

Durham University Evidence Centre for Education
<https://www.dur.ac.uk/dece>



Attracting and retaining teachers in hard to staff areas: What does the evidence say?

Beng Huat See, Stephen Gorard, Rebecca Morris and Nada El Soufi

b.h.see@durham.ac.uk

s.a.c.gorard@durham.ac.uk

rebecca.e.morris@warwick.ac.uk

Attracting and retaining teachers in hard to staff areas: What does the evidence say?

Beng Huat See, Stephen Gorard, Rebecca Morris and Nada El Soufi

Abstract

This describes a systematic review of international research evidence to identify the most promising approaches to attracting and retaining teachers in hard-to-staff areas and schools. The aim was to identify approaches that had been tested and shown to have benefits. For this reason, we included only empirical studies that employed a causal or suitable comparative design, and had robust measurements of recruitment and retention outcomes. Studies were quality assessed to take account of threats to trustworthiness which may bias the results. A search of 13 electronic databases and Google/Google scholar identified 20 distinct research reports that met the inclusion criteria. The only approach that seems to work in attracting teachers to challenging schools or areas was the offer of financial incentives. However, such financial incentives are not enough to keep teachers in these schools once the financial payments are removed. A caring, supportive and congenial working environment may be helpful in keeping teachers when they are under pressure. There is little evidence that approaches such as mentoring, support, or teacher development are effective. This is largely because much of the research on these approaches is so weak. More robust research capable of addressing causal questions is therefore urgently required to determine their impact in attracting and retaining good teachers in areas where they are most needed. But longer, the solution is to introduce policies improving hard to staff areas and schools, so that the problem of staffing does not arise.

Background

Education systems worldwide attempt to provide good quality education for their citizens, and this requires a supply of high-quality teachers. Supply has reportedly become more difficult in recent times because of challenges in recruiting and retaining teachers, and teacher shortages are a global concern. Many European countries and the US have reported serious shortages of teachers (European Commission, 2015). Widespread media reports of teacher shortages both in England (Sky News, 2017; Boffey & Helm 2015; Hazell 2018a) and the US (Williams, 2018; Caitlin 2017; Passy 2018; Strauss 2016) have dominated newspaper headlines in the last few years.

In England and the US teacher shortages are predicted to get worse as pupil population is rising and more teachers in the profession are leaving before retirement. Shortages are particularly acute in some subjects and regions. But for some schools and regions this overall shortage of supply is more serious because they are already facing great difficulties in attracting and retaining teachers by virtue of their undesirable location and student intake. Schools in isolated and economically deprived areas, rural schools, inner city schools, schools with a high proportion of disadvantaged and low performing pupils and schools with challenging intakes often struggle to recruit and retain teachers.

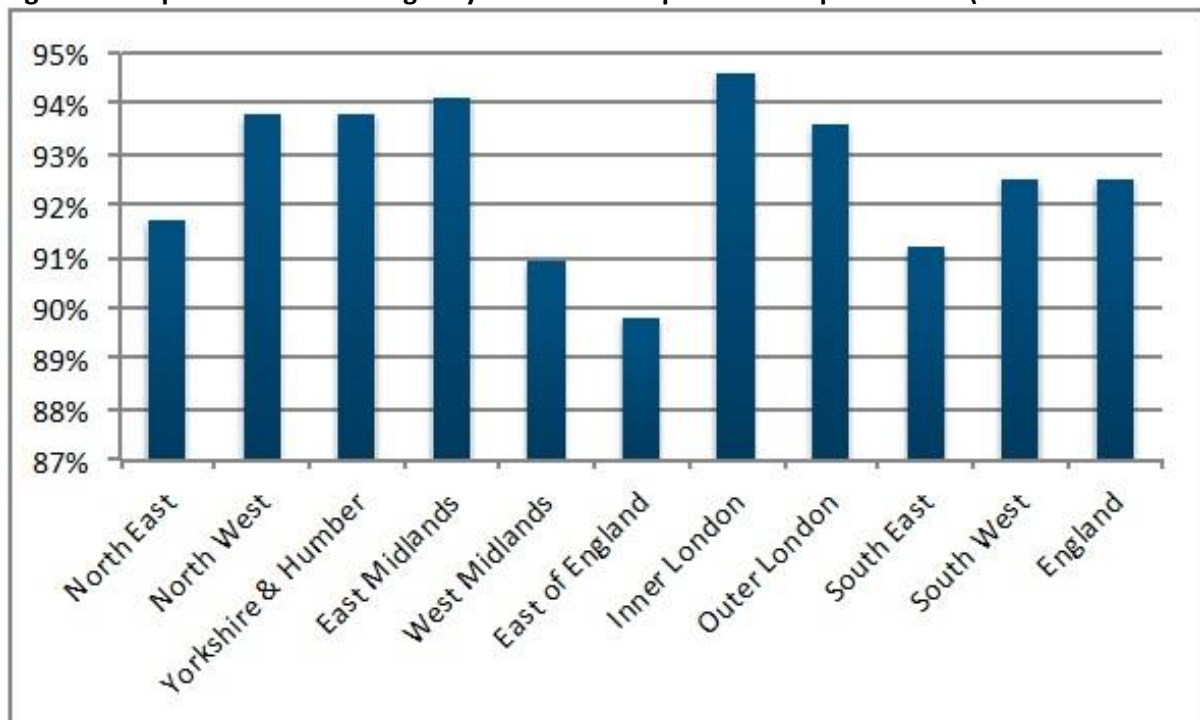
Attracting and retaining suitably qualified teachers in *some* subjects and geographical areas is a challenge common to the school staffing policies of many developed countries. More than half of the countries in Europe and almost all school districts in the US report problems, and shortages and oversupply can coexist because of the uneven distribution of teachers across phases, subjects and regions. In Germany and England, for example, there is an oversupply in some subjects and a shortage

in others (European Commission/EACEA/Eurydice 2018). In Greece, there is a shortage on some small remote islands while there is a general oversupply of teachers in the rest of the country. Teacher shortages related to the remoteness of some regions are mentioned in half of the countries that participated in the European Commission survey. In other countries, it was the high cost of living and high proportion of disadvantaged pupils in some large urban cities (such as Brussels and London) that reportedly made it difficult to attract and retain teachers.

Large urban schools and small rural schools can find it difficult to recruit teachers for different reasons. Small rural schools with their economically less efficient small class sizes, geographical isolation with lack of social amenities and poor transport system, and relatively lower salaries have made it difficult for them to recruit and retain teachers. Large urban schools, on the other hand, may find it difficult because of the higher cost of living, or higher proportions of socio-economically disadvantaged pupils, and transient populations. In England, less affluent areas have had greater difficulties in attracting qualified teachers than many other parts of the country.

There are geographical cold spots in England where schools are rated as least likely to have teachers in shortage subjects with a relevant degree. In coastal rural areas, which can be highly deprived, 7% of secondary teachers are unqualified, compared with 4.6% in more affluent inland rural areas (Social Mobility Commission 2017). Regionally, the North East, West Midlands and East of England are less likely to have teachers with a relevant degree teaching shortage subjects compared to London (Figure 1). For example, only 17% of physics teachers in poorer schools outside London have a relevant degree, compared with 52% in affluent areas in the rest of the country (Sibieta 2018).

Figure 1 - Proportion of hours taught by a teacher with post A-level qualification (maths and science)



In the House of Commons (2017) 5th Report on teacher recruitment and retention, the government acknowledged that there were wide regional variations in teacher supply. While there have been plans to encourage more teachers to work in areas most in need, these have not been very successful. The pilot for a National Teaching Service, for example, which was set up to get teachers to teach in areas most struggling to recruit, had to be abandoned after managing to recruit only 54 of 1,500 intended teachers.

To tackle these challenges, many education systems have offered incentives and implemented a range of programmes to attract potential teachers to, and retain existing teachers in, difficult-to-staff regions and for some high demand subjects. For example, the Department for Education in England recently announced that early career maths and science teachers in Yorkshire in the North East of England will receive £2,000 to test whether incentivising such teachers will keep them in the profession. Teachers in other Opportunity Areas will also benefit from an injection of £72 million (Department for Education [DfE] 2019). This approach was based on evidence from the Gatsby Foundation (Sims 2019) and Education Policy Institute (Sibieta 2018).

Some of these programmes have been evaluated and tested but there is, so far, no synthesis of the research findings, so the evidence of their effectiveness is still unclear. Many of these incentives and programmes are expensive and it would be a waste of taxpayers' money and the country's resources to continue using them if there is no evidence that they work. There is also an opportunity cost as the money used for these incentives could be otherwise channelled to more effective programmes. If they show promise it is important to know how they can best be implemented, and the extent to which they can be deployed in other countries facing similar challenges. It is therefore crucial that these strategies are robustly evaluated and tested before more money is spent on them worldwide.

As far as we know, there has been no large-scale comprehensive single-study review of the evidence on teacher recruitment and retention policies with a view to addressing the recurring problems in teaching supply. Previous reviews have taken a narrower focus and/or have not taken into account the quality and design of research in each study (e.g. Wheeler and Glennie, 2007). This paper presents the findings of a systematic review of international empirical research to identify the most promising approaches in attracting and retaining teachers in hard-to-staff schools and areas.

Method

Search strategy

Our review began with a broader search for studies that address teacher recruitment and retention issues in general, and from these we identified and analysed the studies relevant to hard-to-staff areas separately for this paper.

The search included substantive terms about teacher supply (e.g. teacher recruitment or retention) and policy initiatives, incentives, approaches and schemes (and their synonyms). As the purpose of this review was to identify approaches that show evidence of impact only studies that employ a causal design were included. Therefore, the key words also included any causal term (or a synonym) or any research design that would be appropriate for testing a causal model, such as experiments, quasi-experiments, regression discontinuity and difference-in-difference. A scoping review was first conducted to test out the sensitivity of the search terms on well-known sociological, educational and psychological databases to ensure that the search terms picked up relevant pieces of literature and also known studies on this topic. Following this, a very general and inclusive statement of search terms was generated for each database. These were adjusted to suit the idiosyncrasies of each. For different databases we had to modify the syntax but used similar key words.

Because the scoping review and previous reviews of literature suggested that there were few robust experimental evaluations of policy initiatives or approaches that aim to improve recruitment and retention of classroom teachers, we included any empirical studies with at least some type of

comparative design, many of which will subsequently have low ratings for trustworthiness in terms of causal claims.

The search terms were applied to the main educational, psychological and sociological electronic databases. These included:

- Education Resources Information Clearinghouse
- JSTOR
- The Scholarly Journal Archive
- Social Sciences and Education Full Text
- Web of Science
- Sage
- Science Direct
- Proquest Dissertations and Theses (<http://library.dur.ac.uk/record=b2044198~S1>)
- British Education Index
- ERIC (Educational Resources Information Center)
- IBSS (International Bibliography of the Social Sciences)
- Ingenta Journals (full text of a large number of journals)
- EBSCOhost (which covers the following databases: PsychINFO, BEI, PsycARTICLES, etc, ProQuest, IBSS
- plus Google and Google Scholar.

The search was limited to studies published in the English language. We intentionally did not set any date limits, to keep the search open. To avoid publication bias, the search included any material published or unpublished that mentions both substantive and causal terms. A total of 6,708 research reports were identified and exported to EndNote for screening. This review was completed at the end of 2018 and therefore would not include studies that come after 2018.

There is no one definition of teacher retention. While most studies considered retention as involving teachers staying within their current school, in others retention referred to teachers staying within the school district, the state, state-funded schools, or even within teaching as a profession. The same mix appears in claims about teacher wastage in England (See and Gorard 2019). In this review, we included any studies that look at retention of teachers regardless of how this is defined.

Screening

Each identified study was screened to remove duplicates, and for relevance on the basis of title and abstract. Only studies that related specifically to recruitment and retention for hard-to-staff areas were retained. This process removed 6,161 studies, leaving 547 which were read in full.

We screened the studies using the inclusion and exclusion criteria. Studies were included if they were:

- Empirical research
- About activities aimed at attracting people into teaching or about retaining teachers in teaching
- Specifically about recruitment and retention of classroom teachers
- About incentives/initiatives/policies or schemes on teacher recruitment and retention
- About mainstream teachers in state-funded /government schools
- Studies that had measurable outcomes (either retention or recruitment)
- Studies that relate to mainstream education

- Recruitment and retention of traditional shortage subject teachers(e.g. maths, science and design and technology)

Studies were excluded if they were:

- Not primary research
- Not published or reported in English
- Not actually a report of research at all
- Simply descriptions of programmes or initiatives with no evaluation
- Not about strategies or approaches to improve recruitment or retention of teachers (e.g. observational or correlational studies of factors influencing recruitment and retention)
- Studies that have no clear evaluation of outcomes
- Studies with non-tangible or measurable outcomes (e.g. teachers' attitude or beliefs or perceptions)
- Ethnographic, opinion pieces, guidance briefs or manuals on how to attract and retain teachers
- Outcome is not teacher recruitment or retention
- Not about recruitment and retention of teachers
- If it is specifically about school leaders, school administrators or teaching assistants
- Outcome is about student achievement (e.g. Cowan & Goldhaber 2016)
- Not about mainstream teachers, e.g. special education teachers or ethnic minority teachers
- Not relevant to the context of English speaking developed countries (e.g. Duflo et al. 2007)
- Not relevant to the research questions
- Anecdotal accounts from schools about successful strategies

A large number involving surveys or comparisons before and after with no comparison groups were eventually excluded because they do not add to the evidence base. There were also many studies about recruitment and retention initiatives and what some schools or school districts have adopted, but with no evaluation of the outcomes. These were excluded.

At this stage the full reports were skim-read by one researcher. Any studies now thought not to meet the inclusion criteria were then reviewed by the other three members of the research team for consensus. Further, in order to establish inter-rater reliability, all four members of the team independently reviewed 10 randomly selected reports to decide if they agreed on their inclusion or exclusion.

Only 52 studies were retained, deemed to be relevant to the research question and at least partly concerned with staffing in hard-to-staff schools and subjects.

Data extraction

Information from these papers was summarised, including details on research design, cases used, allocation to groups, outcome measures, missing data, analysis and the results. A further 17 studies were then excluded when it became clearer that they were not evaluations but narrative discussions of previous research or suggestions of strategies. Of the remainder, 13 studies were excluded because they merely involved asking respondents such as headteachers which strategies they thought worked or were important to them. Three reports were of different approaches to evaluating the same intervention by the same set of authors. These were treated as being one report here.

Quality assessment

In total, 20 reports were assembled that were deemed to be both relevant and research-related. These were passed through a quality assessment ‘sieve’, a tool used to judge the trustworthiness of each research report. If public investment is to be made on recruitment and retention programmes or incentives, it is crucial that the most robust evidence is given the most weight. For this reason we synthesised the evidence by first assessing the trustworthiness of the findings in each report based on five criteria (Table 1). These were a research design appropriate to address a causal question (e.g. whether it is an RCT with random assignment of cases, if there is a fair comparator group), scale of the study (smallest cell size), level of attrition, quality of outcome measurement (e.g. self-report or administrative data) and any other threats to validity. Each study is then assigned a score between 1* (the minimum standard to be given any weight, including some kind of comparison) and 4* (Gorard et al 2017). Studies that are not empirical would be rated 0 and excluded; these are not discussed here. Four-star studies are the most secure, meaning that the evidence is most reliable or trustworthy.

Table 1- Quality assessment ‘sieve’

Design	Scale	Dropout	Data quality	Threats	Rating
Strong design for RQ	Large number of cases (per comparison group)	Minimal attrition, no evidence of impact on findings	Standardised, pre-specified, independent	No evidence of diffusion, demand, or other threat	4★
Good design for RQ	Medium number of cases (per comparison group)	Some attrition (or initial imbalance)	Pre-specified, not standardised or not independent	Little evidence of diffusion, demand or other threat	3★
Weak design for RQ	Small number of cases (per comparison group)	Moderate attrition (or initial imbalance)	Not pre-specified but valid in context	Evidence of diffusion, demand or other threat	2★
Very weak design for RQ	Very small number of cases (per comparison group)	High attrition (or initial imbalance)	Issues of validity or appropriateness	Strong indication of diffusion, demand or other threat	1★
No consideration of design	A trivial scale of study, or N unclear	Attrition huge or not reported	Poor reliability, too many outcomes, weak measures	No consideration of threats to validity	0

We ignored the source of publication, reputation of the researchers/authors and their reported outcomes and conclusions in our assessment. It is very common to find studies reporting a positive impact despite having no relevant data and studies that produced conclusions unwarranted by the data. The study outcomes are classified as relevant to either recruitment or retention, or both. Approaches with the most highly rated studies showing positive effects are considered the most promising.

Synthesis

The research reports were classified according to whether they were about recruitment, retention or both. These were then further sorted according to the types of incentives or initiatives. A broad classification of incentives/initiatives was created. These include financial incentives (e.g. signing bonuses, district benefits, wage up lifts, scholarships and loans), other not directly financial incentives (e.g. housing benefits, retirements, pension, health care and child care benefits) and other non-financial incentives (e.g. alternative routes into teaching, staff development, mentoring & induction and workload reduction) or a combination.

It has to be made clear that approaches with no evidence of impact does not mean that they are not effective, but rather that the existing evidence is such that it is unable to establish impact.

The results

There are a total of 20 studies of interventions relevant to staffing in difficult areas, which had 26 individual outcomes relevant to either or both the recruitment and retention of teachers (Table 2). Most involved some kind of financial incentives to teach in hard-to-staff schools and, on balance, such approaches appear to work. Many are from the US, while very few are from England. None are of the highest quality.

Table 2 – Quality rating of all included studies on 26 outcomes

Quality of study	Positive outcome	Unclear outcome	Negative or neutral outcome
4*	-	-	-
3*	2	-	1
2*	7	-	2
1*	8	4	2

Note: Some studies evaluated both recruitment and retention and reported different results for each outcome. For this reason, the table presents the ratings for the evaluation of each outcome.

Improving recruitment

Overall, the results of recruitment studies are mixed. Of the nine study outcomes that met at least our minimum quality for a causal claim, five reported positive outcomes, but only three were of a higher quality (i.e. 2* and above). The highest quality study (3*) showed positive outcomes for recruitment (Table 3), but not for retention. Hough and Loeb (2013) assessed the effect of awarding higher salaries/bonuses to teachers teaching shortage subjects, in schools with a high proportion of poor and ethnic minority students, in the San Francisco Unified School District. Teachers were given a rise of \$500 to \$6,300 (depending on the salary scale), and a \$2,000 bonus for teaching in hard-to-staff schools, a retention bonus of \$2,500 if they stayed on after the fourth year, and \$3,000 after the eighth year. Using a difference-in-difference approach, the authors compared the recruitment and retention of 1,611 applicants with teachers in different school districts before and after the introduction of the policy (a comparison made more difficult because of the economic downturn in 2008). There was an increase in the number (from 49% to 54%) and proportion (27% to 37%) of shortage subject teachers in hard-to-staff areas. However, there was no difference in the retention rates of targeted and non-targeted teachers. Over 90% of teachers stayed on in the district and over 85% stayed in their school, in both groups. The authors suggested that a policy aimed at retaining teachers in a competitive labour market when the economy is doing well may not then be necessary when unemployment is higher.

Table 3 – Quality rating of studies on recruitment

Quality of study	Positive outcome	Unclear outcome	Negative or neutral outcome
4*	-	-	-
3*	Hough and Loeb 2013	-	-
2*	Steele et al. 2010 Glazerman et al. 2013	-	-
1*	Clewell and Vegas 2011 Waters-Weller 2009	Gordon and Vegas 2005 Goldhaber et al. 2010 Fowler 2003	Dwinal 2012

The next two medium strength pieces (2*) also showed positive results for recruitment, but not for retention. Steele et al. (2010) evaluated the Governor’s Teaching Fellowship (GTF) scheme, involving a \$20,000 incentive to attract and retain new teachers to low-performing schools for four years after becoming licensed. The teachers had to repay \$5,000 for each year that they did not meet the commitment, in a period when the average starting salary for California teachers was \$33,121. An instrumental variable design was used, based on 718 GTF teachers, excluding those who could not be tracked, were missing data, or not enrolled at recognised institutions. GTF recipients were not randomly selected, and so may have had a predisposition to teach in low-performing schools. More teachers were enrolled during GTF, around twice as many as in the years before and after, and 28% more taught in low performing schools. So it seemed that money was an attractor. However, there was no difference in retention rates (75% over four years) between recipient and non-recipients, despite the penalty clause.

Glazerman et al. (2013) examined the impact of the Talent Transfer Initiative, which offered bonuses to the highest performing teachers (those ranked in the top 20% in terms of raising student attainment from year to year using a value-added approach for each grade and subject) for agreeing to move to and stay in low-performing schools. The incentive was \$20,000 paid in instalments over a two-year period. Some teachers were already teaching in low-performing schools, and they received a \$10,000 retention stipend if they remained in the school over the two-year period. The participants included 85 teacher pairs matched on school characteristics and randomised to intervention or not, across 114 elementary and middle schools. Because the teacher pairs changed their personnel between randomisation and the start of the school year, the two groups were no longer equivalent at the beginning of the study. Of the vacancies assigned to the scheme, 88% were filled, compared to 44% the year before, and 71% in the comparison group. Retention after one year was 93% (70% in the comparator group), and 60% after two years (compared to 51%). The results suggest that while the transfer incentive may have had a positive impact on teacher recruitment and then retention rates during the payout period, the effect did not last once the payment stopped.

The weaker studies (in terms of design for a causal question) are more mixed in results. Fowler (2003) examined the Massachusetts Signing Bonus Program for New Teachers, offering a \$20,000 bonus for highly qualified people switching careers to teaching. Initially, recipients began teaching after seven weeks of training, although this was changed to a year-long programme in 2002, before being assigned to high-need schools, and provided with further training, support and mentoring. There was no explicit comparison group. The programme failed to recruit candidates from outside the area, and despite advertising across states only seven candidates outside Massachusetts were recruited over four years. This may be partly because other states were also experiencing severe teacher shortages, and some offered higher salaries. The programme also failed to place all teachers, mostly from Massachusetts, in high-need schools (only 71% from the first cohort, and 48% and 35% in following years). Dropout among bonus recipients was higher than the national average (46% by the third year), and highest in the high-need districts (55%). A survey of head teachers suggests that bonus recipients on the scheme

were more attracted to the fast-track scheme than the bonus incentive (Churchill et al. 2002). Evaluation of signing bonus incentives in general suggests that any effect tends to be short-lived (Choi 2011).

Gordon and Vegas (2005) analysed the impact of the Fund for the Maintenance and Development of Basic Education and Teacher Appreciation - a funding reform in Brazil which stipulated that at least 60% of additional funds be allocated to teacher wages. It has been linked to increased positive trends in student enrolment, and reduction in grade retention and dropout (World Bank 2002, Castro 1998). For this review we considered only the impact of funding on teacher numbers and qualification. The study was a longitudinal retrospective cohort study. Because the intervention coincided with major education reform in Brazil, such as increased economic changes, educational resources for some municipalities and the legislation that all teachers must be qualified, it is difficult to attribute any causal effect. Many of the patterns reported were present before the intervention, which was linked to an increase in the number but not the qualification of teachers, and a reduction in student: teacher ratios in some areas, but not the poorest ones. The impact of the programme is therefore difficult to assess.

In America, Goldhaber et al. (2010) compared salaries in private and public schools using data from the 1999-2000 School and Staffing Survey, the 2000 Common Core of Data, and the 2000 Census. The survey contains responses from 56,354 teachers in 5,465 public schools and 10,760 teachers in 3,558 private schools. Findings showed that private schools tended to pay slightly more for more qualified and experienced teachers than public schools, for teaching large classes, longer hours, and in more disadvantaged schools with a high proportion of ethnic minority pupils. For example, private schools with a high proportion of poor students paid their teachers 17% higher salaries than schools with an average number of poor students. This is more than in public schools, which often have similar schemes, and despite teachers working in public schools with a high proportion of ethnic minority students being paid a slightly higher salary. There are other differences between the two sectors, and teachers expressed concerns about working conditions, but one implication, as Goldhaber et al. concluded, could be that teachers will need to be paid more to get them to teach and stay in challenging schools.

The impact of initiatives through alternative routes to teaching remains unclear. Dwinal (2012) conducted a case study of Teach for America (similar to England's Teach First scheme) in the rural Mississippi-Arkansas Delta region, where there is a teacher shortage, geographical isolation and a heavily ethnically segregated school population. The programme recruited potential school leaders from university graduates and professionals, through an intensive selection process, and they committed to teach for at least two years in state schools. The low response rates (under 20%) to interviews with principals, and a comparison between regions over time using the weaker measure of vacancy rates rather than number of teachers recruited, made it difficult to establish the impact of recruitment. There was no decrease in vacancy rates relative to other areas, partly because the programme imposed limits to the number of participants in each district (so directing them elsewhere). This was a very weak study. Other studies suggest that Teach for America teachers tend to leave teaching after a couple of years (e.g. Glazerman et al. 2006, Decker et al. 2004, Raymond et al, 2002, Clark et al. 2017, Henry et al. 2014).

Clewell and Villegas (2001) reported a six-year evaluation of the Pathways to Teaching Careers programme, including paraprofessionals and noncertified teachers, and Peace Corps Fellows strands. The paraprofessional and noncertified programmes involved identifying non-qualified staff already working in schools and offering them scholarships as well as other support services to help them obtain qualified teacher status, after which they are committed to continue teaching in the schools for a specified period. The Peace Corps Fellowship identifies and supports potential teachers from

returning Peace Corps volunteers (similar to the Troops to Teachers programme in England). Fellows are placed in schools on a full-time contract and paid a salary where they work towards a teaching qualification. The study was largely based on self-report, with a high level of missing data. Only 44% reported where they were teaching initially, and only 31% after three years. Pathway teachers reported higher completion rates than traditionally certified teachers (75% to 60%). A high proportion (84%) ended up teaching in hard-to-staff schools, and had better retention rates over three years compared to national average (81% to 71%). They were also perceived to be more effective than typically qualified teachers.

Waters-Weller (2009) explored the relationship between improvement in working conditions (which they defined as reduction in class size and teaching load, plus more planning time), retention bonuses, and teaching and staying in high poverty schools. This was an exploratory cross-sectional study looking at the relationship between school intakes and attrition rates, the attitudes of teachers towards low socioeconomic status schools, and the kind of incentives likely to increase retention. The survey of 3,525 teachers in two urban districts only had a 29% response rate. The majority of teachers indicated that they would stay in their current school for the next year, including those who were in high poverty schools. They generally indicated that extra money for salaries and bonuses were not necessarily needed to keep them if the school had an excellent administrator, but money was an inducement to transfer to a poor school. The design of the study could not establish a causal link between improvement in working conditions or retention bonuses on retention, hence it was rated only 1* for strength of evidence.

Improving retention

There were 17 studies that examined the impact on teacher retention. Again, none were of the highest quality. The strongest study (3*) by Hough and Loeb (2013) suggested no lasting benefit from financial incentives for retention of teachers in hard-to-staff schools. Other studies discussed in the section on recruitment, also indicated no effects on retention (Steele et al. 2010 and Glazerman et al. 2013, Fowler 2003). There were a further 11 studies that dealt solely or mostly with retention of teachers in hard-to-staff schools (Table 4). Almost all of these reported positive effects, but these were largely very weak studies.

Table 4 – Quality rating of studies on retention

Quality of study	Positive outcome	Unclear outcome	Negative or neutral outcome
4	-	-	-
3	Clotfelter et al. 2007	-	Hough and Loeb 2013
2	Fulbeck 2011, Fulbeck and Richards 2015, Fulbeck 2014 Fitzgerald 1986 Falch 2010 Feng and Sass 2018 Gold 1987	-	Steele et al. 2010 Glazerman et al. 2013
1	Goldhaber et al. 2010 Lyons 2007 Anthony 2009 Helfeldt et al. 2009 Colson and Sattelfield 2018 Fuller 2003	Fowler 2003	Waters-Weller 2009

Although Clotfelter et al (2008), a 3* study, indicated a positive effect of financial incentives on retention, it concurred with the other studies above that incentives work only as long as they are available and once removed, they have no lasting effect. Clotfelter et al. (2008) examined the impact of the North Carolina annual bonus scheme on the retention of qualified maths, science and special education teachers in high poverty and challenging schools, using a difference-in-difference approach. Teachers received the bonus (\$1,800 per year) for as long as they stayed in the eligible school. This was a reasonably well-conducted study, using administrative data for four years on public school teachers to estimate the likelihood of teachers leaving a particular school. The research compared hazard rates before and after the implementation of the bonus programme, eligible and ineligible teachers in the same schools, teachers in eligible schools and those in schools that narrowly missed out on being eligible. Teachers receiving the bonus were an estimated 15% less likely to leave at the end of the school year compared to other teachers in the same schools.

Five further studies in this section were rated 2*, and suggested positive impacts. Gold (1987) evaluated the New York City Retired-Teachers-as-Mentors Program by comparing mentees with a comparison group of non-mentored teachers. The programme recruited retired teachers as mentors for new in-service teachers. Mentors attended a four-day workshop conducted by staff at a training college, were paid for 66 hours a year for each of three mentees, and assigned to schools with the highest attrition rate among new teachers. The study used Board of Education records and questionnaires completed by teachers, mentors and principals. Retention rates went up for all, but the rates were higher for the mentored teachers (85% and 80% in the second year). It is not clear whether mentors were randomised to new teachers in eligible schools, and no account was taken of missing data in the analysis.

Fitzgerald (1986) looked at offering an annual stipend (of between \$500 and \$2,000) to encourage teachers to teach in schools with a high proportion of pupils eligible for free or reduced lunches, in high priority areas in the US. The study used a difference-in-difference approach to compare the retention rates of teachers in 25 high priority schools with 25 high poverty control schools not receiving the stipend. The groups were similar in terms of pupil and teacher characteristics. Vacancies dropped in treatment schools in the first year, and the drop in retention rates was lower than for control schools (ES = +0.39).

In Norway, Falch (2010) examined the impact on the retention of teachers in high-vacancy schools of paying teachers differential wages, using a difference-in-difference approach. In the period 1993/4 to 2002/3, Norway had a central wage system, but teachers in schools with high vacancies received a wage premium of between 7.5% and 12%. Over the nine years, schools were initially eligible if they had 20% more “shortages” than the previous year. This increased to 30% for the 1996/7 and 1997/8, and then back to 20% for the last four years. In total, 161 schools received the wage premium at least once, and in these schools the attrition rate of teachers was lower than comparison schools by 6%. The reporting of this study, however, was not clear, and the number of schools and teachers included varied considerably over time. This makes it difficult to judge the efficacy of the incentive.

Feng and Sass (2018) considered the effects of the Florida Critical Teacher Shortage Program 1986 to 2011, on the retention of teachers in shortage subject areas (maths, science and special education). Loan forgiveness of up to \$10,000 to pay off their student loan was offered to beginning qualified teachers, if they taught in a shortage subject for at least 90 days. There was a recruitment bonus for new teachers, of up to \$1,200 (to cover removals or equipment), and a retention bonus of up to £1,200 if teachers continued to teach a shortage subject the next year, and had favourable performance appraisal. The latter two were only available from 2000/1 to 2001/2. Since subjects designated as shortage changed over time, the teachers eligible for these incentives also changed over time. These variations were used to compare bonus recipients with non-recipients, in terms of recruitment and

attrition using a proportional hazard model, taking into account student demographics, pupil prior behaviour, prior achievement, class size, teacher gender, race/ethnicity, salary base and experience. There is no report on attrition rates from the study. Loan forgiveness reportedly had a positive effect on the likelihood of teachers staying in teaching the following year, reducing attrition by 12%, but not once funding was removed. The one-time retention bonus for shortage subject teachers also reduced the likelihood of teachers leaving by 25%.

The next three studies have common authors and all examine the same intervention, and so they are treated as one complex study for this review. Fulbeck (2011) evaluated the impact of ProComp (Professional Compensation for Teachers) - a teacher incentive programme in Denver – including 10 financial incentives (seven individual, three school level). School-based incentives were awarded to teachers who teach at schools serving low-income students and high performing schools and schools that make the most progress in maths and reading. Eligibility was restricted to those who were members of teacher unions not working in Charter schools. The total number of teachers included in the retention analyses was 4,145, representing 91% of all Denver Public School District teachers. Retention figures exclude those who retired or whose service was terminated and those made redundant due to reduction in teaching posts. This study used interrupted time-series and difference-in-difference regression models. The average change in retention rate was -0.06% before ProComp and +1.5% afterwards, and participation in ProComp increased retention rates by 2.1 percentage points. It was more effective in hard-to-staff schools (ES 0.25) compared to others (ES 0.08). Retention was higher in high-poverty schools where teachers were eligible to receive a financial incentive to stay.

Fulbeck and Richards (2015) looked at all 7,333 public school teachers in Denver from 2006 to 2010 who were eligible for the ProComp incentive (regardless of whether they did receive it) and who made at least one voluntary move within the district (989). The incentive tended to attract teachers to high growth and high performing schools, and was less successful for schools with high proportion of low income pupils. A limitation of the study is its inability to take account of other factors that may over-estimate the effect of financial incentives, such as principal's hiring preferences and the actual school vacancies advertised.

Fulbeck (2014) looked at participation in ProComp and teacher mobility in high poverty areas, using longitudinal teacher-level data from 2001/2 to 2010/11, and comparing teachers who received ProComp with those who did not, and those who taught in high poverty schools with those who did not. Teachers working in high poverty schools were more likely to move but the odds of leaving the district (and so losing the incentive) were lower for ProComp teachers than for others. The study suggests that the incentive alone was not enough to compensate for poor working conditions, issues with school leadership and school climate.

The remaining studies were rated 1*. Lyons (2007) considered a teacher preparation programme where participants were volunteers, selected for their commitment to the goals of the programme. Unfortunately, much of the reporting is unclear. Findings suggest that teachers exposed to all programme components were less likely than the national average to leave classroom teaching after a year in a high-poverty school.

Anthony (2009) considered the impact on teacher retention of a mentoring system for new teachers in a rural school district in North Carolina. All new teachers were given a two-week intensive session to help them adjust to the school, community and the teaching profession, and were assigned to a trained mentor for three years. The State Board of Education also required all new qualified teachers to complete a three-year induction period to obtain a continuing Standard Professional 2 licence. Both mentors and mentees were given training. Data on retention, measured as a proportion of teachers

returning each year to the school system, was taken from the school system database. The proportion of teachers returning to the school system increased each year from 84% in 2005/6 before the programme to 92% in 2007/8. There was, however, no counterfactual as part of this study, and it is therefore a very weak study for a causal question.

Fuller (2003) examined the Texas Beginning Educator Support System on the retention of beginning teachers - a statewide comprehensive program offering instructional support and mentoring. Although this was a state-wide programme, participation was selective, and it is unclear how selection was organised. Using the state personnel database, the study compared the retention rates of beginning teachers who participated in the scheme with those not participating, from 1999/00 to 2002/03. The participants had higher retention, but this could be at least partly due to the prior selection process.

Helfeldt et al. (2009) described a four-year internship programme aimed at retaining new teachers in high-need urban schools. In this university-school partnership programme, interns were paid, with full teacher benefits, and worked as full-time regular teachers in the classroom. They were assigned an approved trained mentor, and \$8,000 from the intern's salary was paid towards this mentoring scheme. The sample only included 38 interns and 8 mentors, and the bulk of the analysis concerned participant perceptions of the programme. The programme was reported as effective in retaining teachers in high need urban schools with 100% of teacher staying on in teaching one year later, compared to state retention of 81%.

Colson and Satterfield (2018) tested the effects of a teacher compensation plan, known as the Innovation Acceleration Fund, on the retention of SEN (special educational needs), maths, science and language teachers in a small rural district. This was a merit pay system, paying teachers deemed effective based on the contentious Tennessee Value-Added Assessment System. The total potential population was reported as 134. Of these, 93 volunteered for the compensation scheme. Teachers who did not want to have individual teacher effect results were excluded. Only 56 of these were deemed effective. Around 80% of teachers who participated in the compensation scheme were retained compared to 70% who did not participate. The report does not include effect sizes, and the design means that volunteers were compared with non-volunteers.

Conclusion

Most of the work described here concerns financial incentives of some kind. In summary, financial incentives appear to work for recruiting teachers. Offering remission of student loans, higher salaries or premiums for teaching in hard-to-staff areas and schools is effective in attracting teachers. However, it is not clear that such external motivation is desirable, or attracts the best teachers, and it is quite clear that the attraction is not lasting.

While most of the higher quality studies indicate that financial incentives are effective in retaining teachers in hard-to-staff areas, the impact disappears once the incentive is removed. This is the case even though some of the financial incentives used in the US involved a kind of a tie-in, where teachers are committed to staying on in the school or district for a specified period or else incur a penalty. This suggests that financial incentives alone are not enough to keep teachers in challenging schools or in difficult areas. Survey responses from teachers suggest that they may be prepared to stay in less attractive schools or regions if they have supportive leadership and good working conditions (Waters-Weller 2009, Goldhaber et al. 2010, Fulbeck 2014).

Other approaches reviewed included mentoring and induction, teacher development and alternative routes into teaching. There is little evidence that any of these approaches work for recruitment, and no good evidence yet of anything else that works for retention, in high need areas. There were a number of studies that also looked at “grow your own” (training and recruiting from local community), but none of these could establish causation. Almost all these studies were based on stakeholders’ anecdotal reports of successful practice in their own school or district.

Most of the research we found was very weak, and all of the higher quality work involved easier-to-measure, more concrete strategies (such as financial incentives). More research with the kind of designs needed to address causal issues is urgently required to cover mentoring, support, training for teaching in difficult schools, and a host of other alternative approaches that could be combined with financial interventions to attract good teachers and then keep them where they are needed most. In the medium to longer-term a more comprehensive approach would be to change school allocation and economic policies so that there were no longer such clearly defined schools and areas with high levels of poverty (Gorard 2018), meaning that these schools would not be as hard to staff, even though some would remain geographically isolated.

We recognise that in any review of this scale some studies may have been missed, and new and more robust studies may be conducted in the future. This may alter the findings of our review, but given the evidence available at the time of this review the strongest evidence is for financial incentives.

References

- Anthony, J. (2009) *Teacher retention: Program evaluation of a beginning teacher and mentor program*, Education theses, Dissertations and Projects, Paper 100
- Beesley, A., Atwill, K., Blair, P. and Barley, Z. (2010) Strategies for recruitment and retention of secondary teachers in Central U.S. rural schools, *Rural Educator*, 31, 2, 1-9
- Castro, J. (1998) *The maintenance and development of teaching and appreciation of the Magisterium (FUNDEF) and its impact on financing of elementary school*, <http://repositorio.ipea.gov.br/handle/11058/2720>
- Choi, J. (2011) When are signing bonuses more than just 'pay to play'? An experimental investigation, *ProQuest Information and Learning*, 72, 3811-3811
- Churchill, A., Berger, J., Brooks, C., Efrat, A., Griffin, L., Magouirk-Colbert, M., McDermott, K., Sharick, R., and Sheehan, A. (2002) Revised interim report: *The Massachusetts Institute for New Teachers and Master Teacher/NBPTS Programs*, Amherst, MA: Center for Education Policy—University of Massachusetts
- Clark, M.A., Isenberg, E., Liu, A.Y., Makowsky, L. and Zukiewicz, M. (2017). Impacts of the Teach for America investing in innovation scale-up. Princeton: Mathematica Policy Research.
- Clewell, B. and Villegas, A. (2001) *Evaluating the Pathways to Teaching Careers program: Absence unexcused, ending teacher shortages in high-need areas*, T. U. Institute, Washington, D.C.
- Colson, T. and Satterfield, C. (2018). The effects of strategic compensation on teacher retention, *Power and Education*, 10, 1), 92-104, <https://journals.sagepub.com/doi/abs/10.1177/1757743818758782>
- Clotfelter, C., Glennie, E., Ladd, H. and Vigdor, J. (2008) Teacher bonuses and teacher retention in low-performing schools: Evidence from the North Carolina \$1800 teacher bonus program, *Public Finance Review*, 36: 63-87
- Decker, P.T., Mayer, D.P. and Glazerman, S. (2004). *The effects of Teach For America on students: Findings from a national evaluation*. Madison: University of Wisconsin-Madison. Institute for Research on Poverty.

- DeLaat, J. and Vegas, E. (2005) *Do differences in teacher contracts affect student performance? Evidence from Togo*, mimeo, Harvard University and The World Bank
- DfE (2019). *Cash incentives for maths and physics teachers*. London: DfE
- Dwinal, M. (2012) *Teach for America and rural southern teacher labour supply : an exploratory case study of Teach for America as a supplement to teacher labour policies in the Mississippi-Arkansas Delta, 2008-2010*, PhD thesis, <https://ora.ox.ac.uk/objects/uuid:ec68169b-bf6c-4659-82a9-3fe8be3fa883>
- European Commission/EACEA/Eurydice (2018) *Teaching Careers in Europe: Access, Progression and Support. Eurydice Report*, Luxembourg: Publications Office of the European Union
- Falch, T. (2010) *Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment*, Working Paper Series 10910, Department of Economics, Norwegian University of Science and Technology
- Feng, L. and 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
- Fitzgerald, C. (1986) *Report on the High Priority Location Stipend Program*, Miami, FL: Dade County Public Schools Office of Educational Accountability, <https://eric.ed.gov/?id=ED283862>
- Fowler, R. (2003) The Massachusetts signing bonus program for new teachers: A model of teacher preparation worth?, *Education Analysis Archives*, 11, 13, 1-24
- Fulbeck, E. (2011) *Teacher retention: Estimating and understanding the effects of financial incentives in Denver*, ProQuest LLC, https://scholar.colorado.edu/educ_gradetds/100/
- Fulbeck, E. and Richards, M. (2015) The impact of school-based financial incentives on teachers' strategic moves, *Teachers College Record*, 117, 9
- Fulbeck, E. (2014) Teacher mobility and financial incentives: A descriptive analysis of Denver's ProComp, *Educational Evaluation and Policy Analysis*, 36, 1, 67-82, <https://journals.sagepub.com/doi/abs/10.3102/0162373713503185>
- Fuller, E. (2003) *Beginning teacher retention rates for TxBESS and non-TxBESS teachers*, Unpublished, State Board for Education Certification, Texas, unpublished
- Glazerman, S., Mayer, D. and Decker, P. (2006). Alternative routes to teaching: The impacts of Teach For America on student achievement and other outcomes. *Journal of Policy Analysis and Management*, 25,1, 75-96.
- Glazerman, S., Protik, A., Teh, B., Bruch, J., Max, J. and Warner, E. (2013) *Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment. Executive Summary*, NCEE 2014-4004, Washington DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education
- Gold, M. (1987) *Retired teachers as consultants to new teachers: A new inservice teacher training model*, Final report, American Association of State Colleges and Universities, Washington, D.C. of New York. Institute for Research and Development in Occupation Education
- Goldhaber, D., Destler, K. and Player, D. (2010) Teacher labor markets and the perils of using hedonics to estimate compensating differentials in the public sector, *Economics of Education Review*, 29, 1, 1-17
- Gorard, S., See, B.H. and Siddiqui, N. (2017) *The trials of evidence-based education: The promises, opportunities and problems of trials in education*. London: Routledge.
- Gorard, S. (2018) *Education policy: Evidence of equity and effectiveness*, Bristol: Policy Press
- Gordon, N. and Vegas, E. (2005) Educational finance equalization, spending, teacher student outcomes: The case of Brazil's FUNDEF, in Vegas (Ed.) *Incentives to improve teaching: Lessons from Latin America*, Washington: World Bank Press
- Helfeldt, J., Capraro, R., Caparo, M., Foster, E. and Carter, N. (2009) An urban schools' university partnership that prepares and retains high quality teachers for "high need schools", *The Teacher Educator*, 44, 1, 1-20

- Henry, G.T., Purtell, K.M., Bastian, K.C., Fortner, K.C., Thompson, C.L., Campbell, S.L. and Patterson, K.M. (2014). The effects of teacher entry portals on student achievement. *Journal of Teacher Education*, 65, 1, 7-23.
- House of Commons (2017). *Recruitment and retention of teachers fifth report of session 2016-2017. HCC 199*. London: House of Commons,
- Hough, H. and Loeb, S. (2013) *Can a district-level teacher salary incentive policy improve teacher recruitment and retention?*, Policy Brief 13-4, Policy Analysis for California Education, PACE, <https://www.aeaweb.org/articles?id=10.1257/aer.101.3.460>
- Lyons, K. (2007) *Preparing to stay: A quantitative examination of the effects of pre-service preparation on the retention of urban educators*, Ann Arbor, University of California, Los Angeles, 3295729: 171
- Navarro, C. and Verdisco, A. (2000) *Teacher training in Latin America: Innovations and trends*, Sustainable Development Department Technical Paper Series, Washington: Inter-American Development Bank
- Raymond, M., Fletcher, S.H. and Luque, J. (2001). *Teach For America: An evaluation of teacher differences and student outcomes in Houston, Texas*. Palo Alto, CA: Stanford University.
- See, BH and Gorard, S. (2019) Why don't we have enough teachers?: A reconsideration of the available evidence, *Research Papers in Education*, <https://www.tandfonline.com/doi/full/10.1080/02671522.2019.1568535>
- Sibieta, L. (2018) *The teacher labour market in England: Shortages, subject expertise and incentives*, London: Education Policy Institute
- Social Mobility Commission (2017) *State of the Nation 2017: Social mobility in Britain*, London: Social Mobility Commission, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/662744/State_of_the_Nation_2017_-_Social_Mobility_in_Great_Britain.pdf
- Steele, J., Murnane, R. and Willett, J. (2010) Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California, *Journal of Policy Analysis and Management*, 29, 3, 451-478
- Vegas, E. (2007) Teacher labor markets in developing countries, *Future of Children*, 17, 1, 219-232
- Waters-Weller, C. (2009) Attracting veteran teachers to low socioeconomic status schools: Initiatives and considerations, US, *ProQuest Information and Learning*, 69: 2979-2979
- Wheeler, J. and Glennie, E. (2007) *Can pay incentives improve the recruitment of teachers in America's hard-to-staff schools?* A research summary: Policy matters, Durham, NC: Center for Child and Family Policy, Duke University
- World Bank (2002) *Brazil Municipal Education, Resources, Incentives, and Results*, <http://documents.worldbank.org/curated/en/969591468016270639/Research-report>