Effect of Performance Related Pay on turnover-intention of teachers in England

Ajala, Olubunmi A[*] and Asako Ohinata[†]

January 31, 2019

Abstract

We evaluate the impact of the 2013 performance-related pay policy (PRP) in England on teachers' intention to leave the job. Using the 2007-2018 Quarterly Labour Force Survey, and a difference–in–differences estimator, we find that the policy has resulted in a significant reduction in the teachers' intention to leave their job by 1.16 percentage points, which amounts to a reduction in the magnitude of approximately 13%. We found no heterogeneous policy impacts across gender, however, we find differential effects across teachers educational background.

Keywords: Performance-related pay; UK education; LFS; Teachers retention; PRP 2013

JEL classification codes: J33, J45, I21, I28

[*]Corresponding author: Department of Economics, University of Leicester, Leicester, LE1 7RH, The United Kingdom. E-mail address: oaa37@le.ac.uk (Ajala, Olubunmi).

[†]Department of Economics, University of Leicester, The United Kingdom.

1 Introduction

The policy of tying teachers pay to their students' academic performance is often contentious. In the UK, the proposal to implement such a policy has always been heavily attacked by teachers unions and some academic observers (Richardson, 1999). Nevertheless, its implementation is on the increase in Britain and USA, (Brown and Armstrong, 1999)). The proponents of the policy envisage that, higher academic standards can be achieved in the education system through a rigorous performance management system (Storey, 2000).

Prior empirical studies have exclusively focused on the impact of such policies on students' academic performances. (Atkinson, Burgess, Croxson, Gregg, Propper, Slater, and Wilson, 2009; Britton and Propper, 2016). The majority of the evidence on educational outcomes comes from the US (Figlio and Kenny, 2007a; Springer, Ballou, Hamilton, Le, Lockwood, McCaffrey, Pepper, and Stecher, 2011; Fryer, 2011), which have also produced mixed results. Atkinson, Burgess, Croxson, Gregg, Propper, Slater, and Wilson (2009) and Britton and Propper (2016) both investigate the 1999 English performance related pay policy for teachers, which was introduced for experienced teachers and conclude that the policy had a positive impact on the academic performances of students aged between 14 and 16.

In contrast to the existing literature, our study focuses on the teachers. More specifically, we investigate the impact of the 2013 PRP implementation in England on teachers' intention to leave their job. In the UK and US, educational policy debates often suggest that teachers productivity is low in part due to the high turnover rate of teachers, particularly among the new and young (Wayne and Youngs (2003), Hanushek, Rivkin, Rothstein, and Podgursky (2004). Studying the impact of the PRP policy, therefore, allows us to directly investigate one of the policy objectives. In addition, studying the relationship between teachers' pay scheme and teacher retention allows us to investigate whether an implementation of a particular pay scheme can be used as a potential way to reduce teacher shortage problem. Since teacher shortage is an important problem in many countries (Ingersoll, 2003; Sutcher, Darling-Hammond, and Carver-Thomas, 2016; Lindqvist and Nordänger, 2016; Posey, 2017; den Brok, Wubbels, and Van Tartwijk, 2017).

We exploit the 2013 PRP policy, which was introduced in England. The policy affected primary and secondary school teachers whose pay was below a certain threshold on the national pay scale. Prior to the 2013 policy, the pay of these teachers were determined mainly by the years of experience. Teachers were allocated an automatic reward of one point for each year. After 2013, the pay of these teachers became closely tied to their students academic performance. In contrast, those receiving a higher pay have been exposed to performance related pay scheme since 1999. The latter group, therefore, can be used as a control group to evaluate how the 2013 policy influenced the newly affected teachers intention to leave their job.

We use the 2007-2018 Quarterly Labour Force Survey, covering 6 years before and 6 years after the 2013 PRP implementation and employ a difference-in-differences estimator. We find that the probability of leaving teaching significantly declined by 1.16 percentage points, which amounts to approximately 13 % decline in the affected teachers' intention to leave. We also investigate heterogeneous impacts of the policy by region, ethnicity and gender. We find that the effect for those in inner and outer London to be significantly different, as those in inner London have 4 percentage points higher probability of intention to leave the job than those in outer London. The policy impact is also different by race when we compared white with other races. The probability of looking for a new job for white individuals affected by the policy is 3.6 percentage points higher than non-white affected by the policy.

The paper is organised as follows. Section 2 provides background information on the 1999 and 2013 PRP implementations in England. Section 3 reviews the related literature. Section 4 presents the econometrics specification and the data employed. Section 5 reports the results and robustness checks while section 6 provides the conclusions.

2 PRP implementation of 1999 and 2013

Prior to the 1999 PRP implementation, teachers' pay was determined by the year of experience and qualification.Teachers were allocated an automatic reward of one point for each year, with a possibility of a fast track by allocation of two points for a good honours degree on entry (Storey, 2000). The PRP has been implemented twice, once in 1999 and another is 2013.

Under the 1999 implementation, the salary of the experienced teachers was determined by their performance. Teachers in their early teaching careers continued to progress on an incremental scale, until they get to a higher payment threshold. Table 2.1 shows the pay scale in year 2000 (in 2013 equivalent, after we have taken inflation rate into consideration). The 1999 PRP was implemented for those who were above pay scale 9. In contrast, those receiving pay below pay scale 9 continued to receive pay in a similar manner as before 1999. Beyond the threshold, there was no more incremental scale, progression in the pay scale depended on demonstrations of performance.

From September 2013, DfE required every public school in England and Wales to revise its pay and appraisal policies, setting out how pay progression be linked to a teachers performance (DfE, 2013). The new implementation, tacitly moved wage determination of teachers away from Local Educational Authorities (LEAs) to the school level. The 2013 implementation also collapsed the 9 spines under the main pay scale into 6 and further collapsed the 5 scales under the upper pay scale into 3. Table 2.2 shows the pay structure for England and Wales. The implementation also removed the threshold at which performance based progression kicks in, as all teachers can progress across scales, and do not have to mandatorily go through pay scale. Likewise, those already on upper pay scale 1, do not have to go through upper pay scale 2 to get to upper pay scale 3 (if their performance warrants the progression).

We underscore two main differences brought about by 2013 policy, which are in terms of coverage and progression rate. The first difference in terms of coverage is that, the 1999 implementation affected only teachers from the threshold referred to as, the Upper Pay Scale. The 2013 implementation however covers all teachers, thereby including all those excluded in the 1999 implementation. This extended-coverage translates to about 139,260 newly affected teachers. This is induced from Table 2.3: 38.1% of 365,500 teachers that were initially excluded from PRP were affected by the new implementation. Secondly, the guideline that, teachesr do not necessarilty have to progress through M2 – M5 and U2 (as shown in Table 2.2) , and the policy opens up the possibility of a rapid progression for high performers by skipping some of these pay scales. This implies that the relatively young and inexperienced teachers can progress to upper pay scale without waiting for 6 - 9 years, which was previously required.

The 2013 implementation, mostly affected teachers below age 30. This can first be inferred from the fact that, it takes 6 - 9 years to progress through the main pay scale and graduation age range of 21 - 22, for most UK graduates. Those between age 30 and 34 are evenly distributed between the affected and non-affected groups. Table 6.3 shows, the distribution of teachers by age and pay scale in public funded schools. The highlighted area represents, those newly affected by 2013 PRP implementation. The table reveals that the 2013 implementation, mostly affected teachers below age 30.

Table 2.1: Class Teacher Pay Scale Post 1999 PRP Implementation in 2013 Price Equivalent

	Spine Point	Sept. 2000
Main Pay Scale	1	22,257
	2	23,594
	3	24,877
	4	26,231
	5	27,792
	6	29,468
	7	31,236
	8	33,110
	9	35,218
Threshold		
Upper Pay Scale	1	38,160
	2	39,571
	3	41,035
	4	42,552
	5	44,126

Notes: Data is extracted from Guidance Notes of School Teachers Review Body. The table excludes additional payments and allowances pay for teachers. We do not include pay range for those already classified as "Leading Practitioner Teachers" and those on "Leadership Scale". From 1999 PRP implementation,Threshold for PRP starts from Upper Pay Scale

We have converted year 2000 payscale into 2013 equivalent. According to the Office for National Statistics composite price index, prices in 2013 are 46.87% higher than prices in 2000. CPI in 2000 was 671.800 and CPI in 2013 was 986.700, so we used 1.47 factor price. The shaded part represents the group that were not affected by 1999 PRP

Table 2.2: Pay Structure for Qualified Teachers Sept./2013 − Sept./2014

	England & Wales
Main Pay Range	**£p.a.**
M1 Minimum	21,804
M2	23,528
M3	25,420
M4	27,376
M5	29,533
M6 Maximum	31,868
Upper Pay Range	
U1 minimum	34,523
U2	35,802
U3 maximum	37,124

Notes: Data is from Pay Scales 2013 released by National Union of Teachers. The table excludes additional payments and allowances pay for teachers. We do not include pay range for those already classified as "Leading Practitioner Teachers" and those on "Leadership Scale". From 2013 PRP implementation, points M2-M5 and U2 are not now mandatory points

We have used the Fringe areas data because it is the median figures for England and Outer London. Teachers in inner London earn about £4,000 more than teachers in fringe areas, on same pay scale. The shaded part represents the affected group by 2013 PRP

Table 2.3: Class Teacher Headcount prior 2013 PRP Implementation

	Men and Women									
	Percentage of teachers in each age band									
	Classroom teachers' pay scales									
	Main pay scale			Upper Pay Scale			Other			
	1 - 3	4 - 5	6	1	2	3		Leadership	Total %	Total ('000)
TOTAL PUBLICLY FUNDED SCHOOLS										
Under 25	88.9	3.6	0.1	0.1	-	0.1	7.0	0.2	100.0	23.1
25-29	40.2	29.7	14.8	6.2	0.3	0.1	7.2	1.6	100.0	71.1
30-34	12.0	11.3	16.9	23.3	14.2	5.6	7.5	9.1	100.0	65.0
35-39	7.0	5.3	9.4	13.5	13.7	24.9	7.1	19.2	100.0	48.2
40-44	6.7	4.6	7.9	9.6	8.9	30.6	t7.0	24.7	100.0	44.1
45-49	5.1	3.7	8.1	9.2	8.5	33.1	6.4	25.9	100.0	40.1
50-54	2.4	1.9	6.0	7.2	7.4	40.6	6.5	27.9	100.0	36.1
55-59	0.9	0.6	3.9	4.4	6.1	46.4	6.2	31.4	100.0	30.8
60 and over	1.4	0.5	5.6	4.9	5.7	44.9	7.5	29.3	100.0	7.0
All ages	18.2	9.9	10.0	10.5	7.8	20.4	7.0	16.3	100.0	365.5

Notes: Data is from School Workforce Census - 2013 released by Department for Education (SFR15/2013). Full-time regular qualified teachers in public funded schools by salary bands as at November 2012, before the new policy set in. The other column, according to DfE includes a proportion of teachers, on other pay spines (Unqualified Teacher, Excellent Teachers and those that are not specified). The shaded part represents the affected group by 2013 PRP

3   Literature review

The vast majority of empirical studies on performance-related pay for teachers (Lavy (2002), Podgursky and Springer (2007), Figlio and Kenny (2007b), Glewwe, Ilias, and Kremer (2010), Fryer (2013)) have been outside England. While large evidence exists in the US, the application of merit pay for teachers in US is different from its implementation in British schools. Most incentive pays in US are inform of bonuses to existing pay, but in England it comes as a permanent increase to the teacher's salary. Also, while application has been on selective schools or states in the US, its implementation covers all public funded primary and secondary schools in England and Wales. This gives credence to the need to undertake studies into merit pay implication in British schools

There exists few literature on performance pay for teachers in England (Storey (2000), Tomlinson (2000), Cutler and Waine (2000), Chamberlin, Wragg, Haynes, and Wragg (2002), Wragg, Haynes, Chamberlin, and Wragg (2003), Mahony, Menter, and Hextall (2004)), but very limited expirical works and they all focused on impact of the policy on academic performance(Atkinson, Burgess, Croxson, Gregg, Propper, Slater, and Wilson, 2009), Britton and Propper (2016)). To the best of our knowledge, the few empirical works have all focused on the 1999/2000 implementation and not the full implementation of 2013. Similarly, we have no evidence that, any prior paper has investigated the link between full PRP implementation and teachers' intention to remain or leave their job.

As stated earlier, most studies in UK with respect to PRP are not empirical. Storey (2000) analysed the response, from 36,060 individuals consisting of 66% teachers, 31% head teachers and 3% governors obtained by the Department for Education and Employment. Responses were analysed with other submissions made by trade unions, Members of Parliament (MP), Local Education Authorities (LEAs), among other stake holders and concluded that, the new proposals will confront some considerable difficulties and suggested that head teachers and their deputies would need to bend the scheme to make it suitable to their organisational requirements. This study is purely a quantitative analysis ahead of the policy implementation.

Mahony et al., (2004), shared insights into emotional impact of PRP on teachers in England from 76 interviews undertaken in 9 case study primary and secondary schools between

2001-2003. They reviewed, documentary analysis and some tape-recorded interviews with a range of key actors. Their research revealed a number of unintended consequences of PRP, such that, the increased focus on pupil progress carries a danger that, other elements that are educationally important will either become marginal or be in conflict with teachers' personal and professional values. More like Mahony, Menter, and Hextall (2004), Wragg et al., (2003) reviewed views and experiences of randomly sampled 1,000 head teachers, covering over 150 local education authorities in England, who were responsible for implementing PRP. They concluded that 60% of head teachers were against it and 40% for it. They added also that, those who supported it in principle expressed concerns about the assessment procedure.

While earlier papers are considered quanlitative, the first of empirical study in England evaluating the impact of PRP system for teachers on students scores is Atkinson, et al.,(2009). They surveyed 18 schools, covering 181 teachers and about 5000 pupils between year 2000 and 2002. They collected data for their sample lists back to year 1997. They adopted a difference-in-differences method and concluded that the scheme improved test score gains by about 40% on average, except for mathematics, where students scored showed no improvement. Besides, the small sample size of their study, their study only measured the first year of implementation, when the policy came into effect. Question about persistency of policy effect after initial reaction cannot be answered, by this study.

Britton and Propper (2016), covered more years in their study of investigating the impact, wage rigidity across regions, has on school productivity. They explore the gap between the outside labour market wage and regulated wage for teachers, to investigate its effect on students' score. They use data of about 200,000 teachers covering 3000 schools. Using the proportion of pupils in the school, who achieved 5 GSCEs at grades A*-C as a measure of school performance and the difference between inside and outside wage (represented by average wage of local authorities within 30 kilometre radius of the school) and concluded that, a ten percent shock to the wage gap between teachers pay and external wage results in loss of around 2% in average school performance. This study also limits its scope to 1999/2000 PRP.

Besides earlier works limiting their analysis to 1999 PRP, no study, to our knowlede has linked PRP with turnover-intention of teachers in England. But there exists body

of literature in general, investigating teachers pay and their career decisions (Chapman (1984), Adams (1996), Guarino, Santibanez, and Daley (2006)). Chapman (1984), conducted a test on determinants of teacher retention, using discriminant analysis among three groups, (a) those who have been teaching continuously, (b) those who left within 5 years of starting, and (c) those who never taught. He built his influences of retention using, social learning theory interacting, personal characteristics, previous behaviour, and environmental factors. He concluded that those who left teaching within 5 years are closer in characteristics with those who have never taught, while those who have been teaching continuously, reported initial commitment to teaching. Teachers were adjudged to assign less importance to salary as a measurement of professional success. He also found that, those who left teaching tend to be female and older.

Murnane, Singer, and Willett (1989) investigate, factors affecting teachers career path, using data of 5,100 white elementary and secondary school teachers in North Carolina public schools between 1976 and 1978. They estimate a series of hazards models to investigate the relationship between the risk of leaving teaching, and teacher salary and opportunity costs. They concluded that, pay is an important predictor of whether teachers remain in teaching, especially during their first years on the job.

While some of the literature have been analysing salary and teachers disposition to remain in teaching or not, those who try to link PRP with such decision attempt, assessing impact of PRP on teachers satisfaction. Belfield and Heywood (2008) used Schools and Staffing Survey of 2000, to test three hypotheses about determinants and consequences of PRP. Using ordered probit, they found merit pay has a positive effect on teachers satisfaction with the salary but has no significant effect on a teachers satisfaction with either teaching or effort. (Belfield and Heywood, 2008)

Gius (2013), investigated if district-level merit pay has any effect on teacher job satisfaction. Using Schools and Staffing Survey 2007, compiled by the US Department of Education and conducted an ordered probit analysis and concluded that teachers who received merit pay were more satisfied overall than teachers who did not receive merit pay although, the effect is insignificant. He also concluded that, teachers in merit pay districts are more likely to leave the job for better pay.

4 Econometric specification and data

4.1 Difference-in-differences model

The 2013 PRP extends the policy of tying teacher's pay to performance to all teachers in primary and secondary schools in England. This is in contrast to the 1999 policy, which only subjected those on the upper pay scale to the PRP scheme. Therefore, we evaluate the impact of this policy extension on the newly affected teachers' intention to leave the job. In Section 2, we highlighted that the majority of teachers affected by the 2013 PRP policy was aged below 30. We employ a difference-in-differences estimator comparing experienced teachers older than 35 to those aged below 30 and evaluate how the difference in the intention to leave across these two groups of teachers changed over the years.

Our empirical evaluation uses a cross section framework and estimate the following model using a Linear Probability Model (LPM) for ease of interpretation. Our model of turnover-intention y, for an individual i, at a particular time t, is represented by:

$$y_{it} = \boldsymbol{X}'_{\boldsymbol{it}}\boldsymbol{\beta} + \Phi_t + \delta_{DD}I_t + \epsilon_{it} \tag{1}$$

where:

- X'_{it} is the K x 1 vector of individual characteristics, employment characteristics and educational qualification and β is the conformable vector of coefficients. The regressors in X'_{it} are age, gender, race, length of employment with employer, zones of location of job and educational qualification.

- Φ_i is a set of time fixed effects (year dummies)

- I_{it} is our variable of interest which is equal to 1 if individual is affected by the policy after full implementation of PRP (2013). The parameter δ_{DD} is the effect of full PRP implementation on intention of "young teachers" to leave the job.

- ϵ_{it} is the the error term at individual level

We intend to also estimate policy effect differential across gender, race and regions of work by estimating an extended version of equation (1). For example, for gender, we interact I_{it} with female to estimate the heterogeneous effects of the policy.

9

$$y_{it} = \boldsymbol{X}_{\boldsymbol{it}}'\boldsymbol{\beta} + \Phi_t + \delta_{DD}I_t + \delta_{2DD}(Female_{it})I_t + \epsilon_{it} \tag{2}$$

where:

- $\delta_{2DD}(Female_{it})$ is the time policy effect allowed to vary across gender.

4.2 Underlining assumptions

The parallel trend assumption requires that the teachers not affected by the 2013 PRP implementation experience similar trends in turnover-intention as those already affected by 1999 implementation. We test the validity of this assumption by comparing the trends in probability of turnover-intention of those affected by 2013 implementation and those already affected by earlier implementation of 1999. Figure 1 shows the trend of probability of turnover-intention of the treated group and the control group before 2013 implementation. By visual observation, they look parallel except, for the year 2007 and we do not expect this to have any impact on the assumption. Figure 2 shows the trend pre and post implementation, indicating that both groups would have continued with similar trend in the absence of PRP.

Figure 1: Pre Policy Trend
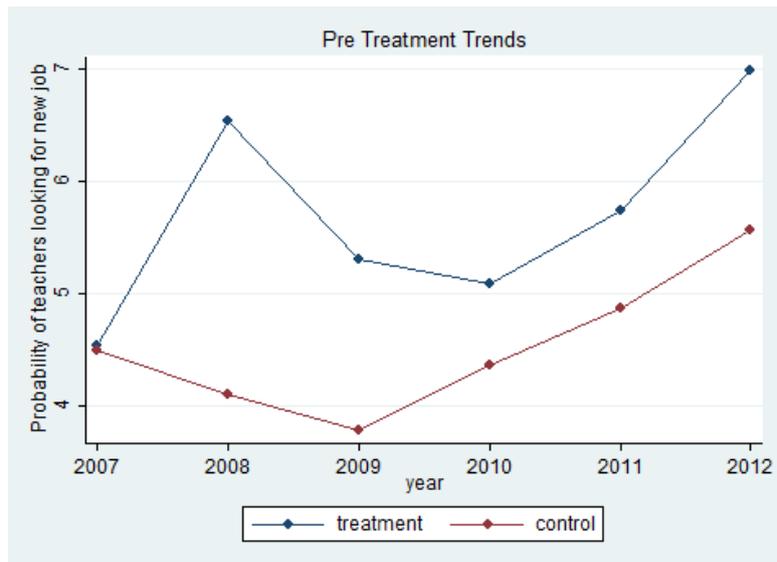
To formally test the assumption we regress the "intention to leave" variable on the set of year dummies, their interactions with the treated group indicator, and all our other covariates. We thereafter, undertook a joint test of significance on the interaction terms

between the treated group dummy and the year dummies. The coefficients of these interactions represent the gaps between the treated trend and the control trend. Testing if the gaps between the two trends is the same between 2007 and 2012 allows us to evaluate whether the two trends are parallel to each other. We fail to reject the null hypothesis that the gaps are different even at 10% level of significance (Prob > F = 0.1007). The parallel trend assumption therefore seems to be satisfied.

Figure 2: Pre and Post Policy Trend
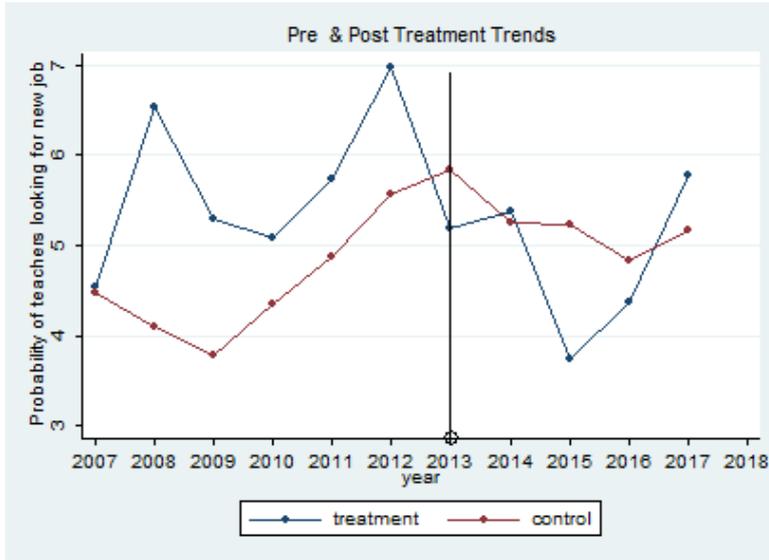
4.3 Data, sample and variable definition

This study employs a pooled cross sectional dataset of the quarterly Labour Force Survey (LFS). LFS is conducted by the Office for National Statistics (ONS). The Quarterly LFS (QLFS) which started in 1992 has a panel design with each sampled individual surveyed for five waves. About 59,000 addresses of about 138,000 respondents are interviewed every quarter. The QLFS collects information on employment and socio-economic characteristics, but also collects information on individuals desire to change or leave their current job. Our final sample covers years $2007 - 2018$, thereby providing us data for 6 years before the policy implementation (2007–2012) and 6 years after the policy implementation (2013–2018)

We select individuals working in primary and secondary schools. We further drop observations for females older than 60 and 65 for males (in-line with the UK retirement policy).

Moreover, we drop those aged 30 to 35. Since these teachers are equally likely to be in the treatment or the control group, we eliminate them from our sample to reduce the chance of wrongly categorising teachers into either groups. In addition, we drop those with missing observations. After eliminating these observations, our final sample size is made up of 43,493 observations consisting of 9,171 in treated group and 34,322 in control group.Our outcome of interest is turnover-intention, which is an indicator variable that equals 1 if individual is looking for a new job to replace present main job. Our focus is on intention to leave, as retention policy is a good step in tackling teachers' shortage before advancing to those who have actually left teaching (turnover).

Table 4.1 reports descriptive statistics of the variables used in our analysis. The treated group is made up 21% of our sample of analysis. With age range between 18 and 65, the average age of our sample is 43 years while the white race constitutes about 92% of the sample. Our data constitutes of 74% female. Our data is consistent with the head count and full-time equivalent numbers of teachers (qualified, unqualified and occasional teachers) in state funded schools (School Workforce Census 2016, 2017) where female constitutes 75.15% in 2016 and 75.55% in 2017. Similarly, the workforce census shows that 88.8% in 2016 and 86.8% in 2017 were British and Irish white.

Table 4.1: Summary statistics of the variables used in the analysis

Variables	Mean	SD	Min	Max
Age	43.05	10.91	18	65
Emp_length	5.6	1.782	1	8
Treated	0.211	0.408	0	1
White	0.916	0.278	0	1
Qualification	3.292	1.246	1	4
Zones	1.629	0.914	1	3
Female	0.74	0.439	0	1
Observations				43,493

Notes: We use the 2007 - 2018 Quarterly Labour Force Survey. We therefore dropped observations with age range 30 to 35.

Table 4.2 shows our sample composition by those below age 30 (treated group)and those above age 35 (control group), before and after the policy implementation. 55.69% of the control group and 55.11% of the treated group were in the pre implementation period, and 44.31% and 44.89% for control and treated group respectively, after policy implementation. This composition, indicates stability between our treated and control group prior and post policy implementation. Table 4.3 and table 4.4 present summary statistics with respect to gender and race respectively.

Table 4.2: Frequency distribution before
and after - treated and control groups

Time	Control Freq (Percent)	Treated Freq (Percent)
Pre	19,113 (55.69)	5,054 (55.11)
Post	15,209 (44.31)	4,117 (44.89)
Total	34322	9171

Notes: We use the 2007 - 2018 Quarterly Labour
Force Survey. We report frequency percentage in
parentheses.
Our treated group is when age is less than 30. We
therefore dropped observations with age range 30
to 35.

Table 4.3: Summary statistics by gender

Variables	Male Mean	SD	Min	Max
Age	44.14	11.18	18	65
Emp_length	5.601	1.844	1	8
Treated	0.188	0.391	0	1
White	0.914	0.28	0	1
Qualification	3.227	1.292	1	4
Zones	1.625	0.91	1	3
Observations				11,313

	Female Mean	SD	Min	Max
Age	42.67	10.79	18	60
Emp_length	5.6	1.759	1	8
Treated	0.219	0.413	0	1
White	0.916	0.277	0	1
Qualification	3.314	1.229	1	4
Zones	1.631	0.916	1	3
Observations				32,180

Notes: We use the 2007 - 2018 Quarterly Labour Force Survey. We therefore
dropped observations with age range 30 to 35.

Table 4.4: Summary statistics by race (white)

Variables	Non-white Mean	SD	Min	Max
Age	40.69	10.68	18	65
Emp_length	5.164	1.764	1	8
Treated	0.261	0.439	0	1
Qualification	3.594	1.014	1	4
Zones	1.541	0.849	1	3
Female	0.735	0.441	0	1
Observations				3,661

	White Mean	SD	Min	Max
Age	43.27	10.91	18	65
Emp_length	5.64	1.778	1	8
Treated	0.206	0.405	0	1
Qualification	3.264	1.262	1	4
Zones	1.638	0.92	1	3
Female	0.74	0.438	0	1
Observations				39,832

Notes: We use the 2007 - 2018 Quarterly Labour Force Survey. We therefore dropped
observations with age range 30 to 35.

5 Results

5.1 Estimation Results

Our Linear Probability Model result (as presented in table 5.1) shows that the probability of looking for a new job to replace current teaching job is reduced by 1.61 percentage points for those likely to be affected by the policy and it was statistically significant at 1%. Controlling for individual characteristics (gender, age and race), our result shows that, the intention to leave teaching fell by about 1.5 percentage points for those likely to had been affected by the 2013 PRP. Controlling also for job characteristics (how long an individual had been working with the employer and the region where the job is located) shows that the policy impact has led to 1.3 percentage points reduction in intention to leave teaching. Similarly, controlling for individual qualification, indicates that the probability of leaving declined significantly at 5% level of significance. When we relate our result with the actual turnover of qualified teachers of 9.7% average between 2013 (when full implementation of PRP commenced) and 2017 (DfE, 2018), a reduction of 1.16 percentage points will translate into 13% decline.

We do not find any evidence that holding a degree or a teaching certificate significantly contributed to increasing the probability of looking for a new job. Our result is in line with the DfE (2018) and the rate of turnover between 2014 and 2017 for qualified teachers has remained constant. We, however, find that those with higher diploma are about 4 percentage points more likely to look for a new job than those with degree and teachers qualification and it is significant at 5%.

The result of checking for the policy effect across sub groups (as presented in table 5.2) shows that, the policy effect is not significantly different for male and female, although affected females are about 0.5 percentage points more likely to want to leave teaching than their affected male counterparts. However, the impact is different by race when we compared white with other races. Probability of looking for a new job for white individuals affected by the policy is 3.6 percentage points higher than non-white affected by the policy. Although, the proportion of non-white in our sample is 8.42% and it is possible that the distribution could have affected the conclusion with respect to race because a slight change in 91.58% (which is white), can certainly posit significant effect relative to non-white.

	Baseline	After controlling for individual characteristics	After controlling for employment characteristics	After controlling for educational qualification
Diff-in-Diff	-0.0158***	-0.0147***	-0.0129**	-0.0130**
	(0.00259)	(0.00505)	(0.0134)	(0.0134)
Individual Characteristics				
Female		-0.0146***	-0.0135***	-0.0134***
		(0.0000)	(0.0000)	(0.0000)
Age		-0.00189***	-0.00114***	-0.00113***
		(0.0000)	(0.0000)	(0.0000)
White		-0.00536	-0.00172	-0.00174
		(0.2)	(0.679)	(0.676)
Length with employer:				
3 months < 6			-0.0281***	-0.0280***
			(0.00964)	(0.00968)
6 months < 12			-0.0243**	-0.0244**
			(0.0144)	(0.014)
Zones:				
Outer Londonn & South England			-0.0180***	-0.0178***
			(0.00289)	(0.00325)
Rest of England			-0.00396	-0.00395
			(0.107)	(0.107)
Qualifications:				
Higher Diploma				0.0408**
				(0.0162)
Teachers Qualification				-0.00682
				(0.145)
Other Qualification				-0.00631
				(0.263)
Observations	43,493	43,493	43,493	43,493
R-squared	0.001	0.005	0.013	0.013

Notes: *** Significant at 1%; ** significant at 5%; * significant at 10%. We report p-values robust to heteroskedasticity in parentheses. The estimated coefficients of the full set of "length with employer" are not reported for the sake of brevity and are available from the authors upon request.

We use the 2007 - 2018 Quarterly Labour Force Survey and estimated Linear Probability Regression. Dependent variable:1 if individual is looking for new job to replace present job. The reference group for zones is Inner London and reference group for qualification is Degree. We controlled for time fixed effect by including year dummies. Our treated group is when age is less than 30. We therefore dropped observations with age range 30 to 35.

With respect to location, the probability of looking for a new job, for affected teachers in London fell by about 1 percentage point and only significant at 10%. The policy effect is however, significantly different for those in Inner London and those we grouped as outside London. Probability of looking for a new job is 4 percentage points lower than for those in Inner London and significant at 1% level. We do not however find any significant policy impact difference between inner London and the rest of England (excluding outer London and South of England), even though teachers in the rest of England is about 1 percentage point less likely to look for a new job relative to inner London after the policy implementation.

It could therefore be concluded that, the policy has been able to stem adverse intention to quit teaching, particularly for the newly affected teachers. This is in-line with one of the objectives of the policy. We are however unable to answer if the policy has attracted more young teachers to the profession (because of the policy implementation), but our result showed that turnover intention before and after the policy are statistically different

for the young teachers.

Table 5.2: Policy Impact Difference

	Gender	Race	Zones
DID	-0.0340***	-0.0636***	-0.0279***
	(0.0000)	(0.0000)	(0.0000)
Female	-0.0137***		
	(0.0000)		
Policy effect interacted with gender (Female)	0.00471		
	(0.587)		
White		-0.0071	
		(0.122)	
Policy effect interacted with race (White)		0.0368***	
		(0.0000)	
Outer London			-0.0157**
			(0.0164)
Rest of England			-0.00319
			(0.202)
Policy effect interacted with zones (Outer London)			-0.0395***
			(0.0000)
Policy effect interacted with zones (Rest of England)			-0.0113
			(0.247)

	Policy Difference Test
	Prob> F
Treatment effect male = Treatment effect female	0.195
Treatment effect non-white = Treatment effect white	0.0000***
Treatment effect London = Treatment effect outer London	0.0080***
Treatment effect London = Treatment effect rest of England	0.9223
No of observations	43,493

Notes: *** Significant at 1%; ** significant at 5%; * significant at 10%. We report p-values robust to heteroskedasticity in parenthesis. We use the 2007 - 2018 Quarterly Labour Force Survey and estimated Linear Probability Regression. Dependent variable:1 if individual is looking for new job to replace present job. We controlled for individual characteristics (gender, age, and race), controlled for employment characteristics (length with employer and zone of employment location) and controlled for educational qualification of individual teacher (Degrees, Higher Diploma, Teachers' Qualification and Others).

5.2 Robustness Checks

To check the robustness of our findings, we conduct various sensitivity analyses. We first undertake a placebo test by reducing the age for categorising individuals into the treated group from below 30 years to below 25 years. We then excluded those between age 25 and 30 and utilise those above 30 years as the control group. We expect the result to be similar with our original conclusion, since all those still within the newly created placebo treated group are all members of our original treated group. Results in table 5.3 show that the conclusion is similar. On the baseline, and when we controlled for individual characteristics, we found the policy to have resulted in reducing the intention to leave the job by 1.85 percentage points and significant at 5% level. When we controlled for employment characteristics and educational qualification, the effect reduced to 1.52 percentage points.

Secondly, we increased the age for categorising individuals into treatment from below 30 years of age to below 40 years of age. We expect the effect of the policy to diminish or

flip, since part of people now included in the newly-created placebo treated group are actually within the control group. Our result in table 5.3 confirms that the policy impact is reduced to 0.73 percentage point and significant only at 10% (on the baseline). When we controlled for employment and qualifications, the policy effect fell to 0.33 percentage point and is no longer significant.

We further undertake a placebo test by varying the implementation period to year 2009. Our apriori expectation is that, the policy impact will be different from results obtained from the genuine categorisation done with 2013, since category of newly affected teachers by the 2013 policy (young teachers) were not affected by the 2009 implementation . Our results, similar to what is presented in table 5.1 is presented in table 5.5, and we find no evidence at all that, the hypothetical 2009 policy has any significant impact on teachers' turnover intention at baseline, even at 16% level of significance. Same conclusion that, no significant impact on teachers' intention to leave, is reached when we controlled for individual characteristics, employment characteristics and educational qualifications. This result further clarifies the distinction between partial implementation of PRP in 2009 with the full implementation of 2013 which, is our focus in this study.

Table 5.3: Placebo by varying cutoff age of treated group

	Treated: Age less than 25		Treated: Age less than 40	
	Coefficient	P-value	Coefficient	P-value
Policy effect on the baseline	-0.0185**	0.0250	-0.00733*	0.0672
Policy effect after controlling for individual characteristics	-0.0185**	0.0258	-0.00692*	0.0835
Policy effect after controlling for employment characteristics	-0.0152*	0.0671	-0.00334	0.4030
Policy effect after controlling for educational qualification	-0.0152*	0.0661	-0.00328	0.4190
No of observations	45,908		44,626	

Notes: *** Significant at 1%; ** significant at 5%; * significant at 10%. We report *p*-values robust to heteroskedasticity.
We use the 2007 - 2018 Quarterly Labour Force Survey and estimated Linear Probability Regression. Dependent variable:1 if individual is looking for new job to replace present job. We controlled for individual characteristics (gender, age, and race), controlled for employment characteristics (length with employer and zone of employment location) and controlled for educational qualification of individual teacher (Degrees, Higher Diploma, Teachers' Qualification and Others). When treated age is reduced to 25, we dropped observations with age range 26 to 30 and when treated age is increased to 40, we dropped observations with age range 40 to 45.

As part of our sensitivity analysis also, we checked treatment heterogeneity across time and our results show a good level of consistency across time. We interact the treated group across years (2007 - 2018). Results presented in table 5.6 shows that, the treated group when interacted with year dummies (using year 2018 as reference year), all returned negative coefficients between 2.19 percentage points and 5.78 percentage points and are all significant at 1% level except for 2008 interaction that is significant at 5% level.

Table 5.4: Placebo varying implementation year to 2009 - period of partial PRP introduction

	Baseline	After controlling for individual characteristics	After controlling for employment characteristics	After controlling for educational qualification
Diff-in-Diff	-0.00952	-0.00743	-0.00521	-0.00503
	(0.164)	(0.277)	(0.444)	(0.464)
Individual Characteristics				
Female		-0.0145***	-0.0135***	-0.0133***
		(0.0000)	(0.0000)	(0.0000)
Age		-0.00190***	-0.00114***	-0.00114***
		(0.0000)	(0.0000)	(0.0000)
White		-0.00524	-0.00159	-0.00160
		(0.210)	(0.703)	(0.700)
Length with employer:				
3 months < 6			-0.0281***	-0.0281***
			(0.0095)	(0.0096)
6 months < 12			-0.0243**	-0.0244**
			(0.0142)	(0.0137)
Zones:				
Outer Londonn & South England			-0.0178***	-0.0175***
			(0.0032)	(0.0036)
Rest of England			-0.00402	-0.00401
			(0.102)	(0.102)
Qualifications:				
Higher Diploma				0.0402**
				(0.0178)
Teachers Qualificationl				-0.00726
				(0.124)
Other Qualification				-0.00643
				(0.254)
Observations	43,493	43,493	43,493	43,493
R-squared	0.001	0.005	0.013	0.013

Notes: *** Significant at 1%; ** significant at 5%; * significant at 10%. We report p-values robust to heteroskedasticity in parentheses. The estimated coefficients of the full set of length with employer are not reported for the sake of brevity and are available from the authors upon request.

We use the 2007 - 2018 Quarterly Labour Force Survey and estimated Linear Probability Regression. Dependent variable:1 if individual is looking for new job to replace present job. The reference group for zones is Inner London and reference group for qualification is Degree. We controlled for time fixed effect by including year dummies.

Our treated group is when age is less than 30. We therefore dropped observations with age range 30 to 35.

6 Conclusions

This paper studies the impact of the 2013 performance related pay policy (PRP), which ties teachers' pay in England to their performance, on their intention to leave the job. We use difference-in-differences estimators since the 2013 policy affected only those below upper pay scale, while other teachers have been affected by the earlier implementation. We find that the probability of leaving teaching significantly declined by 1.16 percentage points, which amounts to approximately 13 % decline in the affected teachers' intention to leave. We found no heterogeneous policy impact across gender, while policy effect for those in inner and outer London is found to be significantly different. We conclude that, appropriate utilisation of PRP can help in lowering teachers shortage problem, as young teachers see possibility of faster career progression, as an incentive to remain in teaching.

Table 5.5: Treated and Time Effect

Treated & Time	Coefficients
Treated interacted with year 2007	-0.0480***
	(0.0000)
Treated interacted with year 2008	-0.0219**
	(0.0288)
Treated interacted with year 2009	-0.0288***
	(0.00127)
Treated interacted with year 2010	-0.0376***
	(0.0000)
Treated interacted with year 2011	-0.0350***
	(0.0003)
Treated interacted with year 2012	-0.0283***
	(0.0071)
Treated interacted with year 2013	-0.0489***
	(0.0000)
Treated interacted with year 2014	-0.0420***
	(0.0000)
Treated interacted with year 2015	-0.0578***
	(0.0000)
Treated interacted with year 2016	-0.0462***
	(0.0000)
Treated interacted with year 2017	-0.0358***
	(0.0005)
Observations	43,493
R-squared	0.013

Notes: *** Significant at 1%; ** significant at 5%; * significant at 10%. We report p-values robust to heteroskedasticity in parentheses.

We use the 2007 - 2018 Quarterly Labour Force Survey and estimated Linear Probability Regression. Dependent variable:1 if individual is looking for new job to replace present job. We controlled for individual characteristics (gender, age, and race), controlled for employment characteristics (length with employer and zone of employment location) and controlled for educational qualification of individual teacher (Degrees, Higher Diploma, Teachers' Qualification and Others).

References

ADAMS, G. J. (1996): "Using a Cox regression model to examine voluntary teacher turnover," The Journal of experimental education, 64(3), 267–285.

ATKINSON, A., S. BURGESS, B. CROXSON, P. GREGG, C. PROPPER, H. SLATER, AND D. WILSON (2009): "Evaluating the impact of performance-related pay for teachers in England," Labour Economics, 16(3), 251–261.

BELFIELD, C. R., AND J. S. HEYWOOD (2008): "Performance pay for teachers: Determinants and consequences," Economics of Education Review, 27(3), 243–252.

BRITTON, J., AND C. PROPPER (2016): "Teacher pay and school productivity: Exploiting wage regulation," Journal of Public Economics, 133(C), 75–89.

BROWN, D., AND M. ARMSTRONG (1999): Paying for contribution: Real performance-related pay strategies. Kogan Page Publishers.

CHAMBERLIN, R., T. WRAGG, G. HAYNES, AND C. WRAGG (2002): "Performance-Related Pay and the Teaching Profession: A Review of the Literature.," Research Papers in Education: Policy & Practice, 17(1), 31–49.

CHAPMAN, D. W. (1984): "Teacher retention: The test of a model," American Educational Research Journal, 21(3), 645–658.

CUTLER, T., AND B. WAINE (2000): "Mutual Benefits or Managerial Control? The Role of Appraisal in Performance Related Pay for Teachers," British Journal of Educational Studies, 48(2), 170–182.

DEN BROK, P., T. WUBBELS, AND J. VAN TARTWIJK (2017): "Exploring beginning teachers attrition in the Netherlands," Teachers and Teaching, 23(8), 881–895.

DFE (2013): "New advice to help schools set performance-related pay," .

FIGLIO, D. N., AND L. W. KENNY (2007a): "INDIVIDUAL TEACHER INCENTIVES AND STUDENT PERFORMANCE," Journal of Public Economics, 91(56), 901 – 914.

FIGLIO, D. N., AND L. W. KENNY (2007b): "Individual teacher incentives and student performance," Journal of Public Economics, 91(5-6), 901–914.

FRYER, R. G. (2011): "TEACHER INCENTIVES AND STUDENT ACHIEVEMENT: EVIDENCE FROM NEW YORK CITY PUBLIC SCHOOLS," Working Paper 16850, National Bureau of Economic Research.

FRYER, R. G. (2013): "Teacher incentives and student achievement: Evidence from New York City public schools," Journal of Labor Economics, 31(2), 373–407.

GIUS, M. (2013): "The effects of merit pay on teacher job satisfaction," Applied Economics, 45(31), 4443–4451.

GLEWWE, P., N. ILIAS, AND M. KREMER (2010): "Teacher incentives," American Economic Journal: Applied Economics, 2(3), 205–27.

GUARINO, C. M., L. SANTIBANEZ, AND G. A. DALEY (2006): "Teacher recruitment and retention: A review of the recent empirical literature," Review of educational research, 76(2), 173–208.

HANUSHEK, E. A., S. G. RIVKIN, R. ROTHSTEIN, AND M. PODGURSKY (2004): "How to improve the supply of high-quality teachers," Brookings papers on education policy, (7), 7–44.

INGERSOLL, R. M. (2003): "The teacher shortage: Myth or reality?," Educational Horizons, 81(3), 146–152.

LAVY, V. (2002): "Evaluating the effect of teachers group performance incentives on pupil achievement," Journal of political Economy, 110(6), 1286–1317.

LINDQVIST, P., AND U. K. NORDÄNGER (2016): "Already elsewhere–A study of (skilled) teachers' choice to leave teaching," Teaching and Teacher Education, 54, 88–97.

MAHONY, P., I. MENTER, AND I. HEXTALL (2004): "The emotional impact of performance-related pay on teachers in England," British Educational Research Journal, 30(3).

MURNANE, R., J. SINGER, AND J. WILLETT (1989): "The influences of salaries and" opportunity costs" on teachers' career choices: Evidence from North Carolina," Harvard Educational Review, 59(3), 325–347.

PODGURSKY, M. J., AND M. G. SPRINGER (2007): "Teacher performance pay: A review," Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management, 26(4), 909–950.

POSEY, V. (2017): "Solving the Teacher Shortage: It's as Easy as 1-2-3!," Delta Kappa Gamma Bulletin, 84(2), 30–64.

RICHARDSON, R. (1999): Performance Related Pay in Schools: An Assessment of the Green Papers: A Report Prepared for the National Union of Teachers. London School of Economics and Political Science.

SPRINGER, M. G., D. BALLOU, L. HAMILTON, V.-N. LE, J. LOCKWOOD, D. F. MC-CAFFREY, M. PEPPER, AND B. M. STECHER (2011): "TEACHER PAY FOR PERFORMANCE: EXPERIMENTAL EVIDENCE FROM THE PROJECT ON INCENTIVES IN TEACHING (POINT).," Society for Research on Educational Effectiveness, ERIC.

STOREY, A. (2000): "A Leap of Faith? Performance Pay for Teachers.," Journal of Education Policy, 15(5), 509–23.

SUTCHER, L., L. DARLING-HAMMOND, AND D. CARVER-THOMAS (2016): "A coming crisis in teaching? Teacher supply, demand, and shortages in the US," Learning Policy Institute. Retrieved from https://learningpolicyinstitute.org/product/coming-crisis-teaching.

TOMLINSON, H. (2000): "Proposals for Performance Related Pay for Teachers in English Schools.," School Leadership & Management, 20(3), 281–98.

WAYNE, A. J., AND P. YOUNGS (2003): "Teacher characteristics and student achievement gains: A review," Review of Educational research, 73(1), 89–122.

WRAGG, T., G. HAYNES, R. CHAMBERLIN, AND C. WRAGG (2003): "Performance-related pay: the views and experiences of 1,000 primary and secondary head teachers," Research Papers in Education, 18(1), 3–23.