UCL

CENTRE FOR EDUCATION POLICY
& EQUALISING OPPORTUNITIES

Working Paper No. 22-04

The effect of financial incentives on the retention of shortage-subject teachers: evidence from England

Sam Sims
University College London

Asma Benhenda
University College London

School systems often experience shortages of maths and science teachers, reflecting difficulties in both recruiting and retaining people qualified to teach these subjects. In England, teachers with maths and science degrees face a higher outside pay ratio than other teachers and also tend to leave the profession at higher rates. We evaluate a policy aimed at improving retention by providing targeted uplifts in pay worth 8% of gross salary for early-career maths and physics teachers. Leveraging variation in eligibility across time, regions and school subjects, we find that eligible teachers are 23% less likely to leave teaching in state funded schools in years they were eligible for payments. This implies a pay-elasticity-of-exit of -3, which is similar to results from evaluations of similar policies in the United States. Our analysis suggests that the cost per additional teacher retained through the policy is 32% lower than training an equivalent replacement teacher. Taken together, these results suggest that persistent shortages of maths and science teachers can be reduced through targeted pay supplement policies.

VERSION: April 2022

Suggested citation: Sims, S., & Benhenda, A. (2022). *The effect of financial incentives on the retention of shortage-subject teachers: evidence from England* (CEPEO Working Paper No. 22-04). Centre for Education Policy and Equalising Opportunities, UCL.

Disclaimer

Highlights

- Teacher shortages are a widespread problem in publicly funded school systems, especially among maths and science teachers.
- One reason for low retention rates of maths and science teachers is the long-run decline in the competitiveness of their pay relative to other graduate occupations.
- We use comprehensive administrative data on teachers in England to evaluate the effect of offering maths and physics teachers in the first five years of their careers an 8% increase in pay on teacher retention.
- We find that eligibility for these retention payments decrease the probability of attrition in a given year by 23%, which implies a pay-elasticity-of-exit of -3.
- Contrary to existing research, we find no evidence that the effect of the policy varies based on whether a teacher has between one and five years of experience.
- The cost per additional teacher retained by the policy is 32% lower than the cost of training an equivalent replacement teacher.

Why does this matter?

Shortages of maths and science teachers are not inevitable. They can be mitigated through targeted pay supplement policies.

The effect of financial incentives on the retention of shortage-subject teachers: evidence from England

Sam Sims (UCL)
Asma Benhenda (UCL)

School systems often experience shortages of maths and science teachers, reflecting difficulties in both recruiting and retaining people qualified to teach these subjects. In England, teachers with maths and science degrees face a higher outside pay ratio than other teachers and also tend to leave the profession at higher rates. We evaluate a policy aimed at improving retention by providing targeted uplifts in pay worth 8% of gross salary for early-career maths and physics teachers. Leveraging variation in eligibility across time, regions and school subjects, we find that eligible teachers are 23% less likely to leave teaching in state funded schools in years they were eligible for payments. This implies a pay-elasticity-of-exit of -3, which is similar to results from evaluations of similar policies in the USA. Our cost effectiveness analysis suggests that the cost per additional teacher retained is 32% lower than the costs of training an equivalent replacement teacher. Taken together, these results suggest that persistent shortages of maths and science teachers can be reduced through targeted pay supplement policies.

Acknowledgments: This work was generously funded by the Gatsby Foundation. This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

INTRODUCTION

Teacher shortages are a widespread problem in publicly funded school systems (Dolton, 2020). In a recent 48-country survey, 21% of school principals cited shortages of qualified teachers as hindering the quality of instruction in their school (Jerrim & Sims, 2019). The problem is particularly acute among maths and science teachers, resulting in more pupils being taught by teachers lacking specialist knowledge in these subjects (Dee & Goldhaber, 2017; MAC, 2017). These shortages in part reflect high levels of attrition (leaving the profession) among early-career teachers (Sims, 2018). Such attrition also incurs additional costs, since departing teachers take with them valuable knowledge about their pupils and the school context, resulting in reduced pupil achievement (Gibbons et al., 2021; Ronfeldt et al., 2013).

One reason for low retention rates amongst among early-career teachers is the long-run decline in the competitiveness of teachers' pay, relative to other graduate occupations (Chevalier et al., 2007; Dolton & Chung, 2004). This is particularly true for those with maths and science degrees, who face higher average pay outside of teaching than teachers with degrees in other subjects (Britton et al., 2016; LiVecchi, 2017; MAC, 2017). Teachers' probability of leaving the profession is related to the size of this outside pay ratio, particularly among early-career teachers (Gilpin, 2011; Hendricks, 2014; Marachand & Weber, 2020).

This begs the questions of whether such shortages could be reduced through policies that provide targeted increases in the pay of early-career science and maths teachers. In this paper, we provide empirical evidence on this by evaluating one such reform. More precisely, we use data on all teachers in England to evaluate the effects of offering maths and physics teachers in the first five years of their careers a £2,000 per annum increase in pay on whether or not they remain teaching in any state-funded school nationwide. To do so, we exploit variations in eligibility across regions, school subjects, and academic years in a triple-difference model. We also exploit plausibly exogenous variation in the experience levels of teachers at the time of eligibility to investigate heterogeneous effects of the policy.

We find that eligibility for the retention payments (RP) decrease the probability of attrition in a given year by 23%. This result is robust to tests for various year and school subject placebo effects. The RP payments represent 8% or more of eligible teachers' gross annual pay, which implies a pay-elasticity-of-exit of -3. A simple simulation suggests that, taking into account deadweight loss, the policy costs around £63,000 for each additional teacher retained two years after initial qualification. Generating the same number of teachers through additional recruitment would cost approximately £92,000 per equivalent teacher. Taken together, our results suggest that, where policymakers are unable to recruit enough trainee teachers, targeted retention payments may be a cost-effective method of reducing shortages.

This paper contributes to the growing literature on the use of financial incentives to retain teachers. Feng & Sass (2018) study a similar policy in Florida in which early-career science, maths and foreign language teachers were awarded targeted pay increases and estimated a pay-elasticity-of-exit between -2.5 and -3.3, depending on subject taught. Similarly, Bueno & Sass (2018) study a policy in Georgia in which early-career maths and science teachers were awarded targeted pay increases and estimated a pay-elasticity-of-exit of -3.4. Our paper shows that similar results pertain in England. Building on this, a unique contribution of our paper is to exploit plausibly exogenous variation in teacher experience to understand heterogeneous effects, which potentially has direct implications for the targeting of such policies in future. Contrary to Hendricks (2014) and Gilpin (2011) we are unable to detect any such heterogeneity, suggesting teachers within their first five years in the profession are similarly sensitive to additional pay.

SETTING AND POLICY

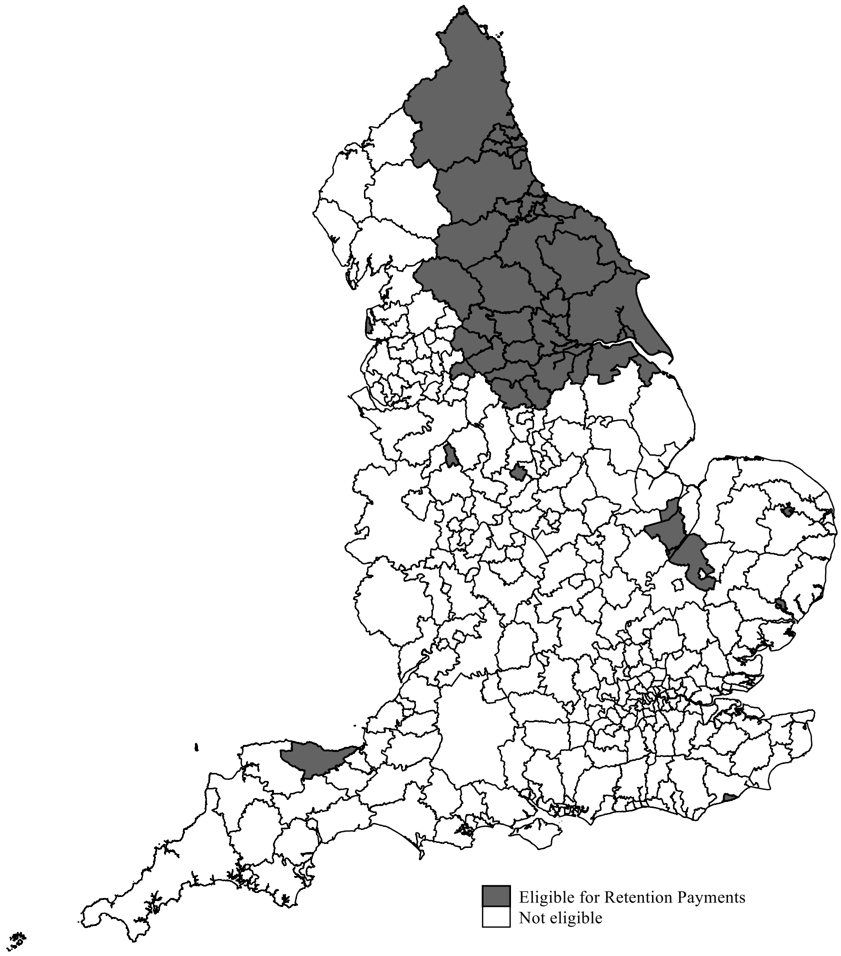
Education policy is England is determined by central government in London. Compulsory education is usually divided into two phases: primary school (ages 4-11) and secondary school (ages 11-16+). Teaching in England is a largely graduate profession, with the vast majority of teachers completing a three-year fee-paying undergraduate degree, followed by a one-year post-graduate teaching qualification. The policy we evaluate in this paper is targeted at secondary teachers, of which there are currently 205,000 working in state-funded schools in England. Secondary teachers train to teach in a specific subject but are not required to have a degree in the subject in which they train to teach. On completion of their training year, they are awarded Qualified Teacher Status. Schools may employ teachers to teach in any subject, regardless of the subject in which they trained to teach.

Once qualified, teachers must be paid above a minimum value, which is determined by central government. In 2018/19 this minimum was set at £23,720 for teachers outside London, which was around £2,000 lower than the average graduate starting salary at the time (STRB, 2019). Subsequent pay progression for teachers is at the discretion of their headteacher (principal). However, headteachers have to cover all costs from their fixed school budget, which is determined by a formula set by central government. Teaching unions also publish advisory pay progression frameworks, setting out their recommended annual pay increases. Some schools stick to these annual pay increments, which makes teacher pay a function of experience. However, the majority of schools deviate from the union advisory pay progression frameworks in some way (Anders et al., 2019).

The Retention Payment Reform

The retention payment (RP) policy involves central government paying £2,000 after tax per annum directly to eligible qualified teachers. In 2018/19, this £2,000 payment was equivalent to 8% of minimum gross teacher pay. Teachers were eligible for RP payments if they met all of the following criteria: they were teaching some maths or physics in a state funded schools in England; had any degree or a teaching qualification in maths or physics; received Qualified Teacher Status between the 2014/15 and 2019/20 academic years; were employed in a school in one of 42 target local authorities (out of a total of 343 local authorities in England). The 42 target local authorities are highlighted in Figure 1 below and were selected by policymakers because they had high levels of educational disadvantage.[i] The eligible areas therefore do not necessarily have higher rates of attrition than non-eligible areas. Eligible areas are comprised of two disadvantaged regions (the North East and Yorkshire & Humber), plus nine other areas chosen from a list of the 39 most educationally deprived local authorities in England. The policy was ‘on’ in the 2019/20 and 2020/21 academic years and ‘off’ in all other academic years. To receive the RP payments, teachers had to apply online in each year in which they were eligible. The Department for Education estimates that take-up was above 90%.

Figure 1. Local authorities eligible for Retention Payment



Contemporaneous Reforms

Two other pay reforms were introduced at around the same time as the RP policy in response to the teacher shortages in England. The first is the Teacher Student Loan Reimbursement (TSLR) policy. Most undergraduate students in England pay for their undergraduate degree course using a government student loan, which they repay once earnings exceed a certain threshold. The TSLR policy involves reimbursing teachers for the full value of their student loan repayments. Teachers were eligible for TSLR payments if they met all of the following criteria: teaching more than 50% of their timetable in physics, chemistry, biology, computer science, or foreign languages; received qualified teacher status between 2013/14 and 2020/21; employed in a state-funded school in one of 25 target local authorities. The policy was 'on' in the 2018/19 academic year onwards. Since the TSLR eligibility criteria are partly overlapping with those for RP (e.g. for certain physics teachers), we run sensitivity checks on our main empirical results in which we exclude all TSLR eligible areas.

The second contemporaneous reform is the Phased Maths Bursary (PMB) policy. The PMB policy involves paying £5,000 or more to eligible teachers in the third and fifth years of their career. Teachers were eligible for PMB payments if they met both of the following criteria: teaching more than 50% of their timetable in maths; received qualified teacher status in 2018/19. The PMB payments were made to eligible teachers in the 2021/22 and again in the 2023/24 academic years. Since the PMB eligibility criteria are partly overlapping with those for RP (e.g. for certain cohorts of maths teachers), we also run sensitivity checks on our main empirical results in which we exclude all PMB eligible cohorts. Eligibility for the RP, TSLR and PMB policies across years and subjects is summarised in Appendix Table A1.

DATA AND DESCRIPTIVE STATISTICS

We use the School Workforce Census (SWC), an administrative dataset containing information on all teachers employed in state-funded secondary schools in England. In particular, we analyse data on the five trainee cohorts of secondary teachers that qualified between 2014/15 and 2018/19 and we observe these teachers from the year they begin employment up to the

2020/21 academic year. Central government collects the information for the SWC from schools at the beginning of each academic year. We therefore infer that a teacher has exited the profession during year *t* if they appear in the SWC in year *t* but not in year *t+1*. This allows us to infer attrition up to the penultimate year in our data (2019/20), which is the first year in which the RP policy is 'on'. The SWC does not include information on teachers employed by independent (fee-paying) schools in England, which are attended by around 7% of pupils at any one time. Hence, our measure of attrition captures whether a teacher remains teaching in any state-funded school in England. The SWC also allows us to observe the criteria that determine eligibility for the RP policy, location of the school in which a teacher is employed, and demographic variables including gender, ethnicity and age.

Table 1 shows descriptive statistics for our sample during their first year in teaching, split by whether they are ever eligible (N=2,111) or never-eligible (N=66,858) for the RP policy. To provide a representative picture of the teachers in our sample, statistics are calculated for the first year in the profession across the five cohorts. First-year attrition in the ever-eligible group (21.9%) is slightly lower than attrition in the never-eligible group (22.5%). Gross teacher pay is 1.6% higher among never-eligible teachers, which likely reflects the fact that teachers based in London (who are not eligible) have higher minimum pay. Gender, age and ethnicity are all very similar between ever-eligible and never-eligible groups. More recent cohorts have more ever-eligible teachers because fewer teachers in these cohorts had left the profession by time the policy was first 'on' in 2019/20.

Table 1. Descriptive statistics for School Workforce Census sample

	Ever eligible for retention payments	Never eligible for retention payments
Attrition	21.9%	22.5%
Gross Pay (Mean (SD))	£24,194	£24,581
Female	64.5%	64.9%
Age (Mean (SD))	28.2 (7.5)	28.3 (7.6)
Ethnicity:		
White	67.6%	66.7%
Black	2.8%	3.2%
Asian	6.7%	6.8%
Other	22.9%	23.3%
Trainee cohort size:		
2014/15 cohort (N)	346	13,658
2015/16 cohort (N)	356	14,319
2016/17 cohort (N)	457	13,740
2017/18 cohort (N)	451	12,916
2018/19 cohort (N)	501	12,225
Total teachers (N)	2,111	66,858

Note: All statistics relate to the first year in employment as a teacher across each of the five trainee cohorts.

EMPIRICAL METHODS

The RP policy targets those teaching subjects (maths and physics) that are particularly likely to leave, working in areas that are particularly disadvantaged. Comparing retention rates among eligible and ineligible teachers is therefore likely to result in a downwardly biased estimate of the effect of the policy. The year in which the RP policy was introduced was also the first year of the pandemic, which temporarily increased retention (Zuccollo, 2021). Comparing the retention of eligible groups before and after the policy was switched ‘on’ is therefore likely to result in an upwardly biased estimate. To account for this, we leverage variation in eligibility across subjects, regions and years in a triple-difference framework. This enables us to capture changes in retention in eligible groups when the policy is switched on (first difference), relative to any changes in retention that occurred in eligible subjects but in ineligible areas (second difference), and in ineligible subjects (e.g. English) but in eligible areas (third difference). Under certain identifying assumptions (see below), this enables us to net out the change in retention for eligible teachers that would have occurred in the absence of the RP payments at the point the RP policy was switched on, leaving us with an estimate of the effect of the RP policy on retention.

We analyse how long a teacher remains in the profession. Such a duration-based outcome measure would be non-normally distributed and right censored (missing) for all teachers who had not left the profession by 2020/21 – both of which would be problematic for OLS regression. Survival analysis can handle both of these problems (Cleves et al., 2016). More precisely, we employ Cox proportional hazard models because they can account for the varying time at risk (years in teaching) across our five cohorts (Cox, 1972). In effect, we estimate the probability that a teacher leaves the profession in each year of our data, conditional on them not having left the profession in a prior year. Standard errors are clustered at the level of treatment assignment i.e., the teacher (Abadie et al., 2017). We implement this empirical strategy using the following estimating equation:

$$h(t) = h_0(t)\exp\,(\beta_1 E_a + \beta_2 \mathrm{E_s} + \beta_3 E_t + \beta_4 E_a E_s + \beta_5 E_s E_t + \beta_6 E_a E_t + \beta_7 E_a E_s E_t + \beta_8 X_{it})$$

where $\mathrm{h}(t)$ is the hazard function, which represents the probability of the teaching spell ending at time t, conditional on that spell lasting until $t-1$ and $h_0(t)$ is the baseline hazard i.e. the value of the hazard if all the covariates were equal to zero. E_a is an indicator variable equal to one for teachers in eligible areas, $\mathrm{E_s}$ is an indicator variable equal to one for teachers in eligible subjects, and E_t is an indicator variable equal to one for years in which the RP policy is switched on. We use subject taught, as opposed to subject of teaching qualification, for our third difference in order to maximise the precision of our estimates. Finally, X_{it} is a vector of sociodemographic covariates: gender, ethnicity, age, age squared.

The coefficient of interest β_7 captures the causal effect of the policy under the identifying assumption that relative attrition of subject-eligible and subject-ineligible teachers in eligible areas trends in the same way as the relative attrition of subject-eligible and subject-ineligible teachers in ineligible areas (Olden & Møen, 2022). Comparing the time series for eligible and ineligible groups is not a relevant test of the identifying assumptions of triple difference estimator (Olden & Møen, 2022). We therefore eschew such graphical tests, instead estimating several placebo tests which directly address the identifying assumptions of the triple difference estimator, albeit in the pre-treatment period. All units in our sample are treated at the same time, meaning that problems with time-varying treatment identified in the recent econometrics literature do not apply in our setting (Roth et al., 2022).

One obvious concern here is the pandemic, which had measurable effects on the teacher labour market in England (Zuccollo, 2021). However, we argue that our triple difference estimator is well-suited to addressing these threats. The general shock from the pandemic, which begins in the same year as the RP policy, is captured by the policy-on fixed effect (E_t) and is thereby netted out of our impact estimates. Any subject-specific Covid shocks, such as pandemic-induced changes in outside job opportunities for maths and physics teachers, is captured by the subject-year fixed effect ($E_s E_t$) and thereby differenced out of our estimates. Finally, area-specific Covid shocks, such as local variations in infection levels among pupils or teachers, is captured by the area-year fixed effect ($E_a E_t$) and thereby also differenced out of our estimates.

RESULTS

Table 2 shows the results from our triple difference models applied to the five cohorts of teachers in our SWC data. The coefficients in the table are hazard ratios, with a value below one representing a reduced risk of leaving the profession. Estimates from three different empirical specifications are presented. Column 1 reports estimates from the basic triple-difference specification with no additional controls. Column 2 incorporates cohort fixed effects, capturing cohort-specific shocks such as macro-economic changes, and region fixed effects, capturing local labour market conditions. Column 3 also incorporates the following four teacher covariates: gender, ethnicity, age, and age squared.

The RP Subject indicator shows that teaching in the targeted subjects (maths and physics) is associated with a 14-16% higher hazard of attrition. The RP Year indicator shows that the hazard of attrition was in general substantially (62-70%) lower than the other years. This likely reflects increased teacher retention following the onset of the COVID-19 pandemic in the 2019/20 academic year (Zuccollo, 2021). We are primarily interested in the coefficient on the 'RP Subject x RP Area x RP Year' interactions which, under our identifying assumption, captures the causal effect of being eligible for RP payments. Across all three specifications, the coefficient indicates that that eligibility for RP reduced the hazard of exit from the profession by 23%. This result is statistically significant at the 1% level. The insensitivity of this finding to the incorporation of additional covariates in columns 2 and 3 suggests that our triple-difference specification is doing a good job of accounting for potential confounders.

Table 2. Impact of being eligible for Retention Payments (RP) on attrition

	(1)	(2)	(3)
RP Subject x RP Area x RP Year	0.77***	0.77***	0.77***
	(0.07)	(0.07)	(0.07)
RP Year x RP Area	1.08	1.06	1.06
	(0.06)	(0.06)	(0.06)
RP Subject x RP Area	1.13**	1.12**	1.12**
	(0.05)	(0.05)	(0.05)
RP Subject x RP Year	0.95	0.93*	0.92**
	(0.03)	(0.03)	(0.03)
RP Year	0.38***	0.29***	0.30***
	(0.01)	(0.02)	(0.01)
RP Area	0.92	1.06	1.07
	(0.03)	(0.06)	(0.06)
RP Subject	1.14***	1.16***	1.15***
	(0.02)	(0.02)	(0.02)
Region fixed effects	No	Yes	Yes
Cohort fixed effects	No	Yes	Yes
Teacher covariates	No	No	Yes
No. of teachers	59,414	59,414	59,414
No. of observations	162,917	162,917	162,917

Note: Each column corresponds to a single regression. *= p<0.1. **=p<0.05. ***=p<0.01. Clustered standard errors shown in parentheses. Teacher covariates are gender, ethnicity, age, age squared.

Heterogeneity

Table 3 shows the results of our empirical tests for three types of heterogeneous treatment effects: by experience levels, by gender, and by degree subject. These three variables are shown in the rows of Table 3. To test for heterogeneity, we interact our main (triple-difference) eligibility indicator with each of the three dimensions of heterogeneity (years of experience, being female, and having a maths degree as opposed to a physics degree). As would be expected, each model also includes the corresponding 'main effect' for the three dimensions of heterogeneity and the main effect for the eligibility indicator. All models include cohort fixed effects, region fixed effects, and teacher demographics.

Prior research suggests that the pay-elasticity-of-exit is higher for less experienced teachers (Hendricks, 2014; Gilpin, 2011). However, the studies documenting this heterogeneity by experience rely on variation in pay across local government units responding endogenously to local labour market conditions. In this paper, since a) all teachers across our five trainee cohorts receive RP in 2019/20 and b) the RP policy was announced in 2018/19, we have plausibly exogenous variation in teacher experience at the time RP is paid. As can be seen from row 1 of Table 3, we find no such heterogeneity by experience within the first one to five years in the profession.

Table 3. Heterogeneity in impact of Retention Payments (RP) on attrition

	Interaction with eligibility	Eligibility main effect	No. of teachers (observations)
1) Experience (continuous)	1.00	0.80	56,392
	(0.01)	(0.11)	[128,211]
2) Female (binary)	1.07	0.76**	56,392
	(0.13)	(0.08)	[128,211]
3) Maths degree (binary)	0.93	0.80**	51,718
	(0.12)	(0.09)	[119,238]

Note: Each row corresponds to a single regression. *= p<0.1. **=p<0.05. ***=p<0.01. Clustered standard errors in parentheses. All models include cohort fixed effects, region fixed effects, teacher gender, ethnicity, age, age squared.

Previous research also suggests that the pay-elasticity-of-exit is higher for males (Hendricks, 2014; Gilpin, 2011). As can be seen from row 2 of Table 3, we find a coefficient consistent with existing findings, indicating a seven percentage point reduction (closer to 1) in the effect of RP on attrition for female teachers. However, this result is not statistically significant at conventional levels. Finally, we test whether having a graduate (or higher) qualification in maths (as opposed to physics) moderates the effect of RP eligibility on attrition. Our rationale for this is that the outside earnings ratio for teachers is a function of subject-specific expertise (Britton et al., 2016). However, once again, in row 3 of Table 3 we find no statistically significant difference.

ROBUSTNESS

These results identify the effect of RP eligibility on attrition under the identifying assumption stated above. Table 4 presents the results from a range of analyses designed to (indirectly) test this assumption by testing for placebo ‘effects’ of the policy in years, school subjects, and geographic areas that were not in fact eligible for the policy. If our models are netting out all influences on attrition besides the RP policy, then there should be no such placebo effects. Note that the number of observations varies across rows of Table 4 because we exclude genuinely eligible units in order to get a clean comparison between control and placebo-treated units.

Rows 1 and 2 test whether there is a placebo ‘effect’ one year and two years prior to the RP policy being ’on’ (2019/20). In both cases, the coefficients are close to one (no effect) and are not statistically significant at conventional levels. In row 3, we add a linear term for time and interact this with the subject (maths and physics) eligibility indicator. This captures any subject-specific differences in trends in the spirit of a comparative-interrupted time series specification (St. Clair & Cook, 2015). The coefficient on this interaction is also close to one (0.98) and not statistically significant (p=0.24). Further, the estimated treatment effect (shown in the table) remains unchanged in this specification. In sum, the evidence in rows 1-3 is consistent with the identifying assumptions, albeit in the pre-treatment period.

Rows 4-5 in Table 4 test whether we detect any placebo ‘effect’ for groups of teachers who taught subjects that were not in fact eligible for the policy. In row 4, we use English teachers as the placebo subject on the grounds that they are the largest single subject-group of teachers outside maths and science. The coefficient is quite close to 1 (1.09) and not statistically significant at conventional levels. In row 5, we combined two of the next largest groups (history and geography teachers) and find a similar lack of any placebo effect. Finally, row 6 in Table 4 tests whether we detect a placebo 'effect' in the two most deprived non-eligible regions: the North East and the East Midlands. The coefficient is once more very close to one. In sum, the evidence in rows 4-6 is consistent with the identifying assumptions.

Table 4. Robustness tests

		Treatment effect	No. of teachers [observations]
Time	1) One year lag placebo	0.95 (0.10)	59,413 [162,914)
	2) Two year lag placebo	1.06 (0.11)	59,413 [162,914]
	3) Subject-specific linear pre-trends	0.77** (0.08)	59,413 [162,914]
Subject	4) English teacher placebo	1.09 (0.14)	39,747 [107,103]
	5) History/geography teacher placebo	0.99 (0.14)	39,747 [107,103]
Area	6) North-east & East-midlands placebo	0.981 (0.09)	54,325 [167,517]

Note: Each row is to a single regression. *= p<0.1. **=p<0.05. ***=p<0.01. Clustered standard errors shown in parentheses. All models include cohort and region fixed effects, teacher gender, ethnicity, age, age squared.

Table 5 presents a final set of checks on our main results, this time testing for sensitivity to various sample restrictions and model specifications. In the sections on contemporaneous reforms above, we described two pay reforms that partially overlap with RP. This raises the concern that our estimates of the effect of RP on attrition may in part reflect the impact of these other policies.

In row 1 of Table 5 we re-estimate our models excluding all areas in which teachers were potentially eligible for Teacher Student Loan Reimbursements (TSLR). In row 2 we exclude the cohort (2018/19 trainees) who were potentially eligible for the Phased Maths Bursaries (PMB). In both cases, the coefficient on RP eligibility remains essentially unchanged and statistically significant at the 5% level.

Recalling Figure 1, it is clear that eligible areas have been carefully selected for the RP policy. A related concern is that policymakers may have selected these areas because they were experiencing particular difficulties around teacher retention, perhaps related to local labour market conditions. Recall that the eligible areas were based on two deprived regions, plus nine further local authorities drawn from a list of the 39 most deprived in England. In row 3 of Table 5, we restrict the sample to the two eligible regions plus the full set of the 39 most educationally disadvantaged local authorities in England. This sample restriction is intended to approximate the geographic aspect of treatment selection. Row 3 shows that the treatment effect is very similar (0.73) and remains statistically significant at the 5% level when we employ this restricted sample.

With respect to our modelling approach, a concern is that our use of Cox proportional hazards models may be sensitive to the presence of tied failure times (see Cleves et al., 2016). To test this, in row 4 of Table 5 we report the results from a Weibull survival analysis model. This specification models the failure times (rather than failure order) and thus does not require assumptions about how to deal with tied failure times. To do so, it imposes parametric assumptions on our model. Despite this, our estimate of the treatment effect is essentially unchanged.

A final sensitivity check relates to our triple-difference specification. Ideally, we would have incorporated a fourth difference - the subject in which a teacher trained - in our research design. However, missing data interacting across the four binary indicators meant this would have left us with inadequate statistical power. For the sake of transparency, in row 5 of Table 5, we report the results from this four-diff specification. The coefficient remains negative, though declines from a

23% reduction in the hazard to a 14% reduction in the hazard. However, the standard error inflates by a factor of six, meaning that the coefficient is very imprecisely estimated and is no longer statistically significant at conventional levels. While this result reflects an important limitation of our data, we take some reassurance from the coefficient remaining negative.

Table 5. Sensitivity checks

	Treatment effect	No. of teachers [observations]
1) Excluding TLSR eligible areas	0.76**	52,363
	(0.04)	[142,217]
2) Excluding PMB eligible cohorts	0.75**	47,962
	(0.09)	[142,330]
3) Disadvantaged areas only	0.73**	15,116
	(0.10)	[40,182]
4) Weibull parametric regression	0.78**	59,413
	(0.08)	[162,914]
5) Four-diff specification	0.86	54,028
	(0.44)	[153,368]

Note: Each row is to a single regression. *= p<0.1. **=p<0.05. ***=p<0.01. Clustered standard errors shown in parentheses. All models include cohort and region fixed effects, teacher gender, ethnicity, age, age squared.

COST-EFFECTIVENESS

Having established how the RP policy affects attrition, we turn now to consider its cost-effectiveness. That is, we aim to establish the cost of each additional teacher retained as a result of the policy (Bellfield & Brooks Bowden, 2018). To do so, we run a policy simulation using a pre-policy (2014/15) maths trainee cohort, using the following three simplifying assumptions. First, maths teachers are eligible for the policy in two years: once if they remain teaching in a state funded school after their first year of employment in the profession, and again after their second in the profession. Second, in line with empirical evidence (Bueno & Sass, 2018; See et al., 2018), we assume that the retention incentives do not influence the decision to train as a teacher and that hazard ratios remain unchanged in all years in which teachers are not eligible for the policy. Third, we adopt the simplifying assumption that the policy costs £2,000 per eligible teacher, per year.

In Table 6, we begin by simulating the number of additional maths teachers that would be retained per cohort of new maths teachers. Panel I reports actual data for the cohort of maths teachers that trained in the 2014/15 academic year.[ii] Panel II reports simulated data for the counterfactual situation with the retention payment policy switched on. Column 1 shows the actual number of recruits to initial teacher training. Column 2 shows the actual number of newly qualified teachers entering the profession in year 1 of their teaching careers. Columns 3 and 4 show the number of teachers who remain in the profession in the second year and third year after qualification. Columns 3 and 4 differ between panel I and panel II because the hazard of leaving has been reduced by 23% in panel II, based on the results from Table 2 above.

Table 6. Simulated effects of the policy on a single maths trainee cohort

		(1) Trainees	(2) Year 1	(3) Year 2	(4) Year 3
I. Factual	A) No. of teachers	2,170	1,746	1,438	1,266
	B) Hazard	0.20	0.18	0.12	
	C) No. of leavers	424	308	172	
II. Counterfactual	D) No. of teachers	2,170	1,746	1,509	1,370
	E) Adjusted hazard	0.20	0.14	0.09	
	F) No. of leavers	424	237	139	
III. Costs/effects	G) Extra teachers	0	0	71	104
	H) Cumulative cost (£k)			3,492	6,509
	I) Cost / extra teacher (£k)			49	63

Note: A) is the actual number of teachers for the 2015/16 maths trainee cohort. C) is the actual number of teachers that leave prior to the next academic year. B) is C/A. E) is B) reduced by 23%, based on results from Table 2. F) and D) are the simulated number of teachers leaving and remaining (respectively) based on E). G) is the difference between D) and A). H) is D) multiplied by £2,000. I) is G/H. Year 1 is what is referred to in England as the 'NQT' (newly qualified teacher) year.

Panel III of Table 6 shows our main results. Row G shows that the policy increases the number of maths teachers remaining in this cohort in year 3 by 104. Row I shows that the cost per additional teacher remaining in this cohort in year 3 is £63k. This is approximately 16 times the total value of payments per teacher (£4,000), which reflects the fact that only 8% of teachers receiving the second retention payment are additional teachers. Based on detailed existing empirical work (Allen et al., 2016), we estimate that costs of recruiting and training an equivalent replacement teacher would be £92,000.[iii] The cost per additional teacher retained through the policy is therefore 32% lower than recruiting an equivalent replacement teacher.

DISCUSSION

Public school systems experience recurring shortages of maths and science teachers, in part because those qualified to teach these subjects can attract higher pay in occupations outside teaching (Dolton, 2020; Gilpin, 2011; Marachand & Weber, 2020). Policymakers in various jurisdictions have responded by implementing targeted pay supplement policies aimed at keeping shortage subject teachers in the profession (Dee & Goldhaber, 2017). While there is a growing literature evaluating such policies, there remains considerable uncertainty around their effects in different settings, the ways in which such policies should be targeted, and their cost-effectiveness. We set out to address these gaps in the literature by evaluating the cost-effectiveness of a retention payment policy worth around 8% of salary per annum for early-career maths and science teachers in England.

We found that this policy improved teacher retention. Indeed, our results suggest a that a one percent increase in pay results in a three percent reduction in attrition, or a pay-elasticity-of-exit of -3. This finding is robust to various placebo tests and specification checks and mirrors the results from evaluations of similar policies in Florida (-2.5 to -3.3; Feng & Sass, 2018) and Georgia USA (-3.4; Bueno & Sass, 2018). Indeed, our results can be considered a conceptual replication of these findings in a different population and labour market setting. Our findings should of course be interpreted in light of the limitations of our research. Two in particular stand out. First, we lack statistical power to estimate the ideal four-diff model. Second, we evaluate only the short-run effects of the policy; our research is silent on what happens to the additional teachers when they are no longer eligible for payments. Nevertheless, taken together these results suggest that policymakers in a range of settings can likely reduce shortages of maths and science teachers by providing targeted salary supplements. Shortages of teachers in subjects with high outside earnings potential are therefore not an inevitable feature of public schooling systems.

Implementing such policies raises interesting policy design issues around targeting of the payments. Existing research (Gilpin, 2011; Hendricks, 2014) suggests that teachers with fewer years of experience in the profession are more sensitive to pay when making decisions about whether to remain in teaching. This implies a greater return on investment for targeting the policy on teachers in their first few years of their career. However, the studies documenting this heterogeneity by experience rely on variation in pay across local government units responding endogenously to local labour market conditions. An important contribution of our paper is that we are able to exploit plausibly endogenous variation in experience at the time the payments are made in order to test this across teachers with 1-5 years of experience in the profession. We find no such heterogeneous effects, which suggest that pay supplement policies are likely to be equally effective within this range. Having said that, the findings from existing research suggest that the effects of increased pay on exit disappear entirely among those with more than ten experience, suggesting targeting early-career teachers (broadly defined) remains a sensible approach (Hendricks, 2014).

A further contribution of our paper is in analysing the cost-effectiveness of the retention payment policy. Our simple policy simulation sheds light on the costs of retaining each additional teacher. Despite the policy paying £2,000 per annum to eligible teachers, the high levels of deadweight loss mean that the cost per additional teacher with two years of experience in the profession is approximately £63,000. This is nevertheless 32% less than the cost of recruiting and training an equivalent replacement teacher in England, which is approximately £92,000 per teacher. However, this assumes that it is indeed possible to recruit such a replacement teacher, which is far from clear in shortage subjects such as maths and science.

Another way of putting this figure into perspective is to compare it to the cost of raising the starting pay of teachers generally. For example, the UK government has recently announced its intention to raise all teachers' starting salaries to £30,000 by 2024, which would represent a 20% nominal increase over 2020/21 levels. However, since teachers in many other subjects do not tend to be in shortage (Dee & Goldhaber, 2017; Worth & Faulkner-Ellis, 2021), such general pay increases are likely to have higher levels of deadweight loss. Taken together, the above suggests that targeted pay supplements are likely to be more cost-effective than alternative policy options when it comes to reducing subject-specific shortages of teachers.

REFERENCES

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2017). *When should you adjust standard errors for clustering? (NBER Working Paper No. 24003).* National Bureau for Economic Research.

Allen, B., Bellfield, C., Greaves, E., Sharp, C., & Walker, M. (2016). *The longer-term costs and benefits of different of different initial teacher training routes*. Institute for Fiscal Studies.

Anders, J., Bryson, A., Horvath, H., & Nasim, B. (2019) *The effects of pay decentralisation on teachers' pay and teacher retention.* https://conference.iza.org/conference_files/Education_2019/22370.pdf

Bellfield, C., & Brooks Bowden, A. (2018). Using resource and costs considerations to support education evaluation: six domains. *Educational Researcher, 48*(2), 120-127.

Bueno, C., & Sass, T. (2018). *The effects of differential pay on teacher recruitment and retention* (Andrew Young School of Policy Studies Research Paper Series No. 18-07). Georgia State University.

Britton, J., Dearden, L., Shephard, N., & Vignoles, A. (2016). *How English domiciled graduate earnings vary with gender, institution attended, subject, and socio-economic background (IFS Working Papers No. W16/06).* Institute for Fiscal Studies.

Chevalier, A., Dolton, P., & McIntosh, S. (2007). Recruiting and retaining teachers in the UK: An analysis of graduate occupation choice from the 1960s to the 1990s. *Economica, 74*(293), 69-96.

Cleves, M., Gould, W., & Marchenko, Y. (2016). *An introduction to survival analysis using Stata*. Stata Press.

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B, 34*(2), 197-202.

Dee, T., Goldhaber, D. (2017). *Understanding and addressing teacher shortages in the United States.* The Hamilton Project.

Dolton, P. (2020). *Teacher supply.* In: Bradley, S., & Green, C. (eds.) *The Economics of Education: A comprehensive overview*. Academic Press.

Dolton, P., & Chung, T. (2004). The rates of return to teaching: how does it compare to other graduate jobs? *National Institute Economic Review, 190,* 89-103.

Feng, L., & Sass, T. (2018). The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects. *Journal of Policy Analysis and Management, 37*(1), 112-135.

Gibbons, S., Vincenzo, S., Shqiponja, T. (2021). Teacher turnover: Effects, mechanisms and organisational responses. *Labor Economics, 73,* 102079.

Gilpin, G. (2011). Reevaluating the effect of non-teaching wages on teacher attrition. *Economics of Education Review, 30*(4), 598-616.

Hendricks, M. (2014). Does it pay to pay teachers more? Evidence from Texas. *Journal of Public Economics, 109,* 112-135.

Jerrim, J., & Sims, S. (2019). *The Teaching and Learning International Survey (TALIS) 2018*. Department of Education, UK.

LiVecchi, A. (2017). *The opportunity cost of teaching for secondary STEM instructors (Doctoral dissertation).* UCL.

Marachand, K., & Weber, J. (2020). How local economic conditions affect school finances, teacher quality, and student achievement: evidence from the Texas shale boom. *Journal of Policy Analysis and Management*, *39*(1), 36-63.

Migration Advisory Committee [MAC]. (2017). *Partial review of the Shortage Occupation List: teachers*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/585998/2017_01_26_MAC_report_teachers_SOL.pdf

Olden, A., & Møen, J. (2022). The triple difference estimator. *The Econometrics Journal*, https://doi.org/10.1093/ectj/utac010

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4-36.

Roth, J., Sant'Anna, P., Bilinski, A., & Poe, J. (2022). *What's trending in difference*-in-*differences? A synthesis of the recent econometrics literature*. https://arxiv.org/abs/2201.01194

School Teachers' Review Body [STRB]. (2019). *Twenty-ninth report*. https://www.gov.uk/government/publications/school-teachers-review-body-29th-report-2019

See, B., Morris, R., Gorard, S., & El Soufi, N. (2020). What works in attracting and retaining teachers in challenging schools and areas?. *Oxford Review of Education, 46*(6), 678-697.

Sims, S. (2018). *Essays on the recruitment and retention of teachers (Doctoral dissertation).* UCL.

St. Clair, T., & Cook, T. (2015). Difference-in-differences methods in public finance. *National Tax Journal, 68*(2), 319.

Worth, J., & Faulkner-Ellis, H. (2021). *Teacher Labour Market in England.* NFER.

Zucollo, J. (2021, June). *Teacher recruitment and retention in the eye of the pandemic* [Blog post]. Education Policy Institute.

APPENDIX

Table A1. Eligibility for teacher pay supplement policies in England

Subject	Trainee Cohort	Academic Year 2017/18	2018/19	2019/20	2020/21	2021/22
Physics	2014/15			RP		
	2015/16			RP		
	2016/17			RP	RP	
	2017/18			RP	RP	
	2018/19			RP	RP	
Maths	2014/15			RP		
	2015/16			RP	RP	
	2016/17			RP	RP	
	2017/18			RP	RP	
	2018/19			RP	RP	PMB
Biology Chemistry Computing Languages	2014/15					
	2015/16					
	2016/17					
	2017/18					
	2018/19					

Note: Shaded grey areas represent eligibility for the Teacher Student Loan Reimbursement policy. RP indicates eligibility for the Retention Payment policy. PMB represents eligibility for the Phased Maths Bursary policy. Subjects not mentioned in the table were not eligible for any supplementary pay policies during the period covered by the table.

[i] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/650036/Opportunity_areas_selection_methodology.pdf

[ii] The number of teachers entering maths ITT in 2014/15 is taken from Table 2 here https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/478098/ITT_CENSUS_SFR_46_2015_to_2016.pdf and the retention figures for that cohort are taken from here: https://department-for-education.shinyapps.io/turnover-and-retention-grids/

[iii] Allen et al. (2016) estimate the costs of training a maths teacher as £29,000 in 2022 prices. Maths trainees also qualify for a £24,000 training bursary. This brings the total cost per trainee to £53,000. Accounting for attrition across initial training and the first two years in the profession, generating an additional 104 teachers with 2 years of experience requires 180 additional initial teacher trainees. This makes the cost per additional teacher £91,730.