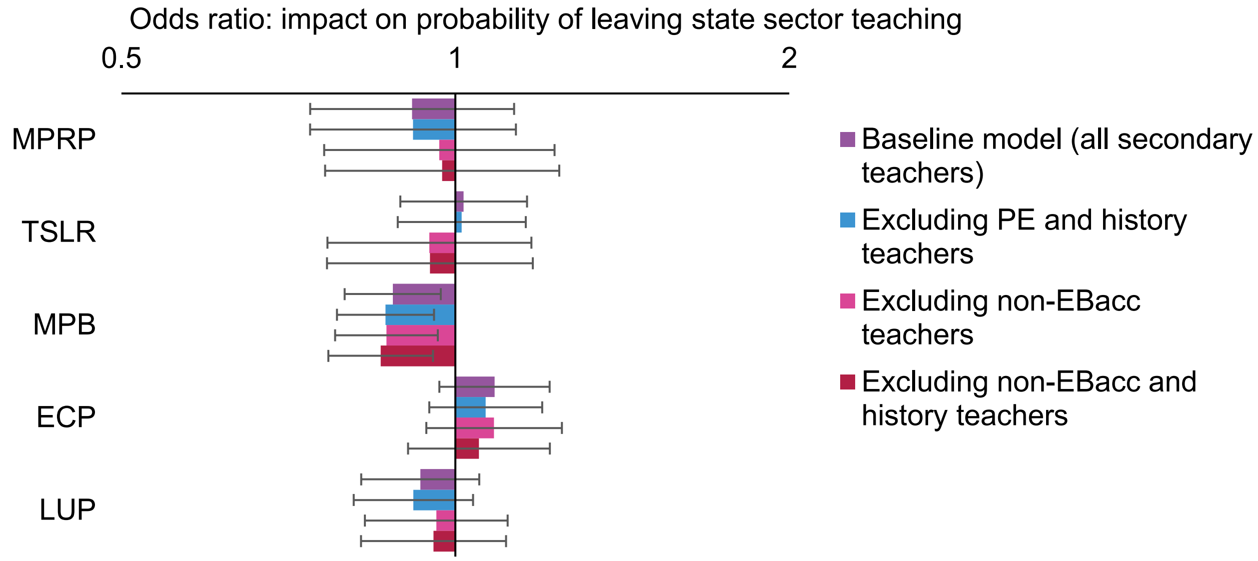
Our first sensitivity check was to explore the effect of changing the teachers included in our comparison group. Our baseline model includes all secondary teachers who were not eligible, or post-eligible, for an ECRP. We use a range of controls to account for underlying differences in retention rates by subject, area, school context, year and cohort effects, plus interactions between these variables. However, it may be that some groups of teachers are unsuitable for the comparison group because their retention behaviour differs systematically from eligible teachers in ways that are not captured by our controls. For example, ECRPs tend to be targeted at shortage subjects that tend to have low retention rates, such as maths, physics and chemistry. It may therefore be inappropriate to include non-shortage subjects with high retention rates, such as physical education and history. However, the downside of excluding teachers is in reduced sample size and thereby reduced estimation precision.

We explore this by defining narrower comparison groups. First we exclude PE and history teachers from the comparison group, on the basis that they are not shortage subjects since they have routinely met or exceeded recruitment targets over many years. Second, we exclude all teachers of subjects that are not EBacc[3] subjects from our analysis. Finally, we exclude all teachers of subjects that are not EBacc[4] subjects and also history from our analysis.

As shown in Figure 9, the results are very similar in shape to the baseline model. All the estimates are in the same direction as in the baseline model, except for TSLR which was very close to one in the baseline model. All the estimates are not significantly different from one, as in the baseline model, except for MPB where they are all significantly different from one, which was also the case in the baseline model. This indicates that the findings are not sensitive to the exclusion of non-shortage subject teachers from the comparison group.

[3] EBacc subjects are English, maths, sciences, modern foreign languages, history and geography.
[4] EBacc subjects are English, maths, sciences, modern foreign languages, history and geography.

Figure 9 The estimates are not sensitive to the exclusion of teachers of non-shortage subjects from the comparison group



Note: The confidence intervals shown are 95 per cent confidence intervals.

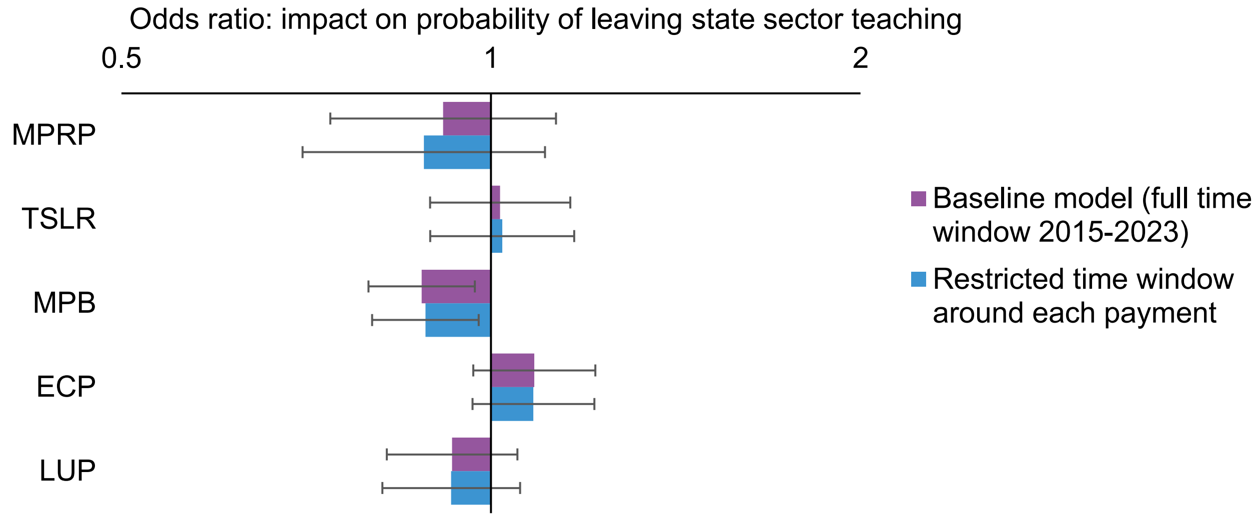
Source: NFER analysis of ITT-PP and SWC data.

Our second sensitivity check was to explore what effect estimating a large regression model that encompassed the entire time period and estimated impacts for all the retention payments may have had. We do this by re-estimating the models and each time restricting the time windows to be the relevant ones around the introduction of the relevant payment. For example, MPRP was introduced in 2019/20 and continued in 2020/21. However, inclusion of data from 2021/22 onwards is largely irrelevant to estimating the impact. We therefore estimate the MPRP impact from a subset of years covering 2017/18-2020/21 (i.e. still including some years prior to its introduction as part of our difference-in-differences approach).

We did this separately for each payment, estimating on subsets of 2017/18-2020/21 data for MPRP, 2016/17-2022/23 data for TSLR, 2018/19-2022/23 data for MPB, 2019/20-2022/23 data for ECP and 2020/21-2022/23 data for LUP. When estimating on these subsets we included all relevant controls and variables for (at least partially) estimating the impacts of other payments in each model, but only report the main impact estimates of interest.

As shows in Figure 10, the findings are very similar to those of the baseline model, suggesting the impact estimates were not sensitive to being estimated using a single large model as opposed to more tailored time periods.

Figure 10 The estimates are not sensitive to restricting the time period to be relevant to the timing of the specific payment scheme



Note: The confidence intervals shown are 95 per cent confidence intervals.

Source: NFER analysis of ITT-PP and SWC data.

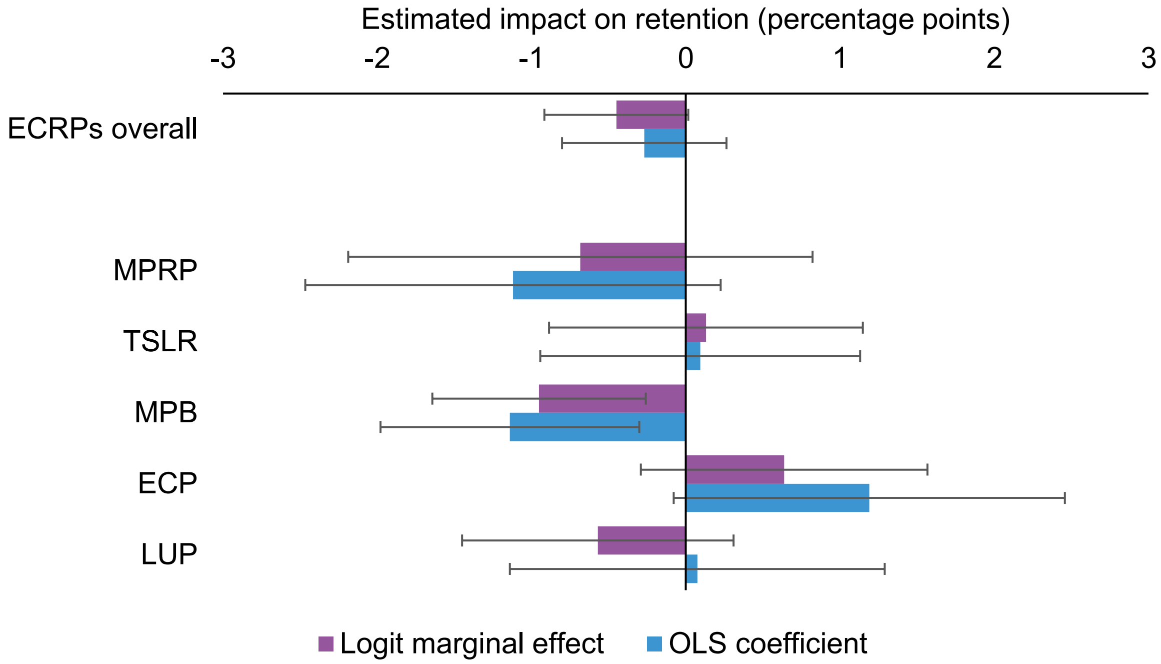
7.2.2. Use of logistic regression model versus ordinary least squares (OLS)

We implemented our difference-in-differences approach by estimating a logistic regression model. Retention is a binary dependent variable (taking values of 0 or 1), so the logistic link function reflects this. Implementing a difference-in-differences model using a binary dependent variable has been questioned by some academics, with a preference for a linear probability model expressed by Wooldridge (2010).

To test the sensitivity of our model to a linear specification, we estimated the same model using ordinary least squares (OLS) regression. Rather than an odds ratio, OLS yields a percentage point differences that is invariant to the underlying retention rate. We compared this with the marginal effect estimated from the logistic baseline model, which is also in percentage point difference terms (as estimated at mean values).

As shown in Figure 11, using OLS gives very similar results. There is no consistent pattern, with estimates being slightly higher and some slightly lower. The main difference is that the LUP impact disappears, but it is still insignificant, as it is in the baseline model. The estimated impact for MPB remains the only significant effect.

Figure 11 The estimates are not sensitive to being estimated using a linear probability model instead of a logistic regression model



Note: The confidence intervals shown are 95 per cent confidence intervals.

Source: NFER analysis of ITT-PP and SWC data.

7.3. Comparisons with existing research

A key difference with the previous literature found in our estimates was on the impact of MPRP. Our model estimated an odds ratio of 0.90, a reduction in the leaving rate of 8.1 per cent and an elasticity of -1.2. However, Sims and Benhenda (2022) estimated a hazard rate of 0.77 (which we estimate to be equivalent in this context to an odds ratio of 0.75), a 23 per cent reduction in the leaving rate an elasticity of -3.

We undertook a replication analysis to identify what effect small differences in our respective approaches might have had. Both analysis were undertaken in their own contexts so can be expected to differ from one another. For example, Sims and Benhenda focussed only on MPRP, whereas our model focussed on estimating impacts from more payment schemes as well. Also, Sims and Benhenda only had access to data up to 2020/21, whereas we have been able to use data up to 2023/24.

We first followed the method laid out in Sims and Benhenda's paper, using the triple difference set up, subject eligibility criterion, cohorts and years of data they used. Our replication analysis found a very similar estimated impact. Any minor extra differences are likely to be driven by other factors such as data cleaning decisions. This included the difference in estimation method, between our use of logistic regression and Sims and Benhenda's use of Cox proportional hazards. As we can replicate the findings using a logistic regression model and the literature has found that the differences are minimal, we did not explore this further.

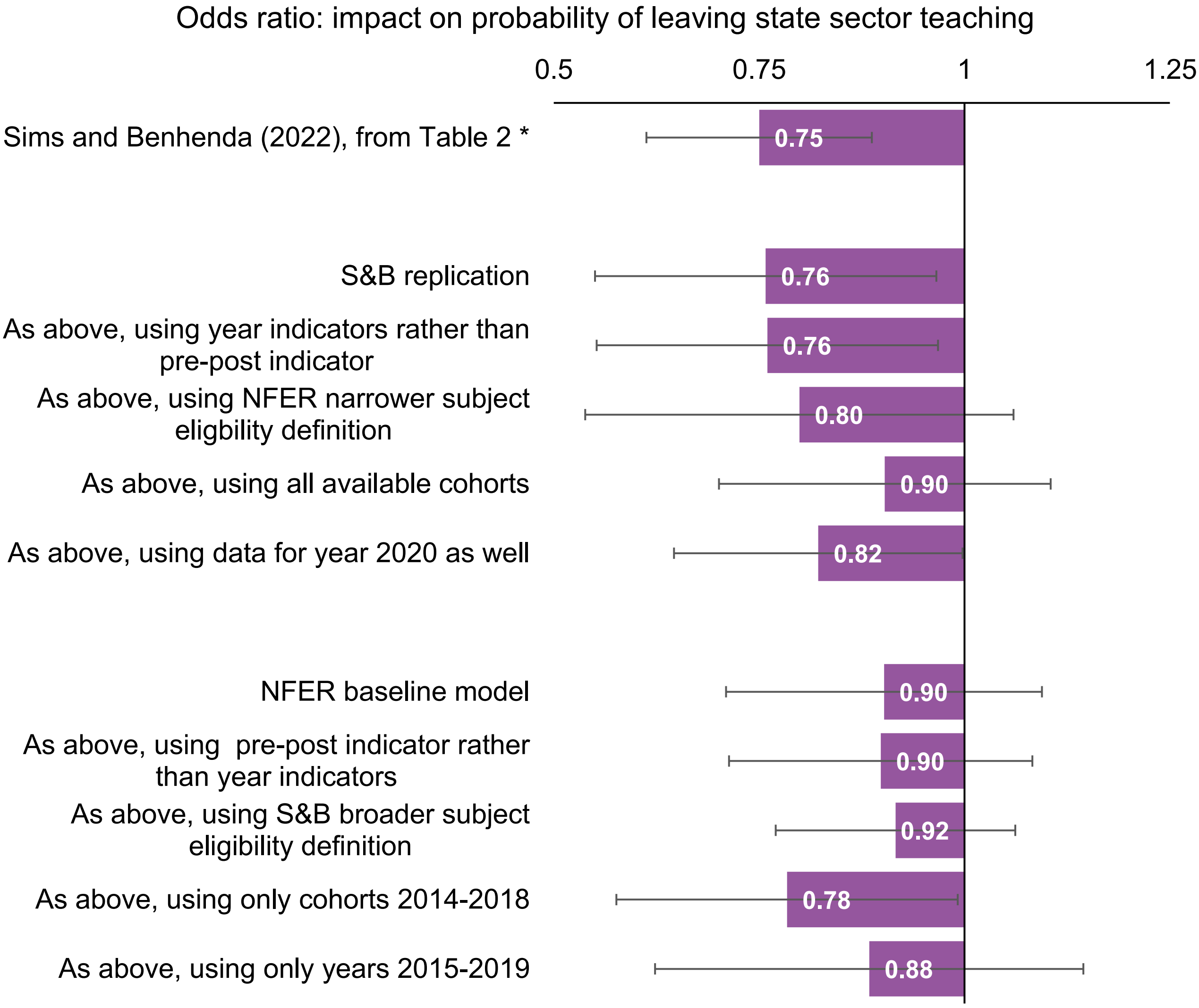
We identified four key ways in which our analyses differed and tested their influence over the estimated impacts by relaxing each approach. First, Sims and Benhenda use a simple before-or-after time period control, whereas we use control indicators for each year to account for changes in retention over time. Second, Sims and Benhenda use a broad subject eligibility criterion based on subject taught, whereas we use a narrower definition based on subject and qualification. We believe our definition is closer to the policy definition, although we understand that the guidance may have been more loosely applied in practice. Neither research teams had access to matched information on which teachers had received the payments to test this. Third, Sims and Benhenda used only teachers from cohorts that qualified from teaching in 2014/15 to 2018/19, whereas we use teachers from all available cohorts (see Table X for an illustration of this). Finally, as noted above, Sims and Benhenda's analysis focussed on years up to 2020/21 and this was all that was available at the time, whereas our baseline model used data up to 2023/24.

We first changed the model according to each of the different approaches from the starting point of the replication model and then did the same from the starting point of our baseline model.

The results are shown in Figure 12. The analysis suggested that some factors appeared to play little role in the differences in estimated impact. For example, whether a before-and-after control or set of year indicators was used made no difference to either model. The subject eligibility criterion used made some difference, but not in a consistent way: using our narrower definition in the replication model shrank the estimate slightly, while using the broader definition in our baseline model also shrank the estimate slightly. Changing the years used in the analysis appeared to make some small difference: adding an extra year in the replication model shrank the estimated odds ratio by 0.06, while using the restricted set of years in our baseline model increased the estimated odds ratio by 0.02. However, this is a small difference in the context of an overall difference in the odds ratios of 0.15.

The most substantial difference was due to the choice of which cohorts to include in the analysis. Using a broader set of cohorts in the replication model shrank the estimated odds ratio considerably, by 0.14, while using the narrower set of cohorts in our baseline model increased the estimate considerably, by 0.12. This suggests that the choice of cohorts to include was a key driver of the differences. All the additional cohorts we include were non-treated, so it only served to expand the size and scope of the comparison group. We chose to do this so that our comparison group included pre-intervention teachers with five or more years of experience. However, this is an arbitrary choice in the context of the evaluation and both approaches are defensible. This suggests that any estimate of the impact on retention is likely to be subject to uncertainty that depends on how the analysis is undertaken, in addition to a high degree of estimation imprecision relative to the size of effect due to the limited available sample size.

Figure 12 Differences in the cohorts included in the Sims and Benhenda (2022) and NFER analyses appear to be the main driver of differences in the impact estimates



Note: *Sims and Benhenda estimated and reported a hazard rate of 0.77, rather than an odds ratio. The equivalent odds ratio at a baseline leaving rate of 7.5 per cent is 0.75, as presented above for comparison. The confidence intervals shown are 95 per cent confidence intervals.

Source: NFER analysis of ITT-PP and SWC data.

Table 12 Differences in the cohorts included in the Sims and Benhenda (2022) and NFER analyses

Year / Cohort	2015/16	2016/17	2017/18	2018/19	2019/20	2020/21
2010/11						
2011/12						
2012/13						
2013/14						
2014/15					MPRP	
2015/16					MPRP	MPRP
2016/17					MPRP	MPRP
2017/18					MPRP	MPRP
2018/19					MPRP	MPRP
2019/20						MPRP

Note: Blue cells indicate cohort-years of teachers included in both our baseline model and Sims and Benhenda's analysis. Purple cells indicate cohort-years of teachers included in our baseline model but not in Sims and Benhenda's analysis. White cells are not present in the data as these cohorts of teachers are still in training, or are yet to train, as teachers. 'MPRP' indicates where maths and physics teaches in eligible areas were eligible for a payment.

7.4 Heterogeneity analysis

Table 13 shows the full set of heterogeneity findings for all the characteristics that we tested, reporting the estimated elasticity and associated standard error. Only characteristics where statistically significant variation was detected are reported in the main report section.

Table 13 Estimated elasticities from the full heterogeneity analysis

Characteristic group	Characteristic	Elasticity	SE
	Overall	-0.7	0.4
Sex	Female	-0.8	0.5
	Male	-0.7	0.5
Experience	0 years' experience	-0.2	0.6
	1 year experience	-0.6	0.5
	2 years' experience	-0.7	0.4
	3 years' experience	-1.1	0.9
	4 years' experience	-1.9	1.4
	5 or more years' experience	-0.8	2.1
Age	24 or younger	-0.3	0.5
	25 to 29	0.1	0.5
	30 to 39	-1.0	0.8
	40 to 49	-2.3	0.8
	50 to 59	-3.2	0.9
Ethnicity	White	-1.1	0.4
	Asian	0.6	0.7
	Black	-1.2	1.1
	Mixed	1.6	1.6
	Other	0.2	1.8
Working pattern	Full time	-0.6	0.4
	Part time	-1.6	0.6
Subject	Physics	-2.0	1.4
	Mathematics	-0.6	0.4
	Computing	1.4	2.1
	Chemistry	1.2	1.6
	Languages	-2.7	1.4
	Biology	-1.0	3.1
Degree class	First	-0.1	0.6
	Upper second	-0.6	0.5
	Lower second	-1.1	0.6
	Other	-1.7	1.0
	Unknown	-1.4	1.1

Characteristic group	Characteristic	Elasticity	SE
ITT route	HE routes	-0.5	0.5
	School-centred ITT routes	-0.1	0.8
	School Direct	-2.2	0.5
	Other employment-based routes (Teach First and PGTA)	1.3	0.9
FSM quintile	First quintile	-0.9	0.6
	Second quintile	-1.4	0.8
	Third quintile	-0.3	0.7
	Fourth quintile	0.1	0.7
	Fifth quintile	-1.2	0.6
EIA status	Not EIA	-0.9	0.4
	EIA	-0.3	0.6
Ofsted rating	Outstanding	-0.7	0.7
	Good	-0.5	0.5
	RI/ satisfactory	-1.7	0.8
	Inadequate	0.6	1.6
	Unknown	-1.1	0.8
Region	East of England	-1.5	1.2
	East Midlands	0.1	1.4
	West Midlands	0.9	0.8
	Inner London	-0.8	1.0
	Outer London	-2.5	0.7
	North East	3.1	1.6
	North West	-0.1	0.8
	South East	-0.7	0.6
	South West	-1.0	0.7
	Yorkshire and the Humber	-1.6	0.9

Source: NFER analysis of ITT-PP and SWC data.

Evidence for excellence in education

Public

The Mere, Upton Park, Slough, Berks SL1 2DQ
T: +44 (0)1753 574123 • F: +44 (0)1753 691632 • enquiries@nfer.ac.uk

www.nfer.ac.uk

NFER ref. NUTR

ISBN. **978-1-916567-32-0**