Intention to Teach: Incentive Impacts of Bursaries

Rebecca Rose *

October 25, 2024

Most recent version

Teachers are an essential input of the education production function. However, amid global teacher shortages, research on policies to attract new teachers remains limited. I assess the impact of financial incentives on the recruitment and retention of trainee teachers. Using a panel of UK teachers, I exploit policy-induced variation in the bursary levels offered across years, subjects, and the trainee's undergraduate classification. Results suggest that a £10k increase in training bursary leads to a 34% rise in trainee recruitment, and a 14% increase in the remaining teacher cohort size three years later. However, trainees are 1.8 percentage points less likely to become a teacher post-training, which is driven primarily by unobservable selection rather than observable personal characteristics. The cost of gaining 5,000 additional teaching years through raising bursaries is roughly equal to the cost through raising teacher pay. Raising training bursaries is a flexible tool to address teacher shortages that can be easily targeted at specific sub-groups, but leads to compositional effects that can impact the long-term motivation and occupation decisions of the resulting teacher workforce.

JEL Classification: I22, J44, J24, J38, H52

Keywords: Teachers, Retention, Recruitment, Financial Incentives

---

*Department of Economics and Centre for Economic Performance (CEP), LSE. Email: r.rose@lse.ac.uk

I am grateful to Stephen Machin, John Van Reenen, and Alan Manning for their invaluable guidance and support. I would also like to thank the faculty, students, and researchers at the CEP and LSE for sharing their insightful advice and knowledge. Lastly, I would like to thank the participants of various seminars and conferences for their feedback and helpful comments.

1 Introduction

Teacher shortages are an urgent issue across both high and low income countries, to such an extent that in 2024 the UN issued a 'global alert'[1]. In the UK, larger incoming pupil cohorts and high teacher attrition rates are putting pressure on recruitment targets. Trainee teacher retention is also particularly low: around 25 percent of trainee teachers do not progress to teaching in a public school. In order to attract more staff, the UK government offer bursaries (financial incentives which are similar to scholarships) as financial support to graduates who undergo a one year course to retrain as teachers. Despite this policy, recruitment targets for trainee teachers have been missed for nine of the last ten years [2]. Evaluation of the effectiveness of such financial incentives is particularly pertinent at a time when the recently elected government has pledged to employ 6,500 additional teachers.

Understanding teacher recruitment and retention is essential because teachers are a key determinant of the educational attainment of students, and therefore impact future human capital and growth (Rivkin et al. (2005), Hanushek (2011)). There is evidence that lower pupil-teacher ratios improve test scores (Angrist and Lavy (1999), Finn and Achilles (1999)), and that teacher value-added increases with experience (Clotfelter et al. (2006), Rockoff (2004)). If there is insufficient supply of teachers, or if due to poor retention teachers leave the profession too quickly, the resulting lower educational attainment could also reduce long run productivity and innovation.

Whilst numerous papers study the effectiveness of tools designed to improve the retention of the existing teacher workforce, evidence on how to increase the supply of new teachers is more limited. In this paper, I evaluate the impact of a UK policy that offers financial incentives to graduates who retrain as teachers. This is a unique setting where financial incentives are large and the training program is a major entry route for new teachers. I examine whether these incentives increase the number of trainees, and how they impact the characteristics of those recruited. Given that bursaries lead to a change in the composition of trainee cohorts, they can also impact the attrition of teachers once they are employed. I therefore evaluate the policy's long-run effect on teacher retention, separating the effects into those stemming from both observed and unobserved selection components.

My analysis combines administrative teacher panel data with trainee records to

[1]Source: UN News at [https://news.un.org/en/story/2024/02/1147067] accessed 29/07/2024

[2]Source: ITT Statistics (DfE) at [https://www.gov.uk/government/collections/statistics-teacher-training]

track teachers across their career. I also use an administrative wage panel data set and labour force survey data to measure local labour market conditions and movements between occupations. The richness of these data sources allows me to exploit policy induced variation in wages and exogenous variation in bursary levels. The bursary itself is an incentive offered to all who enrol onto a one year training course which awards them qualified teacher status at the end. The amount a trainee is eligible for is between £0 to £30,000 and depends on the year of training, the grade classification of their undergraduate degree and the subject they have chosen to train to teach. Trainees are not required to repay the bursary if they do not teach after training.

I find that increasing the bursary level by £10,000 (roughly one standard deviation and a third of a teacher's starting pay) significantly increases entry cohort size by 34 percent. I then separate the analysis by the training subject and the individual's undergraduate classification and find that STEM (Science, technology, engineering, and mathematics) applicants are more responsive. A £10,000 bursary increase raises STEM cohort size by 51 percent, but only increases Non-STEM size by 23 percent. I analyse teacher cohort size three years after training and observe a positive but insignificant increase of 14 percent for the pooled sample. This suggests that the boost in recruitment is reduced as cohorts with bigger bursaries are subject to greater attrition. In STEM subjects, long-term teacher cohort sizes are still 42 percent larger, but growth of non-STEM teacher cohorts is small and insignificant from zero.

I find that retention decreases when the bursary level increases, and that this is driven by the unobservable characteristics of trainees. By running a logistic regression on key retention outcomes, I find that individuals are 1.8 percentage points less likely to teach post-training following a £10,000 increase. The magnitude and significance of this effect is consistent when controlling for personal characteristics and outside wages. Dis-aggregating by subject type, I find that this pattern is strongest and only significant for STEM trainees, who are 2.8 percentage points less likely to become a teacher. Summarising the recruitment and retention results, generous

STEM bursaries are effective in attracting more trainees. But they are less likely to become teachers compared to similar trainees with a lower bursary. The attrition of non-STEM trainees is not significantly affected.

I construct a simple model of dynamic occupational choice where agents have heterogeneous non-pecuniary valuations of teaching (motivation) to interpret the empirical results. The model suggests that a higher bursary reduces the minimum required motivation to enter training and increases the number of trainees, but also reduces the trainee-to-teacher conversion rate through attracting more 'dropouts' who leave after the training is complete. By considering variation in the outside options of individuals, the model also explains the observed differential recruitment and retention impacts across undergraduate classifications, and differential impacts between STEM and non-STEM courses.

Lastly, I construct a cost-benefit analysis that compares the cost of gaining 5,000 additional teaching years through bursaries with a menu of other recruitment and retention policies. Raising the bursary for all trainees by around £6,000 results in an average cost per teacher-year of £58,000. This is approximately the same cost as raising all teacher pay by 0.7%, but more expensive than raising early-career pay[3]. However, bursaries have the additional benefit of flexibility compared to pay rises as they can be easily varied across subjects, years, and personal qualifications[4]. Policy makers should also consider the trade-off between recruiting new teachers versus retaining experienced teachers.

My findings contribute to the teacher recruitment literature by establishing that financial incentives can have both positive recruitment and negative retention effects. Since the training programme is widely available to all high-skill workers, I can also evaluate the policy as the impact of a one-time financial incentive on occupational choice. Teaching is an ideal example of an alternative occupation to high-skill workers due to its availability across geography, wage certainty and low unemployment risk.

[3]Early Career Teachers are those who are in their first 2 years of teaching post-qualification.

[4]Teacher pay is generally set in line with national pay bands which do not vary by subject. Unequal changes to the pay scale risk being politically unpopular and generating union resistance.

My work also serves as an evaluation of a large-scale recruitment policy with large payments that is one of the primary paths into teaching in the UK. The policy itself is an important alternative to general pay rises because bursaries can target shortage subjects and are one-time transfers.

Overall, bursaries are an effective tool whose positive recruitment effects overpower their negative retention effects. However, high bursary cohorts are subject to compositional effects that may not be easily observed or readily apparent when initial recruitment occurs. Marginal trainees are relatively more financially motivated and more analysis is required to establish how this may affect the cohort-level elasticity of labour supply. Teacher cohorts recruited using large bursaries could even be more receptive to different types of retention and motivation incentives. The most important question remaining is whether these marginal candidates differ in their teaching quality. Additional research is required to assess whether high-bursary teachers improve not only the pupil-teacher ratio, but also pupils' educational attainment.

The rest of the paper is as follows: In section 1.1 I discuss the key contributions of this paper. I then provide the context of the policy and discuss the data in sections 2 and 3 respectively. In section 4, I describe the empirical strategy. Section 5 presents the empirical form results and section 7 offers an explanation of these results using a simple model of occupational choice. I discuss robustness in section 8, policy implications in section 8 and conclude in section 9.

1.1 Contribution and Literature

Existing research finds a positive impact of offering financial incentives to the existing teacher workforce on retention. A high wage premium reduces likelihood of attrition (Falch (2011)), and inexperienced teachers are most responsive to this (Hendricks, 2014). Bonuses are also effective at reducing attrition in schools with high poverty rates (Clotfelter et al., 2008) and in short-staffed subjects (Sims and Benhenda, 2022). Bueno and Sass (2018) also find a reduction in attrition through subject-specific bonus awards, but note that the policy had no impact on teaching graduates selecting into those subjects. Feng and Sass (2018) explore a range of financial incentives, including a one-time recruitment payment similar to the one studied in this paper. They find a positive impact of a bonus on retention, but only evaluates this for a small bonus in the year it was awarded.

A handful of existing papers evaluate the impact of financial incentives on new teachers but focus on different types of incentives. Coffman et al. (2023) and Coffman et al. (2019) find that offering a combined grant and loan of up to $1,800 to the most financially constrained trainees increases their uptake into the Teach for America (TfA) program, their completion rate, and their retention in the medium term. Whilst TfA is similar to the UK's PGITT program in that both require an undergraduate degree, TfA is designed as a selective, short term, prosocial program; whereas the PGITT is a major channel of recruitment for permanent teachers. Secondly, Coffman et al.'s work focuses on liquidity constraints which is less relevant in the UK context where all trainees are eligible for postgraduate loans[5]. De Falco et al. (2024) examine the impact of an undergraduate teacher training scholarship in Chile targeted at high-performing students on teacher quality. The policy was effective in attracting more applicants with higher test scores that progress to having higher teacher value-added (TVA) and are no less intrinsically motivated. I complement this work by analysing retention impacts, as Chilean scholarship recipients were required to teach after training. Lastly, a report for the National Foundation for Education Research generates a cost-benefit analysis of the same UK bursary policies (McLean et al., 2023).

The relevance of bursaries as a recruitment tool cannot be understated in the presence of myopic or financially-constrained workers. Christian et al. (2024) found that treating US undergraduate students with information generally corrects their biased beliefs about the relative pecuniary and non-pecuniary payoffs of teaching, however this only had a minor and weak impact on transferring their major to education. Similarly, Biasi (2024) shows that teachers respond four times less to

[5]UK Postgraduate loans are also repaid as a percentage of income over a certain threshold, which eliminates the risk of being unable to repay student debt post training completion.

a change in pension than a change in salary. Bursaries and scholarships can act as a front-loaded incentive that may have a larger impact on recruitment than financial incentives in later years.

My paper adds to the teacher recruitment literature in three ways: Firstly, I analyse the effectiveness of financial incentives on new teachers by exploiting a unique setting where quasi-random financial incentives are large, the training program is a major entry route for new teachers, and training represents a widely available outside option for all graduates. Secondly, I consider the long term impacts of a financial incentive on the underlying characteristics and retention of the trainee population. In other words, I not only estimate the effect of the policy on teacher cohort size, but also explore its composition and retention effects. Lastly, I consider the policy in the context of occupational choice, and evaluate how the financial incentive interacts with outside wages. I develop a simple occupational choice model that complements these results.

I also contribute to the literature on how financial incentives affect the ability and prosociality of new hires. The introduction of a bursary improves the short term return to teaching relative to other occupations, which alters the quality of trainees. In a randomised experiment in Mexico, Dal Bo et al. (2013) show that higher civil service salaries attract more skilled workers, but most importantly do not result in adverse selection effects on motivation. Ashraf et al. (2020) also note that when recruiting Zambian health workers, marginal recruits attracted by career benefits were no less prosocial and led to improved health outcomes, despite the pool of applicants being on average less pro-social. Turning attention to schools, Leaver et al. (2021) study the impact of recruiting teachers for pay-for-performance roles and find worse intrinsic motivation as measured at baseline. But despite negative selection, job outcomes were largely similar, and performance pay led to increased effort, improved student outcomes, and no effect on retention. Conversely, Deserranno (2019) finds that higher paid positions act as a signal and disincentivise prosocial applicants from applying, and as a result retention is worse. On balance, financial incentives seem to affect the pool of applicants, but the effect on job outcomes depend on how applicants are filtered during the hiring process. Lastly, Abebe et al. (2021) measure the impact of providing a small monetary incentive for applying to a position. By reducing the cost associated with making an application, they find an improved ability to recruit quality workers. In my setting, the opportunity cost of training is large and takes one year, therefore the bursary may have a similar effect.

I additionally speak to the literature on labour market conditions and selection into the public sector. When outside wages are higher, the quality of recruits into the public sector fall (Nickell and Quintini (2002), Propper and Van Reenen (2010), Crawford and Disney (2018)). Fullard and Zuccollo (2021) note that due to England's inflexible pay structure for teachers, large local pay gaps discourage workers from entering the occupation. Fullard (2021) estimates the impact of the local unemployment rate on enrolling as a trainee teacher in the UK using centralised admissions data. He finds that a higher unemployment rate does not impact the probability of enrolment, but does impact the diversity of those enrolled. More ethnic minority trainees, more men, and more trainees from 'prestigious' universities are enrolled. Nagler et al. (2020) find that teachers who enter during a recession are more effective at raising student test scores. This suggests that a higher bursary that would result in positive selection at the recruitment stage. However, teacher training is not just a route into teaching. It also presents an opportunity to shield from unemployment and develop human capital. Enrolment into undergraduate education (Dellas and Sakellaris, 2003) and postgraduate education (Bedard and Herman, 2008) increases during times of economic downturn. Therefore a higher bursary could induce a lower transition rate from training to working as a teacher.

2 Context

In 2011, UK education minister Michael Gove expanded the provision of teacher training bursaries, with an aim to raise the status of the profession and the quality of teachers within the UK. The bursary is a financial incentive offered to those who undertake a one-year intensive training programme to become a qualified teacher. The policy intended to target graduates with high performance in their undergraduate degrees (Department for Education, 2011).

School staffing issues have not been resolved since the introduction of these policies. The secondary school student-teacher ratio has been steadily increasing, and the government have since found themselves in a 'recruitment crisis' where recruitment targets have been consistently missed since 2013, with the exception of 2020[6], the pandemic year. In 2023/24 recruitment was only 62 percent of its goal and only 17 percent for physics, a particularly short-staffed subject. Teachers report high levels of dissatisfaction, with 25 percent considering leaving in the next 12 months for reasons other than retirement (Adams et al., 2023). They cite high workloads, government initiatives, and pressures relating to student outcomes or inspection. 61 percent of teachers were also dissatisfied with their pay. Between 2010 to 2022, real wages fell 13 percent for experienced teachers and 5 percent for new teachers (Sibieta, 2023). In 2023, teachers in England went on strike and consequently negotiated a 6.5 percent increase in pay.

The postgraduate initial teacher training course (PGITT) is a one-year postgraduate course that combines theoretical learning (often based in a university) and on-the-job classroom training in at least two schools. Anyone with a UK undergraduate degree or higher, or its equivalent value, is eligible to apply. Training is provided by many different accredited organisations across the country and students apply to each provider separately. Each provider may differentiate itself based on its schooling partners, or the type of support it offers its trainees. A number of training routes are also available, including a school direct route in which trainees are paid a non-qualified teacher salary rather than a bursary. Trainees that complete the course are awarded Qualified Teacher Status (QTS), and depending on the course can also attain a Postgraduate Certificate in Education (PGCE). Whilst a PGCE is

[6]Source: ITT Statistics (DfE) at [https://www.gov.uk/government/collections/statistics-teacher-training]

a postgraduate qualification, it is equivalent to only a third of a masters degree.

PGITT is not the only way to attain QTS, however it is a uniquely interesting route. Individuals of any undergraduate specialism can enrol to re-train and change occupation. This makes teaching a common outside option for higher educated workers with a transparent pay schedule and low unemployment risk. The Postgraduate ITT course is also the most common route of teacher training, representing around 80% of trainees. Other routes include an undergraduate degree in teaching, Teach-first (a Teach for America equivalent), and an assessment-only ITT for those who have had sufficient teaching experience as an unqualified teacher. PGITT is the quickest and most flexible route into teaching. It is also distinct from teach-first, which is more charitable in its mission and operates on a smaller scale.

The bursary level awarded to a trainee varies based on three characteristics: the subject they are training to teach, the year they undertake training, and the classification of their undergraduate degree (or a higher qualification). Classification is the overall grade of a degree, roughly equivalent to a banded version of a GPA[7]. Trainees must also be a UK national to be eligible for funding. The government reviews bursary levels on an annual basis through an opaque process. However, bursaries are generally higher for students with higher degree classifications, and subjects with a distinct teacher shortage. Students receive their bursaries in monthly installments, and payments are not conditional on becoming a teacher after training. Students that drop out are no longer paid monthly, but are not typically required to reimburse their previous payments. The maximum bursary awarded over the time period of this analysis is £30,000, and the minimum is £0. Figure 1b shows the within-subject variation in bursary levels over time for physics and modern foreign languages respectively. As an example, a physics graduate who attained a 2:1 (upper-second) classification would have a bursary of £25,000 in 2016, but a physics graduate with a first-class would be offered a bursary of £30,000. Bursaries are not announced more than a year ahead of time, so potential trainees are not aware of the payoff of

[7]The grades in order of highest to lowest are First-Class (1st), Upper-Second Class (2:1), Lower-Second Class (2:2) and Third-Class (3rd).

waiting an additional year to train.

Figure 2 describes the timeline for teacher training. For a training course that would commence in September of year 2, bursaries are announced in October of year 1. Applications are then open on a rolling basis up until September and involve submission of a personal statement, interviews, and references. Candidates apply directly to training providers, and can re-apply within the same cycle if their first application was unsuccessful. The bursary is paid monthly during training, and individuals typically qualify by June of year 2. Whilst undergoing training, individuals may apply for teaching roles that will commence as early as September year 3 (the next academic year).

The tuition fee for an ITT course is £9,250[8], however all students eligible for a bursary are also eligible for a postgraduate student loan for the full amount. Maintenance loans are also available depending on household income. Student loans in the UK are provided through the UK government and are only repaid after the individual's income is over a set threshold[9]. Past this point, individuals have 6% of their income automatically deducted from their paycheck.

3 Data

I have two main sources of data provided by the UK Department of Education (DofE). The first is the Initial Teacher Training (ITT) dataset which contains the details of ITT trainees from 2013 to 2020. The data contains information on the person's course: their training subject, provider, training route, and their course status: (pass/fail/ongoing). I also observe their sex, ethnicity, age, undergraduate degree subject, and undergraduate classification. I do not observe any funding information, including the level of bursary awarded, so I match each trainee with the amount they are eligible for based on their qualifications, training subject, and year. A trainee's location is inferred by the provider they train with. I include all trainees enrolled

[8]Raised from £9,000 in 2017

[9]A Qualified Teacher's starting salary is above the repayment threshold.

in fee-funded postgraduate training routes: Higher education institution-led, School centered, and School direct [10]

The second source of data is the School Workforce Census (SWC) which is an annual census submitted by schools each autumn containing information on teachers, teaching assistants, and other non-classroom based school support staff. It is a panel that contains information on personal characteristics, pay, qualifications, absences and vacancies, and details on the subject taught, hours worked, and additional roles of teachers. I have access to the SWC from the years 2013 to 2020. Teachers are identified in the SWC using a unique anonymised ID, and can therefore be tracked across time as long as they remain employed within a public sector school. ITT trainees are linked to the SWC data, so I am able to track their career progress. Unfortunately, trainees are not captured in the SWC during their training year as they are not formally employed by their school (excluding school-direct trainees). I therefore do not observe either of the schools that trainees work in during their studies. The SWC also contains no information on pupil performance and therefore I do not construct a measure of teacher value-added[11]. Given the first order concern of teacher recruitment, I focus on the quantity of teachers instead.

My analysis uses data for trainees in the 2013-2019 cohorts. In 2013 the payments became more generous and varied based on undergraduate classification, which provides a richer source of variation to exploit. Analysis stops in 2019 as in 2020 Covid drastically altered the labour market and led to a resulting surge in ITT applications. I focus on trainees in England as course fees and funding structure can vary across countries within the UK and therefore affect results. I also restrict analysis to secondary school focused ITT trainees (educating ages 11-16) as training to be

[10] As of 2022/23, fee-funded routes made up 88 percent of postgraduate trainees and only 11 percent of school direct trainees were salaried. In the robustness section I exclude those enrolled in school direct programmes who may receive a salary rather than a bursary.

[11] The DfE dataset on pupil performance is currently not linked with their history of teachers. Accurately measuring TVA for all teachers in the study would also face additional barriers. Students also take just one set of national tests (GCSEs) at the end of secondary school by which time each student will likely have been taught by multiple teachers for the same subject across different years.

a primary school teacher differs in many respects to secondary school teaching, and the SWC contains more detailed information for secondary teachers.

Lastly, I use several datasets to infer context about labour market conditions that may influence the choices of potential trainees. The Annual Survey of Hours and Earnings (ASHE) is a 1% sample of earnings data provided by firms that includes accurate pay, occupation, and job sector information but limited information on employee's personal characteristics. The Labour Force Survey (LFS) includes self-reported information about pay and occupation, but also includes information about personal characteristics, training and education.ASHE provides rich wage data, but is unable to differentiate wages based on educational qualifications. The LFS therefore supplements ASHE with information about pay and unemployment for the sub-sections of the population relevant to this study[12]. I focus on those eligible to receive a bursary: UK nationals with an undergraduate degree.

3.1 Descriptive Results

I include 95,397 trainees across seven years. Table 1 shows that around half of trainees are early career (under 26), and 81 percent are eligible for a bursary. Whilst the majority of trainees are female, the rate is lower than the overall share of secondary teachers which currently stands at 64 percent[13]. The trainee-to-teacher-conversion rate is 74 percent, meaning that one-quarter of trainees never take up a job as a qualified teacher across the observed time frame[14]. These are also clear distinctions between STEM and non-STEM trainees, which may reflect the composition effects stemming from the differing generosity of bursary payments offered, and the different labour market conditions faced by groups with different undergraduate specialisms.

Overall, I observe 12 different bursary levels across 17 subjects. Figure 3 shows the

[12]Note that these data sets are not linked.

[13]Source: School Workforce Statistics (DfE) at [https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england]

[14]Whilst private schools are not observed in this data, employees working for private secondary schools make up for around 4 to 15 percent of the entire secondary school workforce [Source: IDBR, March 2023].

distribution of the size of bursaries awarded to trainees. STEM graduates generally get higher bursaries and are skewed towards higher values. In contrast, non-STEM trainees are more likely to receive no bursary at all, and bursary levels are skewed towards zero. Trainee retention generally worsens as bursaries increase. By grouping bursaries into four general levels, Figure 4 plots the share of teachers that are present in the SWC in each year subsequent to their training year. For each group, the drop out rate is highest directly after training, with about 35% of trainees dropping out, however retention continues to slope downwards after this. Figure 4 also shows that post-training retention generally worsens with higher awards.

Teacher pay is determined by a nationally set pay schedule, where individuals may progress up the scale based on their performance, experience, and seniority[15]. Any individual considering a career in teaching can therefore transparently observe their potential pay progression. This strengthens the concept of teaching as an outside option, as it is a more predictable and steady career than other private sector occupations. Figure 5 demonstrates how average pay progresses over teaching tenure. Pay progresses fastest during the first ten years of teaching, however after this it stagnates at the same level for junior teaching faculty. Only for senior teaching roles does pay consistently progress. However, not all teachers are guaranteed to proceed to a senior role. The share of teachers in a senior role increases until it is roughly constant at 40-45 percent for each teaching cohort with over 20 years of experience.

4 Empirical Strategy

4.1 Identifying Variation

I exploit variation in the bursary levels offered to students to estimate the impact of financial incentives on a number of trainee outcomes. Bursary levels are determined annually by the Department for Education (DofE) but the process by which they

[15] Also note that teacher pay scales do not vary by undergraduate degree classification, whilst bursaries can.

are set is opaque. Since these levels are unlikely to be randomly assigned, I address two key potential confounders: teacher demand and the alternative labour market opportunities for trainees. For example, the DofE would offer larger bursaries in subjects or years with more demand for new teachers, which may also correlate with teacher workload, vacancy availability, or other non-observable characteristics of the teaching experience. The same argument applies when the graduate labour market is strong: a higher bursary may be required to compensate trainees for their time. I explain how I address these biases in my analysis, and argue that there remains a sizeable amount of variation in bursary levels not accounted for by these confounders that I am able to exploit[16].

I construct a control variable for each of the two confounders. The DofE publishes subject-specific teacher training targets each year[17] which reflect the level of demand for teachers. For each individual I also estimate their outside wages in each year, which I discuss in detail in appendix section A. Regressing bursaries on outside wages and trainee targets confirms that both variables are significantly positively correlated. I therefore directly control for both variables, which leaves only the random variation in bursary levels orthogonal to the labour demand of both teachers and non-teachers. Eighty percent of the total variation in bursary level remains when adding these controls, which suggests that bursary levels are not solely determined by these two factors[18].

I also offer four arguments as to why the remaining variation in bursaries after controlling for confounders contains valid exogenous variation. Firstly, bursary levels vary at the subject-year-degree classification level so I am able to exploit within

[16]It is also likely that these factors bias the effect of the bursary on cohort size towards zero if higher bursaries are offered in cases where supply of teachers is particularly low.

[17]These targets are set by a teacher workforce model that takes into account expected teacher re-entires and exits, the fulfilment of previous year's targets, and changes in student populations

[18]I run a regression of trainee target and outside wages up to the 8th power on bursary level at the individual level to assess what share of the total variation is explained by these two variables. The r-squared of this regression is 0.2054. I also compare the r-squared of two regressions that also control for characteristics and course-related variables; where only one controls for wages and trainee-targets. The increase in r-squared from the inclusion of these variables is 0.0252.

cohort variation across time whilst controlling for year, subject, and degree classification. This means that I can compare two trainees taking the same course in the same year (and therefore experiencing the same teacher labour market), but whose bursaries differ due to different undergraduate classifications. The identifying assumption is that bursary levels are not correlated with any changes at the year-subject-classification level. Secondly, If bursaries were determined using a formula, they would still need to be rounded to generate the values we observe. Therefore there is a bursary component that is orthogonal to its potential determinants. Thirdly, Bursaries are determined at least a year before the course commences, and two years before the first retention outcome is realised. The actual labour market conditions faced by graduates will contain additional random components that were not considered when calculating the bursary level. Lastly, The National Audit Office (2016) stated that "The Department [of Education] has not assessed the impact of bursaries on applicants' success or the number who go on to qualify and teach". Referring to the teacher workforce model algorithm that determines trainee targets, they wrote that "The Department is yet to demonstrate how accurate the model and its own judgements are". If the accuracy of the teacher workforce model was undetermined and the impact of bursary levels is unknown at the time these bursaries were set, it is reasonable to assume that bursaries are not precisely calculated and will contain random variation.

4.2 Cohort-level Characteristics

Firstly, I explore the impact that financial incentives have on the size and characteristics of trainee cohorts. Equation 1 estimates the effect of the average bursary level $\overline{B}_{jt}$ in subject j in year t on the log of the number of trainees enrolled, N_{jt}. β therefore represents the semi-elasticity of cohort size. I control for year and subject fixed effects and the target number of trainees for each subject-year cohort, as reported by the Department of Education. In addition to estimating the impact of the bursary on the log of trainee cohort size (at time = 0), I also run the same regression on the remaining cohort size of teachers three years later (at time = 3). All regressions are run separately for STEM subjects, non-STEM subjects, and all subjects combined. Owing to small sample sizes and a small number of clusters when regressions are run at the subject level, robust standard errors will be biased downwards. I therefore use a wild bootstrap estimation with standard errors clustered at the subject level. In the robustness section of the paper, I discuss alternative methods of estimation to correct for potential biases.

$$log(N_{jt}) = \beta \overline{B}_{jt} + \alpha_j + \alpha_t + \gamma TraineeTarget_{jt} + \epsilon_{jt} \tag{1}$$

I also examine how bursary levels impact the observable characteristics of trainee cohorts. Equation 2 estimates the probability of an individual in subject j in year t having a certain characteristic, where C_{idjt} is the propensity to have characteristic C and ϵ_{idjt} is distributed logit. I run this analysis at the individual level, as it avoids the use of bursary averages and allows me to control for the degree-classification held by the individual, as well as including year and subject fixed effects and trainee targets. I consider the following characteristics: being female, having a non-white British ethnicity, and being under the age of 26. The exponent of the resulting coefficient β can also be interpreted as the impact of bursary level on the expected share of the cohort that holds that characteristic. Errors are clustered at the Subject-Classification-Year level.

$$C_{idjt} = \beta B_{djt} + \alpha_d + \alpha_j + \alpha_t + \gamma TraineeTarget_{jt} + \epsilon_{idjt} \tag{2}$$

4.3 Individual-level outcomes

Secondly, I examine the impact of financial incentives on four measures of trainee retention. The main measure of interest is a whether the trainee ever appears as a teacher in the SWC post-training, as this is typically the stage with the largest rate of dropout. I also estimate (un)conditional retention, which I define as the probability of teaching in each year for up to four years after training has been completed. This measure is a dummy equal to one if we observe an individual in a teacher or senior school position (e.g. head teacher) in any school type in the SWC x years after their training. Conditional retention excludes those who never teach post-training from the sample, whereas unconditional retention includes the full sample of trainees and can be considered a combination of the conditional retention measure and appearing post-qualification. I lastly measure the probability of passing training.

I estimate these measures using equations 3, 4 and 6, where the error terms are distributed logit. I these for the entire sample and separately for STEM and non-STEM subjects. Notation $idjt$ denotes an individual i with undergraduate classification d, training in subject j at time t. Y_{idjt} is a dummy equal to one if the performance indicator is true. I include fixed effects for degree classification, year, subject, and training route r. Errors are clustered at the subject-classification-year level.

$$\begin{aligned} Y_{idjt} =& \beta B_{djt} + \alpha_d + \alpha_j + \alpha_r + \alpha_t + \\ & + \gamma_1 TraineeTarget_{jt} + \gamma_2 \widehat{WageGap}_{idjt} + \epsilon_{idjt} \end{aligned} \tag{3}$$

$$\begin{aligned} Y_{idjt} =& \beta B_{djt} + \alpha_d + \alpha_j + \alpha_r + \alpha_t \\ & + \gamma_1 TraineeTarget_{jt} + \gamma_2 \widehat{WageGap}_{idjt} + \\ & + \gamma_3 Characteristics_{idjt} + \epsilon_{idjt} \end{aligned} \tag{4}$$

$$\widehat{WageGap}_{idjt} = \hat{w}_{out,idjt} - \hat{w}_{teach,idjt} \tag{5}$$

The first regression estimates the overall effect of the bursary level on retention. In equation 4 I control for two key variables that may influence the setting of bursary levels and therefore threaten their exogeneity: trainee targets and the teaching wage gap. The wage gap is calculated as the predicted outside weekly wage for individual

i minus their predicted weekly wage as a secondary school teacher. Using samples of teachers from the SWC and workers employed in other graduate occupations from ASHE, I estimate wages as a function of region, time, characteristics and their interactions. The resulting coefficients are applied to my trainee sample to predict their wages. Additional details of how these wages are approximated can be found in appendix section A. The gap refers to the wage gap in an individual's training year. For the outcomes of (un)conditional measures of retention across teaching years, I additionally control for the wage differential in that given teaching year[19].

$$\begin{aligned} Y_{idjt} =& \beta B_{djt} + \alpha_d + \alpha_j + \alpha_r + \alpha_t \\ &+ \gamma_1 TraineeTarget_{jt} + \gamma_2 \widehat{WageGap}_{idjt} + \\ &+ \gamma_3 Characteristics_{idjt} + \epsilon_{idjt} \end{aligned} \tag{6}$$

The former regression measures the combined effects of observable and unobservable selection into training by controlling for course related fixed effects only (subject, degree classification, and year) as well as confounders. Equation 6 additionally controls for the personal characteristics of the individual (age, sex, ethnicity, region, undergraduate subject) and so only measures the impact of unobservable change the composition of trainees on retention. For example, perhaps a higher bursary attracts more individuals aged over 25 (who are more likely to quit) and less inherently motivated trainees. Regression 4 will identify both effects, whereas the second equation will isolate the sole impact of underlying motivation to teach. The first equation is useful from a policy perspective to assess whether their incentives were effective. However controlling for observable characteristics may be more relevant from an economic perspective, as it isolates the composition effects that admission procedures cannot easily target.

Whilst undergraduate classification is a personal characteristic and not a course related characteristic, I include it in the base specification as it is used to deter-

[19]In robustness checks, the addition of the current wage does not have an impact on the bursary level coefficient. In section 6 I also discuss how the use of different wage estimations impacts my results.

mine the bursary level (or 'treatment'). Using my previous example, whilst a higher bursary could attract older trainees with differing motivation to the same subject-classification group, it cannot in the same way attract trainees with higher classifications as these individuals are offered a separate bursary. The bursary an individual is offered instead depends on their classification. Classification is by construction positively correlated with both bursary level and also positively correlated with most retention outcomes. Failing to control for it in the base specification would lead to an upwards bias and over-estimate the efficacy of bursaries as financial incentives. The outside wage gap can also be considered both a key control variable and a personal characteristic. I therefore first control for trainee targets and then sequentially add outside wages and then other personal characteristics in separate regressions. The overall combined impact of the bursary on retention exists between the coefficient values of the first and second of these regressions.

5 Empirical Results

5.1 Cohort Characteristics

Table 2 shows the impacts of the average bursary level on trainee cohort size, where column 1 estimates equation 1. Overall, a bursary increase of £10k is associated with a significant 34 percent[20] increase in cohort size. This is in line with Worth's 2021 estimates for the impact of bursary size on trainee applications, suggesting that the increase in recruitment is driven by trainee supply, rather than changes in demand. Overall, STEM graduates are more responsive to the financial incentive than non-STEM graduates, as cohort sizes increase by 51 and 23 percent respectively.

I also examine the impact on cohort size 3 years after training to see if offering higher bursaries has a long term impact on the stock of teachers. Overall, there is an insignificant increase in cohort size of 14 percent, much smaller than the initial 34 percent increase, which suggests higher attrition rates for higher bursary awards.

[20]Calculated as the exponent of the regression coefficient

In STEM, a bursary uplift of £10k still results in a statistically significant increase in overall cohort size, but the magnitude has fallen to 42 percent. In non-STEM subjects, there is a small and insignificant positive effect of the bursary on the number of teachers employed 3 years later.

The results produced by estimating equation 2 can be seen in table 3. Across all subjects, the largest significant effect of an increase in bursary is an increase in the share of over-25s, or later career trainees. The share of early career trainees decreases by approximately 5 percentage points following a £10k bursary uplift. The magnitudes are similar in both the STEM and non-STEM subject groups. As bursaries increase, the training year becomes more competitive relative to an individual's outside wage and attracts more later-career individuals who have more experience and therefore a larger wage in their next-best occupation. There is a negligible impact of bursaries on the gender and ethnicity ratios of trainees. The only significant result is observed for STEM subjects, where a £10k increase significantly increases the share of women by approximately 2 percentage points. This is an encouraging result: trainee cohort sizes increase without compromising on diversity within the classroom.

5.2 Retention

In the previous section I established that financial incentives to train had an impact on both the size and composition of trainee cohorts. Next, I explore how trainees attracted by different bursary levels may be negatively selected by estimating the impact of bursaries on individual-level retention outcomes. I present the results of regressions 3, 4 and 6 in tables 4 to 6. Below, I discuss the marginal impact of a £10k bursary uplift (approximately one standard deviation) on the probability of retention, where the margin is estimated using the average probabilities of the population.

Table 4 shows that a £10k raise in funding results in a significant 1.8 percentage point decrease in the probability of appearing post-qualification. The magnitude is unchanged wen controlling for trainee targets and decreases slightly by 0.2 percentage points when controlling for the wage gap on entry into training. The negative impact of a £10k bursary raise increases from 1.6 to 1.8 when controlling for personal characteristics, suggesting that this composition effect is predominantly stemming from unobservable 'motivation' rather than a change in in observable characteristics. Given that the wage may also act as a noisy proxy for characteristics such as age, we can conclude that the observable composition effect accounts for at most 10% of the total negative post-training-retention impact of the bursary[21].

Table 5 disaggregates these effects for STEM and non-STEM trainees. It shows that the negative retention effects are much larger in magnitude and only significant for the STEM sample, who are 2.8 percentage points less likely to appear as a teacher following a £10k increase in bursary. Controlling for trainee targets reduces the magnitude by 0.6, but controlling for characteristics and outside wages in this sub-sample increases the magnitude and leaves the impact of the unobservable negative retention effect at 3 percentage points. Note that controlling for outside wages increases the magnitude of the effect by 0.9, but subsequently controlling for characteristics decreases the effect by only 0.1 and reduces the impact of wages on retention to zero. Again, the wage gap could be proxying for personal characteristics. This implies that for the STEM sample, a £10k bursary raise either generates no observable composition effect, or generates a positive observable composition effect that increases retention by up to 0.7 percent with a £10k bursary increase. The latter is more likely. Referring to tables 3 and 17, higher STEM bursaries attract more female trainees, who have smaller outside wage gaps and are significantly more likely to appear post-training, even when controlling for wages.

[21]Note that controlling for outside wages decreases the coefficient on wages decreases to almost zero when controlling for personal characteristics. This suggests that for this outcome, the wage itself does not convey any additional useful variation that the characteristics themselves do not generate (despite the outside wage being a function of additional variables). The wage can act as a noisy proxy for age or sex. Therefore the 10% drop in coefficient from column (2) to column (3) in table 4 could be attributed to sex or age. Since this cannot be disaggregated from the confounding effect of outside wages, we can summarise that the negative composition effect is between 0 and 10% of the overall effect.

I also estimate the impact of financial support on the likelihood of appearing as a teacher at each year post-training in table 6. I construct a conditional measure where the sample is comprised of those who appear as a qualified teacher at any time after qualification. The coefficient can therefore be interpreted as the additional retention effect separate from post-training dropout. I present unconditional retention results in the appendix table 15 for a combined measure. For each year, I measure the total impact of bursaries in column (1) and only the effect stemming from unobservable changes in trainee composition (by including personal characteristic controls) in column (2).

Overall, there are generally negative but insignificant effects of the bursary on appearing in each year. The sign of the marginal effects suggests that on the intensive margin, trainees who become teachers may be present in schools for fewer years. I find an insignificant 0.5 percent decrease of being present as a teacher in year one for the full sample, which increases in magnitude to 1 percent and becomes significant when controlling for personal characteristics. This suggests that bursaries lead to negative selection in terms of unobservable 'motivation', but positive selection effects in observable characteristics. The change in the probability of STEM teachers appearing the year after training following a £10,000 uplift when (not) controlling for characteristics is significant at (-1.3) -1.9 percentage points. Wages are less likely to proxy for the impact of other characteristics in this case as the wage gap coefficients generally increase in magnitude and become significant when controlling for characteristics.

One remaining feature of interest in table 6c is that STEM trainees with higher bursaries are increasingly likely to appear as a teacher with each additional year. By raising the bursary by £10k, the probability that a STEM teacher is present in year 4 increases by around 4 percentage points. Intuitively this would suggest that for those who do eventually enter teaching, STEM teachers with higher bursary levels are less likely to enter teaching straight away.

Trainee targets have markedly different impacts on retention outcomes for the

STEM and non-STEM samples. A 100 unit rise in the trainee target is significantly associated with a 0.5 percentage point increase in the probability of teaching post-training for STEM subjects. However, the same change in targets decreases retention for the non-STEM sample by 0.5 percent. There are two explanations as to why targets could have either a positive or negative effect. For example, a large number of teacher vacancies suggest it may be easier to find a job that matches a trainee's preferences and increase the number of trainees that become teachers after qualifying. However higher targets also imply more under-staffing in those subjects, and so trainees experience larger workloads and more stress during their school-based training, making them less likely to apply for teaching roles post-qualification. If non-STEM subjects feature more qualitative assessments that are time consuming for teachers to mark, then under-staffing may generate a larger penalty in non-STEM subjects that cause this channel to dominate.

5.3 Local Labour Market Conditions

Local wages are an important factor to consider when evaluating teacher retention for two reasons. Firstly, the outside options of potential teachers impact their decision to enrol into teacher training. Secondly, the Department for Education may also take this into account when designing bursary incentives. Figures 6 and 7 summarise the age-earnings profile of teachers versus university graduates using different data sources. Note that teachers do not necessarily earn less than non-teachers, particularly in the early phases of their career. Therefore, it may not be surprising that the teaching:outside wage ratio for trainees (individuals that have selected into teaching) in their training year is 1.11 (but not statistically significant from one)[22]. As individuals age, the outside wage becomes slightly more competitive on average, but is also more variable.

[22]Whilst this result is a departure from some of the literature measuring average teacher pay gaps, my sample is unique as it refers to teachers at the very start of their career, rather than focusing on the outside options of existing and more senior staff. A recent study of the wage outcomes of individuals who were marginally rejected from teaching also finds a teaching wage premium (Tsao, 2024).

Figure 8 represents the regional disparity in graduate wage outcomes. Teacher salaries are more uniformly distributed due to national pay scales, meaning that the relative return to teaching mostly depends on the outside wage. Teaching is therefore most attractive in northern regions where outside wages are relatively lower. Note teachers within and around London also are subject to an additional pay supplement, so whilst teaching is least competitive in this area, the ratio is not as low as it otherwise would be.

I control for the wage gap upon entry into training as defined in equation 5 on the probability of appearing post-qualification in the fourth column of tables 5 and 5. If weekly teaching wages become £100 more competitive (around one standard-deviation) during the training year, trainees are 4.5 percentage points less likely to become a teacher post-training[23]. This effect becomes small and insignificant from zero when controlling for characteristics, suggesting that the wage effect was driven by differences in wages between groups, and not within groups. It is also of a similar magnitude for STEM and non-STEM groups.

The wage gap during an individual's trainee year has a significant positive effect on teaching in any given year, conditional on becoming a teacher post-training (table 6). A £100 larger wage gap during training increases the probability of teaching in any given year by between 2 and 3 percentage points. This means that those recruited into teacher training when it is less financially competitive are generally less likely to become a teacher, but those that do teach more on the intensive margin. Note that this effect is significant only when controlling for personal characteristics, implying that this wage effect is instead driven by differences in wages within groups.

Outside wages are a useful source of variation to examine how bursaries could have differential impacts depending on the relative competitiveness of teaching. In table 7, I investigate the regional impact of the bursary level on appearing post-qualification by running regression 6 but excluding wage gaps and replacing the bursary with a full set of interaction terms of the bursary level with region. Seven out of ten regions

[23]Whilst the bursary coefficient is robust to the definition of outside wage, the wage gap coefficient varies in magnitude depending on the wage measure. More details in the appendix

are significant, and marginal effects range from being insignificant from zero in Outer London and the North West to -4.7 percentage points in the North East. With the exception of Inner London, the regions with the largest impact are more rural regions (North East, East of England, South West), whereas those with the least impact are regions housing major cities (Outer London, North West, West Midlands).

In a separate set of regressions, I interact the wage gap and the bursary level. The bursary level coefficient can now be interpreted as the impact of a bursary increase on an individual with a wage gap equal to zero. The interaction term is the additional change in retention that an increase in the bursary level of £10k causes for an individual with a £100 greater wage gap. The results show that for individuals with no wage gap, an increase in the bursary level reduces their likelihood of appearing post-qualification by 2 percentage points. Consistent with table 4, the larger the wage gap upon entry into training, the less likely a trainee is to appear post-qualification. However, there is a positive significant additional marginal effect of the bursary level as the wage gap increases. In other words, the negative retention effect of the bursary is largest for those with small wage gaps. The impact of offering a higher bursary can become positive for those who have a weekly wage gap of roughly £250 or more, however for STEM graduates this value is much larger at roughly £570. In section 7 I introduce a model with rational agents that can explain the main results of the empirical section. The model, however, cannot explain why those with the highest outside options are the least negatively impacted by the bursary level.

6 Robustness

6.0.1 Cohort size

The selected specification for cohort size regresses the average bursary on the total cohort size at the subject level. However, the bursary varies at the subject-classification level. To further disaggregate this impact, I estimate the semi-elasticity of cohort size at the subject-classification level. This allows the bursary level B_{djt} to be fixed for each observation, rather than an average across a trainee cohort. I run equation 7 separately for each degree classification d, with errors clustered at the subject level, where N_{djt} is the number of trainees enrolled with classification level d in that subject-year. I additionally include year and subject fixed effects. All regressions are run separately for STEM subjects, non-STEM subjects, and all subjects combined. I weight subject-level regressions by the average cohort size across all years.

$$log(N_{djt}) = \beta_d B_{djt} + \alpha_{dj} + \alpha_{dt} + \gamma_d TraineeTarget_{jt} + \epsilon_{djt} \tag{7}$$

The results of equation 7, which disaggregates the impact of a bursary increase at the subject-classification level, are shown in table 16[24]. Since the confidence intervals for each estimate are large, the impacts of the financial incentive across degree classifications on recruitment are not statistically significant from one-another. The same conclusion can be reached in tables 16b and 16c, which run the same analysis for the non-STEM and STEM samples respectively. However, the magnitudes are all roughly consistent with the average impact of the bursary level calculated in table 2.

Using an average bursary level as the independent variable in regression 1 when estimating the overall impact on cohort size could lead to biased results; For example, if the bursary offered to graduates with a lower second (2:2) increases, let us assume that this increases the number of 2:2 graduates enrolled on to the course. Keeping all else equal, this could also reduce the average bursary level of that cohort if the existing trainees had higher classifications and larger bursaries. This example would demonstrate recruitment increasing whilst the average bursary level fell, despite an underlying positive correlation between bursaries and recruitment. I explore this bias by running regression 7 with a pooled sample of all subject-classes to estimate the average impact of the bursary across all classifications. The results, which will be included in a later draft, suggest that the impact remains significantly positive

[24]Note that third class degrees are not included due to insufficient variation in the bursary level.

for total trainee cohort size, but the magnitude of estimates are reduced.

I also address cross-subject substitution. Prospective trainees may prefer to teach the subject that they were chiefly trained in during their undergraduate degree. However if the bursary of a related subject is more generous, this would prompt them to apply to train as a teacher in this subject instead. Econometrically, a positive coefficient of the bursary on trainee cohort size may be generated by the redistribution of existing trainees, rather than attracting new trainees. I therefore control for the average bursary level of competing subjects[25] in the pooled estimation of equation 7. These results will be included in detail in a later draft, but estimates generally become larger and more precise.

I perform two remaining robustness checks. Firstly, I control for the average outside wage gap of trainees upon entry into training within the specified cohort. Coefficients and p-values are largely unaffected. Secondly, I limit the sample of subject-years included in the regressions of trainee cohort size to those who are still observed in the sample three years later. This allows a direct comparison between the change in trainees recruited and the change in teachers employed three years later. The coefficients on average decrease by a third, and as a result the difference in impact on cohort size during training versus teaching is less pronounced. However, the sample sizes of this wild-bootstrap estimation generate large confidence intervals which make attrition difficult to measure by comparing cohort sizes. The regression results on retention are more informative for this purpose.

6.0.2 Cohort Characteristics

Table 3 showed that the share of trainees under 26 and the share of males in STEM trainee cohorts falls with an increase in bursaries. In additional analysis, I examine whether the number of trainees under 26 and the number of STEM male trainees also decreases using equations 1 and 7, where N_{jt} is the count of individuals that possess

[25] Competing subjects are all subjects in that year-classification group that share the same STEM status. E.g. for first-class physics, I estimate the average bursary offered for first-class trainees in all other STEM subjects.

the relevant characteristic. I find that all sub-groups increase in size following a rise in bursary, and that only the relative shares are significantly effective.

I additionally use alternative specifications to estimate the impact on cohort characteristics. I use a linear regression where the left hand side variable is the share of individuals in the cohort possessing the characteristic, and each observation represents a given year-classification-subject cohort. I use a wild bootstrap where errors are clustered at the subject level, and observations are weighted by average cohort count across years. I find that the magnitude of effects are unchanged, all under-26 coefficients remain at the same level of significance, but the impact on being female is no longer significant for any category. Finally, I add a control for undergraduate classification in the original specification. Results are of a similar magnitude and significance is unchanged, with the exception of the share of non-white individuals. The magnitude of the effect roughly doubles to -0.83 for the non-STEM sample and becomes significant at the 10 percent level.

6.0.3 Retention

One concern when estimating the retention results is that not all trainees are awarded bursaries. Trainees may also apply for a 'school direct salaried' route, in which they are instead paid a taxable salary by a school. Due to the format of the data, I am able to identify school-direct trainees but in some years I am unable to identify whether these trainees are salaried or not. For robustness, I run the same regressions but excluding all school-direct trainees. The results reported in tables 4, 5 and 6 become stronger: all significant coefficients are between 5 percent smaller and 50 percent larger. However, it is notable the probability of appearing post-qualification doubles in size and is significantly negative in column 1 for non-STEM trainees in table 5.

Given that outside wages are calculated as a function of characteristics of the individual, the wage gap will be correlated with characteristics (although not co-linear due to additional fixed-effects and interaction terms). I therefore also compare how the coefficient on bursary level changes when controlling for characteristics before controlling for the outside wage gap. The effect continues to be minimal: the majority of the effects observed for appearing post-qualification STEM from unobservable composition effects.

I also explore the impact of adding a dummy equal to one when a trainee is likely to experience a drop in income when moving from training into full-time teaching on the probability of appearing post-qualification. Whilst this has a negligible effect on the bursary coefficient and is itself insignificant for the pooled and STEM samples, it does have an effect for the non-STEM population. The bursary coefficient becomes significantly negative and the wage drop coefficient becomes significantly positive. This result seems counter-intuitive, as we would assume that an income drop would make an individual more likely to exit the profession. It is instead likely that this wage drop term is acting as a proxy for non-linearity in the effect of bursary level, since a wage drop is most likely at higher bursary levels. Indeed, the effect is no longer significant when controlling for the square of the bursary[26].

7 Modelling Occupational Choice

I develop a simple model to interpret the results presented in section 5 and explore whether they can be replicated with rational agents. This model is based on the standard occupational choice framework as developed by Keane and Wolpin (1997), and builds on models of teacher labour supply presented in Stinebrickner (2001*a*) and Stinebrickner (2001*b*). It is a dynamic model of individual behaviour where in each period individuals maximise lifetime utility by selecting one of two occupations, but must undergo training before they are able to move into the teaching occupation.

Each period, an individual chooses to either work as a teacher (T), or work in

[26]Given the potential of non-linear effects, in a future draft I will also present results on whether differences between STEM and non-STEM samples are driven by differences in the average value of bursaries awarded to these groups. A future draft will also feature the results of tables 4 and 5 where observations are kept constant across columns 1 to 4.

another occupation (O). Individual i at time t derives a utility from occupation j which is a linear combination of wage w_{it}^{j} and non-pecuniary benefit m_{i}^{j}.Given that there are only two occupations, I normalise m_i to be the relative non-pecuniary benefit of teaching, otherwise referred to as motivation. I assume that motivation is randomly drawn from a normal distribution, and known to the individual. Utility is denoted below, where d_{it}^{j} is a dummy variable equal to one when teaching is chosen. Each individual lives for a fixed number of periods.

$$U_{it}^{j} = w_{it}^{j} + d_{it}^{j} m_i \tag{8}$$

Each period, an individual receives a wage offer for occupation j which is a function of qualifications q_i, experience e_{it}^{j} endogenously accrued within the occupation, and a random AR(1) error component ϵ_{it}^{j}. The errors follow an AR(1) progression in order for a positive wage shock to be maintained across time periods. Although teacher pay is banded, a random component is included because a teacher's pay varies through band sub-steps, speed of promotion, and additional responsibilities. Schools also have some autonomy on deciding which band a teacher is classed under[27]. For simplicity, labour demand is not modelled and the parameters of the wage function are considered as exogenous.

$$w_{it}^{j} = f_j(q_i, e_{it}^{j}) + \epsilon_{it}^{j} \tag{9}$$

An individual selects their occupation for that period in order to maximise their expected lifetime utility, according to the value function below:

$$V_{it}^{j} = U_{it}^{j} + \beta E_{max}\left[\{V_{it+1} | d_{it}^{j}\}\right] \tag{10}$$

$$d_{it} = \begin{cases} 1 & \text{if } V_{it}^{T} > V_{it}^{O} \\ 0 & \text{if } V_{it}^{T} < V_{it}^{O} \end{cases} \tag{11}$$

[27] Academies in particular are not required to adhere to national pay structures

Individual i will choose to teach in any year when the present discounted value of teaching is greater than the present discounted value of the other occupation. Rearranging this inequality, each individual trades off the relative wage gap of teaching versus other occupations and their own motivation for teaching. The lower the expected lifetime wage gap, the lower the minimum motivation $\underline{m_i}$ required to select into teaching. For simplicity, I assume no switching costs. The key friction is instead that individuals must undergo a training year that is only funded by the bursary before they can be paid a teacher's wage. The bursary offered to the individual could be as small as zero, which implies that the training year has a high opportunity cost, or larger than the first year pay of either occupation[28].

Teacher wages vary less with respect to undergraduate classification compared to other occupations, and better qualifications are rewarded more in occupations other than teaching[29]. Therefore, an individual with better qualifications will experience a higher wage gap when selecting into teaching. As a result, the minimum required motivation $\underline{m_i}$ for a highly qualified individual is higher than the minimum required motivation for an individual with lower qualifications, all else being equal.

7.1 Model impact of Bursaries

Teacher training is considered as the first year of teaching. Therefore, the bursary amount $B_{i,t}$ is replaces the first year teaching wage. I simulate the occupation choices for 30 periods under the model for a population of individuals with different outside wages and outside wage growth. For all individuals, the initial average teaching wage and teaching wage growth is the same. Each individual is randomly assigned their motivation of teaching from a normal distribution, and their wage for both occupations vary subject to an AR(1) sequence of wage shocks. I then observe how the number and characteristics of individuals who decide to train in year t=1 changes

[28] For additional simplicity, I do not incorporate liquidity constraints or a negative bursary offer. Whilst in reality some trainees may experience liquidity constraints, maintenance loans of a limited amount are available.

[29] Based on regressions of wage against undergraduate classification in the SWC data and LFS data.

when the bursary level is varied. I summarise the key results of this exercise below, and more details can be found in the appendix B.

In this model, the share of individuals in one particular occupation eventually stabilises. This is because as long as each occupation experiences positive wage growth, it is more profitable to the individual specialise and remain in one occupation and continue to benefit from the return to the occupation-specific experience they have accrued. Switching between occupations is more likely in earlier time periods when experience is lowest and wage gaps between teaching and the 'other' occupation are smallest. The earlier the time period, the more likely it is that an exogenous wage shock can change the lower-utility occupation choice into the dominant choice. Therefore, we can imagine that the bursary can impact an individual's lifetime choices in two ways: Firstly, raising the bursary could permanently pull more individuals into teaching who otherwise had a very weak preference for the outside occupation. In other words, a higher bursary attracts individuals whose expected lifetime payoff from the outside occupation was only marginally larger than that of teaching. We can informally name this group the 'converts'. Secondly, a larger bursary can attract people into teaching temporarily, particularly if the bursary is worth more than the second year wage of teaching. Individuals will train whilst the teaching payoff is generous and then move to the alternative occupation to accrue experience in their long-term most profitable occupation. In other words, the front-loaded payoff of the bursary is large enough to compensate the individual for the lifetime income lost by giving up one year of experience in the alternative occupation. I call this second group 'dropouts'. In figure 13, we see that the shares of individuals in teaching stabilises for all bursary levels by around year 10. For bursary levels above £30,000 (the starting wage of teaching), the share employed in teaching is initially high and falls over time, whereas for bursary levels below the starting wage of teaching the share increases over time.

7.1.1 Result 1: Increase in the number of trainees

An increase in bursary increases the relative financial payoff of teaching and reduces the minimum motivation threshold required to select into teaching. This induces marginal applicants to apply for teaching. Given that training targets are largely unmet, the increase in supply will lead to an increase in the number of trainees recruited. This can be seen in figure 9, and aligns with the results in section 5. Figure 13 also shows that a larger bursary attracts a larger share of the population to train in year t=1. Figure 14 attempts to replicate the results in the empirical analysis by measuring the share of the population who train in the first year, and following this training cohort across time. A higher bursary consistently increases the number of trainees.

Note that the in figure 13, the bursary level increases in £5,000 intervals but the increase in trainees is not uniform. This may happen due to randomness in the outside wage offers of individuals, but it can also occur because of the distribution of motivation in the population. In figure 9, we can see that the number of applicants attracted by an additional £5,000 in bursary is initially small when starting at the right-most point of the graph, is then increasing till the motivation threshold is at the central average level of motivation, and then decreasing on the left hand side of the distribution. By observing the gap between each line in figure 14, we see that the marginal impact of the bursary in the model simulations is largest for the median bursary amounts.

7.1.2 Result 2: Increase in the long-term number of teachers

As previously discussed, a raise in bursary level attracts both 'converts' and 'dropouts' into the training programme. Converts are individuals who prior to an increase in bursary only marginally preferred their outside option. The increase in the bursary is sufficient to make teaching the dominant occupation and they will remain in teaching and benefit from the return to their occupation-specific experience. The empirical results confirm that even after four years, teacher cohort size is larger for higher bursary levels[30]. Figures 14 and 15 demonstrate this using the model simulations: the share of the population engaged in teaching is higher for larger bursary levels any number of years later.

7.1.3 Result 3: Composition Effects based on Qualifications

Given that the teaching wage gap varies with personal qualifications, there are separate motivation thresholds for entering teaching for those with different undergraduate classifications. For example, those with a first will have a higher threshold than those with an upper second since their outside wages are higher (Britton et al., 2022). When the bursary increases, the number of marginal applicants from each of these groups depends on the relative population density at that part of the motivation distribution. Figure 10 demonstrates the case where offering a uniform bursary increase to all applicants attracts mostly upper-second class candidates. The empirical results show that the higher bursaries were most effective at attracting candidates with a first-class degree. The simulation results show that those with higher outside wages (greater than £30,000) are less likely to train than those with lower wages (less than £30,000) at any bursary level. Figure 15 disaggregates the impact of the bursary on training cohort size for these two groups. Note that the axis for the high and low wage groups are separate so the difference in the share of the population training in year one is around 74 percent.

STEM and Non-STEM results can also differ when outside wages for the alternative occupation differ between those who apply to teach STEM and non-STEM subjects. This is plausible given that STEM subjects are mostly taught by those who studied STEM subjects themselves at the undergraduate level and that this group has higher average wages in the labour market (Britton et al., 2022). Section 6 established that in terms of magnitude, STEM bursaries were most effective in attracting those with first class degrees whereas non-STEM bursaries mostly attracted those with lower-second class degrees[31]. The model simulations can also reflect this

[30] Though not always significant.

[31] Note that these coefficients were significant, but not statistically significant from the estimates for other clas-

by assuming that STEM graduates have a higher proportion of trainees with high starting wages.

One additional assumption is required to avoid contradictory results. As seen in figure 11, attracting predominantly highly qualified trainees requires motivation thresholds to sit to the left of the mean on the population distribution. Conversely, in order to attract candidates with lower classifications motivation thresholds must sit to the right of the mean. Combining this with the empirical results would suggest that the motivation threshold for a lower-second non-STEM graduate would be above the threshold for a first class STEM graduate. This implies that the outside wages for a lower-second non-STEM graduate would also be higher than those of a first class STEM graduate. As this is not the case in the data, my results would require the additional assumption that the distribution of motivation in the STEM graduate population is to the right of the distribution for non-STEM graduates (as seen in figure 12). Those that select into studying STEM courses are on average more motivated (or have a higher relative non-pecuniary valuation of teaching) than those that choose to study non-STEM subjects. Note that this does not necessarily mean that more STEM graduates will select into teaching, as STEM graduates also have higher wage offers outside teaching.

7.1.4 Result 4: Increase in attrition between training and teaching

The empirical results found that higher bursary levels reduced the likelihood of appearing post-qualification, even when controlling for observable characteristics. The model suggests this is due to higher bursaries attracting more 'dropouts' into training: those that seek a higher paid training year without the intention to remain as teachers. As previously discussed, this occurs when the bursary is large enough to compensate the individual for the lifetime income lost by forgoing one year's experience, but teaching is still not competitive enough compared to their outside option to induce the individual to remain in teaching after training.

sifications.

Figure 16 replicates the survival rates in the data shown in figure 4 using simulations and shows the share of the trainee cohort that remain teaching in each subsequent year. The simulations show that the biggest drop in retention occurs in the first year post-training, and that higher bursaries lead to higher attrition. One clear distinction is the magnitude of attrition. The empirical data shows that post-training retention is between 65 to 75 percent. However the simulations report 95 to 99 percent retention. In the final paragraph of this section I discuss some alternative models that could predict a larger rate of attrition. Table 10 shows that whilst moving form the lowest to the highest simulated bursary increases trainee cohort size by 14%, the number of converts only increases by 10%. Meanwhile the number of dropouts has increased by 322%. Overall, the share of trainees that are converts decreases by 3.3% over this interval, from 98.2 to 94.9%.

7.1.5 Result 5: Reduced marginal and average motivation

I have established that as the bursary level increases, the threshold level of minimum motivation required to select teaching falls. This is because the financial payoff of teaching has increased whilst the compensation of the alternative occupation is unchanged. The marginal individuals who are persuaded to teach by the increase in bursary have lower motivation, and will also decrease the average motivation of the pool of trainees. Table 9 confirms that the average motivation for all trainees, as well as for converts only, is lower at higher bursary amounts.

7.1.6 Result 6: Retention varies by Outside Option

Retention of trainees depends on their likelihood of being a never-takes versus a convert. In order for someone with a high expected outside wage to permanently select teaching, they would require at least one or a combination of the following: a very low wage growth for their alternative occupation, a very high motivation for teaching, a large negative wage shock to their alternative wage, or a large positive wage shock to their teaching wage. In the model simulations, individuals are equally distributed across 10 outside wage values and 5 wage growth values, and wage shocks and motivation are randomly allocated from a normal distribution. Table 11 shows that in these simulations, the higher the outside wage value the higher the share of trainees that are dropouts. This increases from 0.4% when the bursary is zero, to 24.4% at a bursary of £55,000[32]. In line with previous results, the average motivation of trainees with large outside wages is also higher, and the number of trainees is fewer.

The model backs up the key differences between STEM and non-STEM attrition in the empirical results. Figure 15 displays the probability of retention at each bursary level for each value of expected outside wage. By comparing the gap between two neighbouring bursary levels as the wage level increases, we can see that the effect of raising the bursary level is more negative for higher outside wages. In section 5, we observed that bursaries had the largest negative impact on post-training retention for STEM graduates, but a smaller and non-significant effect on STEM graduates. The model can explain this as long as STEM graduates have higher outside wages, or a higher variance in outside wages.

There are two empirical results that are in direct contradiction to the model's predictions. Firstly, the model suggests that controlling for bursary levels, those with higher classifications in their undergraduate degrees are generally less likely to remain in teaching due to higher outside wages[33]. Whilst this is in line with most findings in the literature[34], this is not reflected in the empirical results (see figure 17). In the results section, I also established that the negative retention effect of the bursary was largest for those with smaller wage gaps. The converse is true in the model: the negative retention effect is strongest for those with large wage gaps. This can be seen in figure 15 as the gap between two neighbouring bursary levels increases as the wage level increases.

[32]Whilst the largest bursary actually offered to trainees was £30,000, this is equivalent to an income of just under £43,000 following student loan repayments and tax deductions.

[33]This can be seen by comparing the retention rate in figure 15 at the same bursary level across different wage levels. A higher outside wage worsens retention and makes it less likely that a trainee will appear as a teacher.

[34]For example, see Podgursky et al. (2004), Stinebrickner (1998)

7.1.7 Implications and Limitations of the model

Overall, the model predicts that raising the bursary is effective in recruiting more teachers in both the long term and short term. However, the higher the bursary the higher the rate of 'dropouts'. A higher bursary will also lower the average motivation of trainees, and so marginal trainees will be more sensitive to future teacher pay cuts or raises. The magnitude of the recruitment and retention effects depends on the outside wages and underlying distribution of motivation of the potential trainees. In light of these findings, in section 8 I evaluate whether bursaries are cost effective and discuss which sub-groups are most effectively targeted by this policy tool.

The model is successful in generating both an increase in trainee recruitment and trainee-to-teacher attrition, in addition to explaining why there are markedly different results observed for different sub-groups of trainees. However the model does not explain the large magnitude of the post-training attrition rates observed in reality, nor does it explain the second order interaction effects between outside wages and the bursary level. Discount rates are important in determining the impact of the bursary on both recruitment and retention, however an improvement to the model could incorporate myopic agents to generate larger attrition rates. Alternatively, a model that combines risk averse agents and uncertainty over teaching motivation may also provide further insight into the observed behaviour of individuals under this policy, as the bursary reduces the inherent risk associated with learning your motivation for teaching in the training year.

This model does not establish the optimal wage-setting behaviour of the firm (or in this case government). However, the results imply that a front-loaded financial incentive may not be optimal if the objective is to maximise long-run employment, subject to budget constraints. A growing literature on wage-tenure contracts suggests that although a high bursary may be effective in inducing more job-to-job transitions into teaching, the design of the wage-tenure profile in the early career stages, and the extent to which pay is back-loaded, matters for retention [35]. The wage schedule for

[35] See Burdett and Coles (2003), Shi (2009), and Bagger et al. (2014) for model examples.

teachers remains relatively rigid, and the effectiveness of customised bursaries could remain limited unless wage offers and return to tenure can also be personalised.

8 Policy Evaluation

I evaluate the cost-effectiveness of bursaries compared to a menu of other policies in increasing the supply of teachers. Through a back-of the envelope calculation that incorporates elasticity estimates of UK teachers from Sims and Benhenda (2022) and Worth et al. (2022), I estimate the cost of attaining 5,000 additional teaching years for three policies: Raising the teacher training bursary, increasing early-career pay (the pay of teachers in their first 2 years of teaching), and raising all secondary school teacher pay. Each policy has effects on the recruitment and the retention of teachers, and I consider wages, bursary costs, and training costs. For each policy I evaluate the additional cost of the policy and the additional teaching years gained. I use these figures to generate the average cost per additional teaching year. Additional details and assumptions made are included in the appendix.

Table 12 presents these results. I find that the cost of raising an additional 5,000 teaching years would require a bursary raise of around £6,500 and would cost £58,000 per additional teaching year. This is just over the average cost of raising all pay: which would require a 0.7% pay increase and cost £57,700 per additional teaching year. The most cost-effective tool is raising early-career pay. Raising early career pay by 2.5% for one teacher cohort would result in 5,000 additional teaching years at a cost of £37,000 per year.

The estimates show that bursaries are roughly just as cost effective as raining pay for all teachers. This may be surprising since a pay rise is targeted at the entire teaching workforce, despite attrition being highest in the first five years of teaching. Pay rises are more expensive for experienced teachers with higher pay, who have lower attrition rates to begin with. Alternatively, bursaries are effective in recruiting more individuals into teaching, but come with the additional disadvantage of lower post-training retention and higher fixed costs. A pay rise for early career trainees is cheaper than a general pay rise because it is targeted at teaching cohorts during the lowest retention years, in addition to their pay being lower. It also avoids the negative post-training attrition effect observed with raising bursaries because the financial incentive is delayed until years one and two of teaching. Raising pay for early-career teachers is unlikely to attract such 'dropouts' as long as pay growth remains positive for the cohort after the policy window ends.

One additional benefit of bursaries is they can be easily targeted at specific subgroups of trainees, whereas varying teacher pay by subject is more difficult to implement politically. Table 13 demonstrates this heterogeneity by estimating the cost to increase the number of total teaching years by 10 percent for all trainees and for STEM and non-STEM[36]. The average costs per teaching year are lower for STEM trainees (£56,000) than Non-STEM (£59,000). STEM subjects are more cost-effective as trainee recruitment is more responsive to financial incentives, however the gap between these costs is narrow because STEM trainees are already offered larger bursaries than STEM trainees, resulting in a larger fixed cost. Note that bursaries are particularly cost-effective when applied to smaller groups: for example, the average cost of a teaching-year for first-class maths trainees is £50,000.

There are three key considerations to be made when targeting bursary increases to raise the long run size of teacher cohorts. Firstly, when increasing the bursary offered to a specific degree classification group, post-qualification attrition will increase. The cost-effectiveness of doing so will depend on the current bursary level, the cost of training and the effectiveness of trainee teachers compared to experienced teachers. Secondly, bursaries only significantly raised cohort size for certain degree classification groups. Attracting the highest number of trainees for the lowest possi-

[36]This measure is used for comparison since cohort sizes vary by subject, and so gaining 5,000 trainees is more difficult in STEM than Non-STEM. Because the average cost of a teaching year is increasing in the number of trainees recruited (in other words, there is a diminishing marginal return to raising the bursary), this measure is more comparable. Bursaries face diminishing marginal returns because as retention rates worsen, trainees spend on average less time in teaching. This means that the fixed costs of the bursary and training become a larger share of total costs. Total costs increase at a faster rate than total teaching years.

ble increase in bursary level requires an understanding of the underlying distribution of motivation and outside wages. In the case of STEM candidates, the population of marginal applicants is largest for first-class degree holders whereas most non-STEM marginal applicants hold a lower-second degree. At the margin, attracting the same number of first-class degree holders requires a larger non-STEM bursary than is currently offered. Lastly, raising the bursary level may trade off the unobservable non-pecuniary motivation of trainees in favour of financial motivation. This will affect the elasticity of supply of the teacher workforce and so policy-makers should consider this in future pay-setting decisions.

All three of these policies are most effective when used in tandem with each other. Whilst raising early-career pay is the cheapest policy, it is limited in how much it can be raised before it becomes larger than the expected pay of a third-year teacher. Bursaries are a cost-effective tool to target particularly low-staffed subjects and can increase and decrease between years in a way that pay cannot. However, bursaries also are subject to diminishing returns[37]. If bursaries are used as the sole policy in recruiting a large number of teachers, the average cost per teaching-year will increase above the average cost of raising teacher pay. Lastly, there is a clear experience trade-off when policy focuses on attracting new teachers versus retaining experienced ones. Employing a sufficient number of teachers is a first-order concern, but consideration for teacher quality is also essential to maintain teaching standards in the long-term.

9 Conclusion

This study evaluates the impact of UK teacher training bursaries, a financial incentive, on the composition of trainee teachers and the resulting impact on workforce retention. I found that higher bursaries increase trainee cohort size and that teacher cohort size remains slightly higher three years later. However, an increase in the

[37] See previous footnote.

bursary level reduces the probability that an individual will appear as a teacher post-training. This negative retention effect is primarily driven by unobserved motivation, as the coefficient of interest is relatively constant when controlling for outside wages and personal characteristics. The effect is strongest for STEM graduates. Non-STEM graduates have smaller recruitment effects and no significant negative retention effects.

By developing a simple model of occupational choice to explain my results, I find that raising bursary levels increases the long term number of teachers by attracting marginal individuals into teaching with a lower non-pecuniary valuation of teaching, or 'motivation'. A higher bursary level also increases the number and share of 'dropouts' in the trainee cohort: Individuals who are attracted to train by the financial incentive, but who move to their alternative occupation directly afterwards. This increases the post-training attrition rate.

Through a cost-benefit analysis, I found that bursaries are just as cost-effective as raising teacher pay, and cost less per teaching-year when targeted at the most receptive sub-groups of trainees (e.g. STEM first class trainees). However even at their most effective, bursaries are still more costly per teaching year than raising early career teacher pay. Bursaries have the additional benefit of flexibility compared to pay rises as they can be easily varied across subjects, years, and personal qualifications. Policy makers should also consider the trade-off between recruiting new teachers versus retaining experienced teachers.

My findings contribute to the teacher recruitment literature by establishing that financial incentives can have both positive recruitment and negative retention effects, and I am able to evaluate this in the context of the outside option wages of trainees. By framing teaching as a widely available occupation for university graduates, I am also able to evaluate the effect of a one-time financial incentive on occupational choice. My work also acts as an evaluation for a large-scale recruitment policy that is one of the primary paths into teaching in the UK.

Overall, bursaries are an effective tool whose positive recruitment effects overpower their negative retention effects. However, high bursary cohorts are subject to compositional effects that may not be easily observed or readily apparent when initial recruitment occurs. Marginal trainees are relatively more financially motivated and may react more strongly to the relative stagnation of public sector teacher pay. Alternatively, additional research could explore whether salary increases are a more effective retention tool for high-bursary teacher cohorts. Future research could also examine how wage-setting practises within the UK schooling system impacts the distribution of new teachers. The most important question remaining is whether these marginal candidates differ in their teaching quality. Additional research is required to assess whether high-bursary teachers improve not only the pupil-teacher ratio, but also pupils' educational attainment.

References

Abebe, G., Caria, A. S. and Ortiz-Ospina, E. (2021), 'The Selection of Talent: Experimental and Structural Evidence from Ethiopia', *American Economic Review* **111**(6), 1757–1806.

Adams, L., Coburn-Crane, S., Sanders-Earley, A., Keeble, R., Harris, H., Taylor, J. and Taylor, B. (2023), Working lives of teachers and leaders – wave 1, Technical report, Department for Education.

Allen, R., Belfield, C., Greaves, E., Sharp, C. and Walker, M. (2014), The Costs and Benefits of Different Initial Teacher Training Routes, Technical report, Institute for Fiscal Studies.

Angrist, J. D. and Lavy, V. (1999), 'Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement', *The Quarterly Journal of Economics* **114**(2), 533–575.

Ashraf, N., Bandiera, O., Davenport, E. and Lee, S. S. (2020), 'Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services', *American Economic Review* **110**(5), 1355–94.

Bagger, J., Fontaine, F., Postel-Vinay, F. and Robin, J. M. (2014), 'Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics', *American Economic Review* **104**(6), 1551–96.

Bedard, K. and Herman, D. A. (2008), 'Who goes to graduate/professional school? The importance of economic fluctuations, undergraduate field, and ability', *Economics of Education Review* **27**(2), 197–210.

Biasi, B. (2024), Salaries, Pensions, and the Retention of Public-Sector Employees: Evidence from Wisconsin Teachers.

Britton, J., Walker, I. and Waltmann, B. (2022), How much does it pay to get good grades at university?, Technical report, Institute for Fiscal Studies, London.

Bueno, C. and Sass, T. R. (2018), The Effects of Differential Pay on Teacher Recruitment and Retention.

Burdett, K. and Coles, M. (2003), 'Equilibrium wage-tenure contracts', *Econometrica* **71**(5), 1377–1404.

Christian, A., Ronfeldt, M. and Zafar, B. (2024), 'College Students and Career Aspirations: Nudging Student Interest in Teaching'.

Clotfelter, C., Glennie, E., Ladd, H. and Vigdor, J. (2008), 'Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina', *Journal of Public Economics* **92**(5-6), 1352–1370.

Clotfelter, C. T., Ladd, H. F. and Vigdor, J. L. (2006), 'Teacher-student matching and the assessment of teacher effectiveness', *Journal of Human Resources* **41**(4).

Coffman, L. C., Conlon, J. J., Featherstone, C. R. and Kessler, J. B. (2019), 'Liquidity Affects Job Choice: Evidence from Teach for America', *The Quarterly Journal of Economics* **134**(4), 2203–2236.

Coffman, L. C., Conlon, J. J., Featherstone, C. R., Kessler, J. B. and Mixon, J. (2023), 'Liquidity for teachers: Evidence from Teach For America and LinkedIn', *Economics of Education Review* **97**, 102460.

Crawford, R. and Disney, R. (2018), 'Wage Regulation and the Quality of Police Applicants', *Economica* **85**(340), 701–734.

Dal Bo, E., Finan, F. and Rossi, M. A. (2013), 'Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*', *The Quarterly Journal of Economics* **128**(3), 1169–1218.

De Falco, A., Hattemer, B. and Sierra Vásquez, S. (2024), Recruiting Better Teachers? Evidence from a Higher Education Reform in Chile.

Dellas, H. and Sakellaris, P. (2003), 'On the cyclicality of schooling: theory and evidence', *Oxford Economic Papers* **55**(1), 148–172.

Department for Education (2011), Training our next generation of outstanding teachers Implementation plan, Technical report.

Deserranno, E. (2019), 'Financial incentives as signals: Experimental evidence from the recruitment of village promoters in Uganda', *American Economic Journal: Applied Economics* **11**(1).

Falch, T. (2011), 'Teacher Mobility Responses to Wage Changes: Evidence from a Quasi-natural Experiment', *American Economic Review* **101**(3), 460–65.

Feng, L. and Sass, T. R. (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.

Finn, J. D. and Achilles, C. M. (1999), 'Tennessee's Class Size Study: Findings, Implications, Misconceptions', *Educational evaluation and policy analysis* **21**(2), 97–109.

Fullard, J. (2021), Bad economy, good teachers? The countercyclicality of enrolment Into Initial Teacher Training Programmes in the UK .

Fullard, J. and Zuccollo, J. (2021), Local pay and teacher retention in England, Technical report, Education Policy Institute.

Hanushek, E. A. (2011), 'The economic value of higher teacher quality', *Economics of Education Review* **30**(3), 466–479.

Hendricks, M. D. (2014), 'Does it pay to pay teachers more? Evidence from Texas', *Journal of Public Economics* **109**, 50–63.

Keane, M. P. and Wolpin, K. I. (1997), 'The career decisions of young men', *Journal of Political Economy* **105**(3), 473–522.

Leaver, C., Ozier, O., Serneels, P. and Zeitlin, A. (2021), 'Recruitment, Effort, and Retention Effects of Performance Contracts for Civil Servants: Experimental Evidence from Rwandan Primary Schools', *American Economic Review* **111**(7), 2213–46.

McLean, D., Tang, S. and Worth, J. (2023), The impact of training bursaries on teacher recruitment and retention: An evaluation of impact and value for money, Technical report, National Foundation for Education Research.

Nagler, M., Piopiunik, M. and West, M. R. (2020), 'Weak Markets, Strong Teachers: Recession at Career Start and Teacher Effectiveness', *Journal of Labor Economics* **38**(2), 453–500.

National Audit Office (2016), Training new teachers, Technical report, Department for Education.

Nickell, S. and Quintini, G. (2002), 'The Consequences of the Decline in Public Sector Pay in Britain: A Little Bit of Evidence', *The Economic Journal* **112**(477), F107–F118.

Podgursky, M., Monroe, R. and Watson, D. (2004), 'The academic quality of public school teachers: an analysis of entry and exit behavior', *Economics of Education Review* **23**(5), 507–518.

Propper, C. and Van Reenen, J. (2010), 'Can pay regulation kill? Panel data evidence on the effect of labor markets on hospital performance', *Journal of Political Economy* **118**(2), 222–273.

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005), 'Teachers, Schools, and Academic Achievement', *Econometrica* **73**(2), 417–458.

Rockoff, J. E. (2004), 'The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data', *American Economic Review* **94**(2), 247–252.

Shi, S. (2009), 'Directed Search for Equilibrium Wage-Tenure Contracts', *Econometrica* **77**(2).

Sibieta, L. (2023), 'What has happened to teacher pay in England?'.

Sims, S. and Benhenda, A. (2022), The effect of financial incentives on the retention of shortage-subject teachers: evidence from England.

Stinebrickner, T. R. (1998), 'An empirical investigation of teacher attrition', *Economics of Education Review* **17**(2), 127–136.

Stinebrickner, T. R. (2001*a*), 'A Dynamic Model of Teacher Labor Supply', *Journal of Labor Economics* **19**(1), 196–230.

Stinebrickner, T. R. (2001*b*), 'Compensation Policies and Teacher Decisions', *International Economic Review* **42**(3), 751–780.

Tsao, C. (2024), It's Not (Just) About the Money: The Value of Working Conditions in Teaching.

Worth, J. (2021), Do bursaries change who applies to teacher training?, Technical report, National Foundation for Education Research.

Worth, J., Tang, S. and Galvis, M. (2022), The impact of pay and financial incentives on teacher supply, Technical report, NFER.

10 Tables

Table 1: Within-sample Trainee Characteristics

	All trainees	STEM trainees	Non-STEM trainees	STEM Difference
share female	0.60	0.52	0.65	-0.13
mean age	28.53	29.75	27.62	2.13
share under 26	0.52	0.47	0.55	-0.07
share non white	0.21	0.29	0.15	0.14
share census matched	0.74	0.72	0.76	-0.04
share with bursary	0.81	0.92	0.74	0.18
Total Observations	95397	39347	56050	16703

All differences are significant at the 1% level

Table 2: Cohort Size Regression Results: Subject-Year Level

	All Subjects		Stem Subject		Non-Stem Subject	
	Trainees	Teachers	Trainees	Teachers	Trainees	Teachers
Bursary	0.292**	0.136	0.413	0.352***	0.208	0.066
	[0.036]	[0.186]	[0.190]	[0.000]	[0.122]	[0.748]
Observations	383	215	140	80	243	135

P-values of Wild-Bootstrapped regressions in brackets. *p <0.10, **p<0.05, ***p<0.01. Bursary is in units of £10k. All regressions control for subject-specific trainee targets, with wild-bootstrapped standard errors clustered at the subject level.

Table 3: Marginal Impact of Bursary on Cohort Characteristics

All Subjects

Cohort Share	Female	Non-White	Under-26
Bursary	0.000	-0.006	-0.045***
	(0.005)	(0.005)	(0.007)
Observations	84,857	81,498	84,857

Non-Stem Subjects

Cohort Share	Female	Non-White	Under-26
Bursary	-0.008	-0.006	-0.048***
	(0.008)	(0.006)	(0.008)
Observations	49,007	47,078	49,007

Stem Subjects

Cohort Share	Female	Non-White	Under-26
Bursary	0.021**	-0.012	-0.058***
	(0.009)	(0.008)	(0.009)
Observations	35,850	34,420	35,850

Logistic regression with standard errors in parentheses. *p <0.10, **p<0.05, ***p<0.01
Marginal effects evaluated at the average probability. Units of £10k. All regressions control for subject-specific trainee targets.

Table 4: Marginal Effect on Becoming a Teacher Post-Qualification

	Appear Post-Qualification			
	All Subjects			
	(1)	(2)	(3)	(4)
Bursary^{+}	-0.018***	-0.018***	-0.016**	-0.018**
	(0.006)	(0.006)	(0.007)	(0.008)
Trainee Target^{++}		-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.002)
Outside Wage Gap^{++}			-0.045***	0.002
			(0.002)	(0.003)
Course Controls	X	X	X	X
Characteristics	-	-	-	X
Observations	95,397	95,397	76,651	76,651

Standard Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01

Marginal effects evaluated at the average probability. $^{+}$ Units of £10k. $^{++}$ Units of 100. Outside wage gap controls for the predicted [outside wage - teacher wage] in the training year.

Table 5: Marginal Effect on Becoming a Teacher Post-Qualification (By Stem Status)

	Appear Post-Qualification							
	Stem				Non-Stem			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Bursary^{+}	-0.028***	-0.022***	-0.031***	-0.030***	-0.011	-0.011	0.000	0.001
	(0.007)	(0.006)	(0.007)	(0.007)	(0.009)	(0.093)	(0.011)	(0.011)
Trainee Target^{++}		0.004**	0.005**	0.004**		-0.002	-0.005**	-0.004**
		(0.002)	(0.002)	(0.002)		(0.002)	(0.112)	(0.002)
Outside Wage Gap^{++}			-0.044***	0.002			-0.052***	0.002
			(0.002)	(0.004)			(0.003)	(0.004)
Course Controls	X	X	X	X	X	X	X	X
Characteristics	-	-	-	X	-	-	-	X
Observations	39,347	39,347	32,459	32,459	56,050	56,050	44,192	44,192

Standard Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01. Marginal effects evaluated at the average probability. $^{+}$ Units of £10k. $^{++}$ Units of 100. Outside wage gap controls for the predicted [outside wage - teacher wage] in the Training year.

Table 6: Marginal Impact on Remaining in Teaching (Conditional on Entry Post-Qualification)

All Subjects

	Probability of Conditional Retention by Year							
	Year 1		Year 2		Year 3		Year 4	
Bursary†	-0.005	-0.010*	-0.004	-0.009	-0.001	-0.001	-0.016	-0.016
	(0.003)	(0.005)	(0.003)	(0.007)	(0.006)	(0.008)	(0.011)	(0.011)
Trainee Target††	-0.001*	-0.002*	-0.000	-0.000	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Wage Gap (Entry year)††	-0.002	0.031*	-0.002	0.022***	0.006	0.027***	-0.001	0.029***
	(0.009)	(0.016)	(0.003)	(0.007)	(0.005)	(0.007)	(0.006)	(0.008)
Observations	50,766	50,766	44,456	44,456	38,849	38,849	29,091	29,091

Non-Stem Subjects

	Year 1		Year 2		Year 3		Year 4	
Bursary†	-0.002	-0.004	-0.002	-0.009*	-0.005	-0.008	-0.020	-0.022
	(0.002)	(0.004)	(0.003)	(0.005)	(0.005)	(0.008)	(0.013)	(0.013)
Trainee Target††	-0.001**	-0.003**	0.000	-0.001	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Wage Gap (Entry year)	-0.007	0.0019*	-0.003	0.001	0.004	0.014**	0.003	0.024***
	(0.005)	(0.010)	(0.002)	(0.006)	(0.004)	(0.007)	(0.006)	(0.008)
Observations	30,213	30,213	26,362	26,362	22,964	22,964	17,310	17,310

Stem Subjects

	Year 1		Year 2		Year 3		Year 4	
Bursary†	-0.013**	-0.019**	-0.006	-0.011	0.012	0.016	0.035	0.038*
	(0.006)	(0.009)	(0.009)	(0.0016)	(0.0014)	(0.0017)	(0.022)	(0.021)
Trainee Target††	0.002	0.002	0.000	0.002	0.002	0.000	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
Wage Gap (Entry year)	0.036*	0.044	0.003	0.052***	0.008	0.043***	-0.007	0.0032**
	(0.021)	(0.039)	(0.007)	(0.013)	(0.009)	(0.012)	(0.011)	(0.013)
Observations	20,553	20,553	18,094	18,094	15,885	15,885	11,781	11,781
Course Controls	X	X	X	X	X	X	X	X
Characteristics	-	X	-	X	-	X	-	X

Standard Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01

Marginal effects evaluated at the average probability. † Units of £10k. †† Units of 100. Y variable is the probability an individual is present in the school workforce census X years after training ends. This conditional measure excludes those who don't appear as a teacher at all post-training from the regression. Outside wage gap controls for the predicted [outside wage - teacher wage] in the year of entry into teacher training.

Table 7: Teaching Post-qualification: Bursary-Region Interaction Marginal Effects

	Appear Post-Qualification		
Region x Bursary Level⋆	All	Non-Stem	Stem
South West	-0.034***	-0.037***	-0.030**
	(0.009)	(0.013)	(0.013)
South East	-0.024**	-0.012	-0.023**
	(0.007)	(0.012)	(0.012)
Outer London	-0.003	0.014*	-0.022**
	(0.007)	(0.009)	(0.009)
Inner London	-0.029***	-0.006	-0.042***
	(0.012)	(0.016)	(0.014)
East of England	-0.033***	-0.038***	-0.032***
	(0.007)	(0.010)	(0.012)
East Midlands	-0.021***	-0.008	-0.025**
	(0.009)	(0.012)	(0.015)
West Midlands	-0.010	-0.009	-0.003
	(0.007)	(0.010)	(0.010)
North West	-0.002	0.003	-0.007
	(0.008)	(0.009)	(0.010)
Yorkshire and the Humber	-0.023**	-0.012	-0.038***
	(0.009)	(0.013)	(0.011)
North East	-0.047***	-0.026**	-0.047***
	(0.010)	(0.011)	(0.011)
Course Controls	X	X	X
Characteristics	X	X	X
Trainee Targets	X	X	X
Outside Wages	-	-	-
Observations	76,893	32,557	44,336

Reports the marginal change in probability evaluated at the regional average. Standard Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01.

⋆ Bursary is in units of £10k. All regressions control for year, region, trainee targets and subject fixed effects.

Table 8: Appearing Post-qualification: Wage Interactions

	Appear Post-Qualification		
	All	Non-Stem	Stem
Bursary^{+}	-0.020***	-0.002	-0.031***
	(0.008)	(0.011)	(0.007)
Wage Gap^{++}	-0.012***	-0.014**	-0.009
	(0.008)	(0.006)	(0.006)
Bursary x Wage Gap^{+++}	0.001***	0.013***	0.005**
	(0.002)	(0.002)	(0.002)
Course Controls	X	X	X
Characteristics	X	X	X
Trainee Targets	X	X	X
Observations	76,651	44,192	32,459

Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01. Marginal effects evaluated at the average probability. $^{+}$ Units of £10k. $^{++}$ Units of 100. $^{+++}$ Units of (100*10k) Outside wage gap controls for the predicted [outside wage - teacher wage] in the training year. All regressions control for year, region and subject fixed effects.

Table 9: Simulation Results: Motivation by Bursary Level

Bursary Level	Average Trainee Outside Wage	Average Trainee Motivation	Average Complier Motivation
0	25871	3056	3107
5000	25861	2937	2998
10000	26022	2923	2997
15000	26080	2780	2858
20000	26163	2789	2872
25000	26282	2720	2808
30000	26302	2717	2843
35000	26536	2634	2748
40000	26574	2681	2787

Table 10: Simulation Results: Compliers and Never-Takers by Bursary Level

Bursary Level	Trainee Complier Share	Δ Total Trainees	Δ Total Compliers	Δ Total Never-Takers	Δ Complier Share
0	0.982	1.000	1.000	1.000	0.000
5000	0.979	1.002	1.000	1.144	-0.003
10000	0.976	1.031	1.025	1.349	-0.006
15000	0.973	1.043	1.033	1.589	-0.010
20000	0.968	1.069	1.053	1.895	-0.014
25000	0.964	1.091	1.072	2.182	-0.018
30000	0.956	1.103	1.074	2.694	-0.026
35000	0.952	1.129	1.095	3.000	-0.030
40000	0.949	1.140	1.102	3.215	-0.033

Table 11: Simulation Results: Compliers and Never-Takers by Wage Level

Wage	Number of Trainees	Average Trainee Motivation	Average Complier Motivation	Share Never Taker
20000	40844	917	982	0.004
25000	31791	2312	2443	0.010
30000	20451	3770	3987	0.019
35000	11233	5472	5844	0.031
40000	5095	7711	8261	0.045
45000	1763	10147	11025	0.074
50000	416	12236	13296	0.144
55000	45	13889	15000	0.244

Table 12: Cost-Benefit Policy Comparison: Cost of 5,000 Additional Teaching Years

Policy	Value Raise	Marginal Cost	Cost per Teacher-Year
Bursary Raise (All Subjects)	£6,175	£290,042,633	£58,001
Raise all pay (1 year)	0.67%	£288,559,712	£57,817
Raise early career pay (1 cohort)	2.5%	£185,958,682	£37,620

Figures are based on teacher cohort characteristics and pay in the 2022/2023 academic cycle.

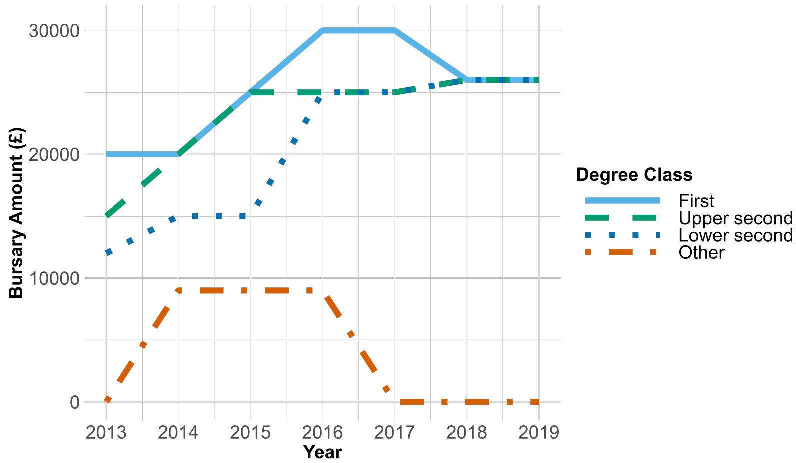
Table 13: Cost-Benefit of a 10 Percent Increase in Trainees

Bursary Type	Value Increase	Additional years	Additional Cost	Average Cost
General Bursary	£3,596	2,923	£165,867,494	£56,763
Best Case Bursary	£2,587	2,923	£150,918,726	£51,645
Worst Case Bursary	£5,904	2,923	£200,517,472	£68,608
STEM Bursary	£2,588	907	£51,021,155	£55,786
Non-STEM Bursary	£4,908	2,016	£120,537,846	£59,352
Maths First Class Bursary	£1,150	101	£5,049,268	£49,789

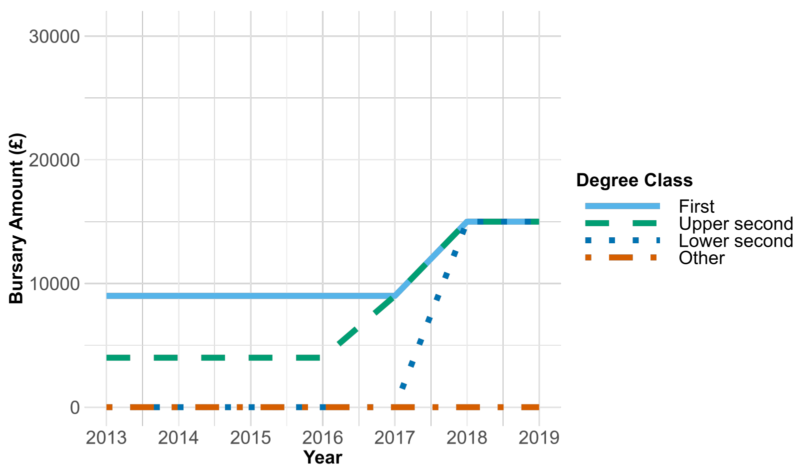
Figures are evaluated for a 10% increase of the specified cohort size from the existing average bursary level for that group in the 2022/2023 training cycle. The group refers to the specific training cohort that would be given a bursary uplift. Best case estimate uses the upper bound of coefficients, and the worst case uses the lower bounds.

11 Figures

Figure 1: Bursary Levels over Time



(a) Physics



(b) Modern Foreign Languages

Each line demonstrates the evolution of the bursary level offered over time for an individual with a certain classification in their undergraduate degree.

Figure 2: Training Timeline

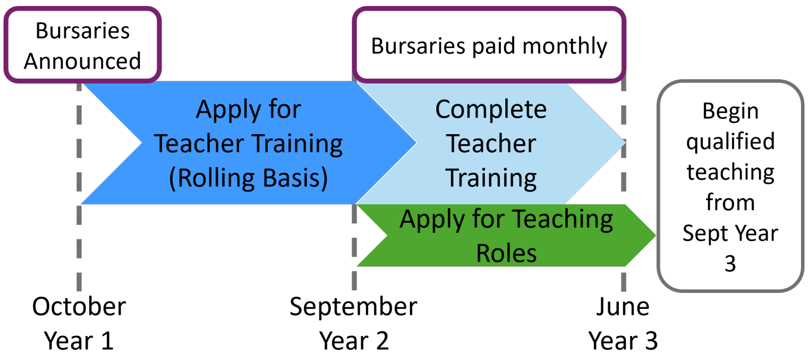
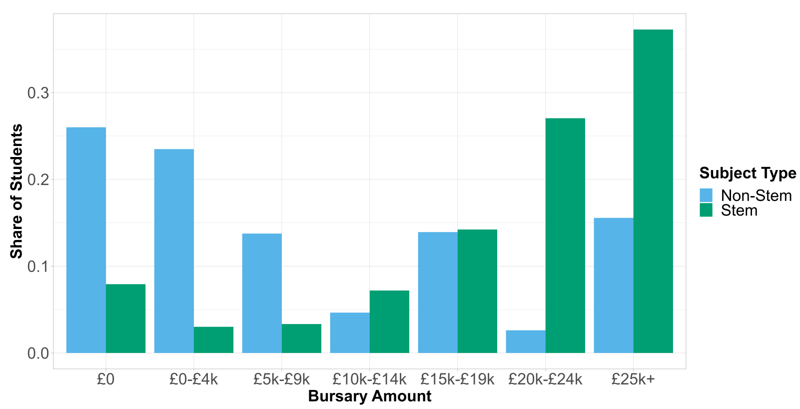
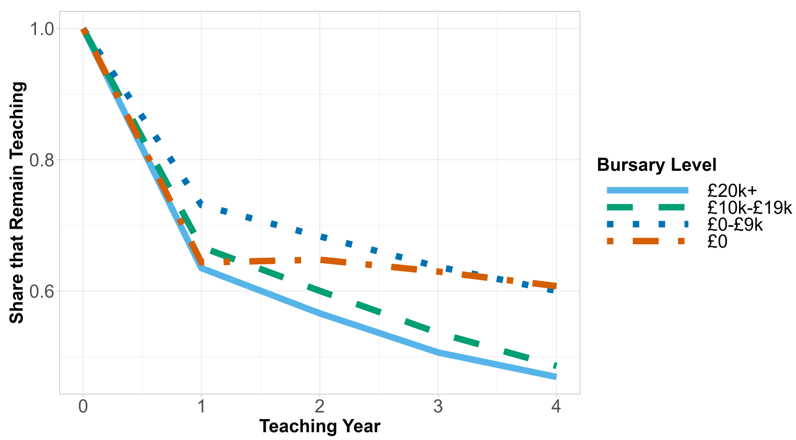


Figure 3: Distribution of Bursary Awards



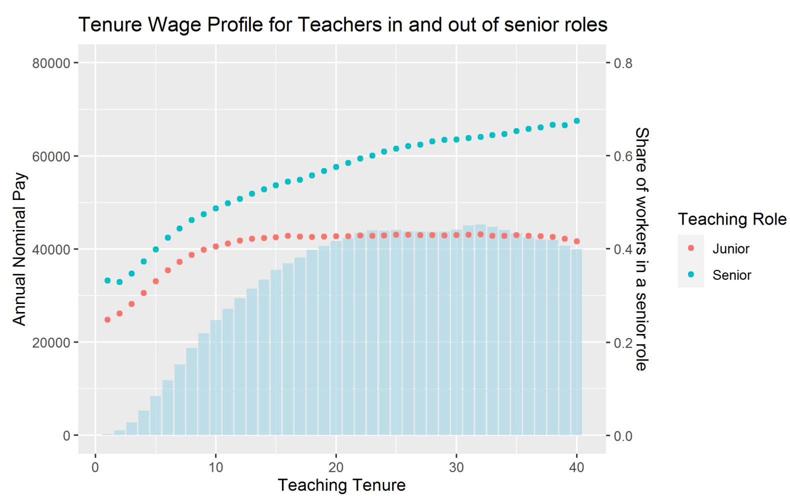
Initial Teacher Training (ITT) Data, 2013-2020, DfE. Bursary level is matched to trainees based on their characteristics and subject chosen.

Figure 4: Retention Rate by Level of Bursary Award



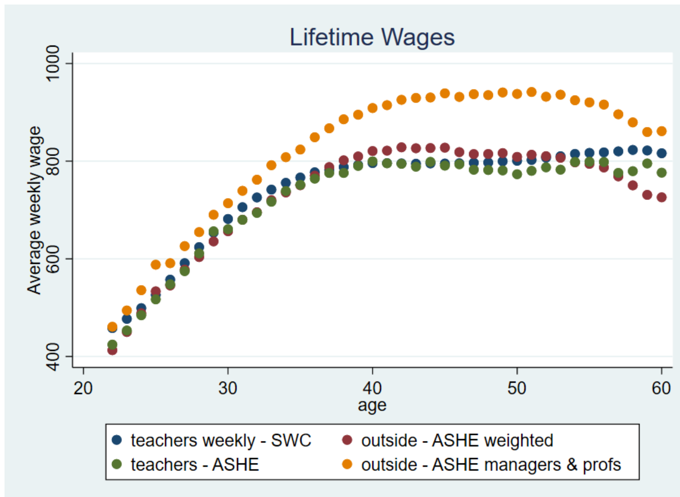
Data: School Workforce Census (SWC) and Initial Teacher Training (ITT) Data, 2013-2020, DfE. An is teaching year in x if they are present in the school workforce census x years after their training.

Figure 5: Average Teacher Salaries over Teaching Tenure



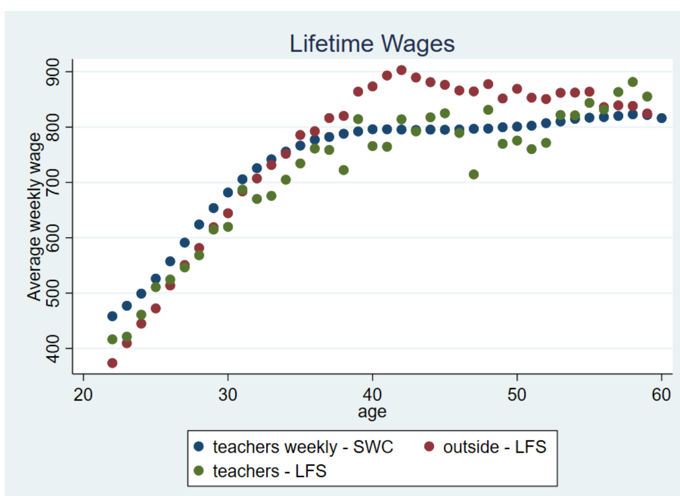
Data: School Workforce Census (SWC) 2013-2020, DfE. Blue bars represent the share of the teaching workforce at each tenure level that holds a senior role. Teaching tenure is calculated as years since qualified teacher status was attained. Senior staff roles are heads of department, heads of year, deputy head-teachers and head-teachers.

Figure 6: Teachers and Non-Teachers Age-Earnings Profile (ASHE)



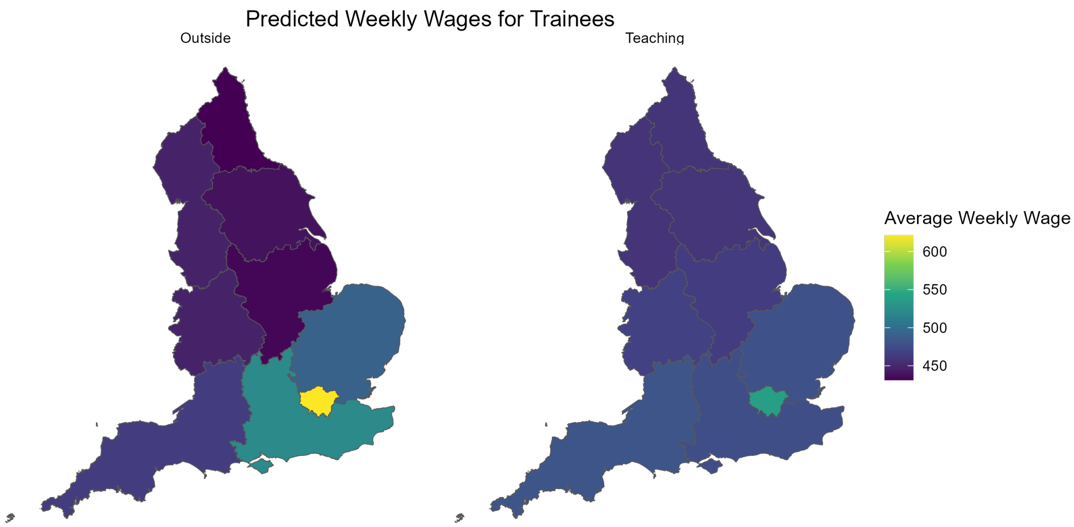
Data: SWC and ASHE 2013-2020. ASHE weighted represents earnings weighted according to a job's likelihood of belonging to a graduate in the LFS. Managers and Professionals represent earnings for SOC levels 1 and 2. More details in Appendix.

Figure 7: Teachers and Non-Teachers Age-Earnings Profile (LFS)



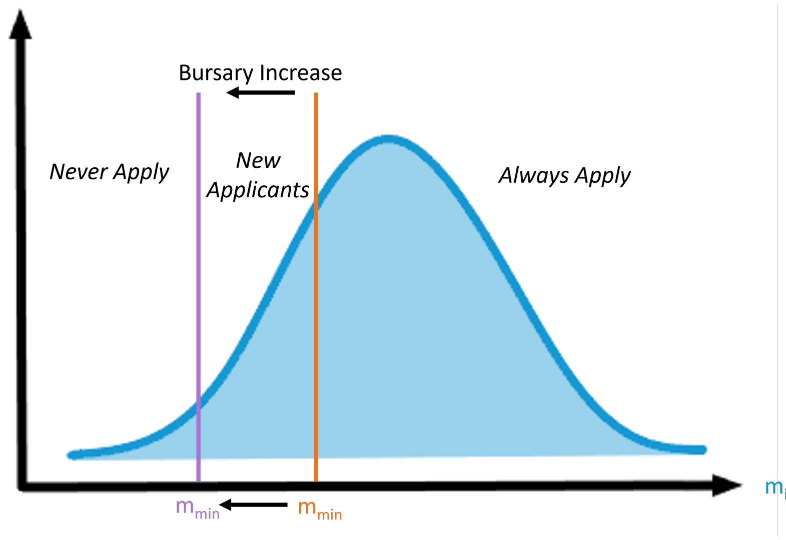
Data: SWC and LFS 2013-2020. LFS outside wages represent earnings for all UK-born graduates.

Figure 8: Wages for Teachers and their Outside Options



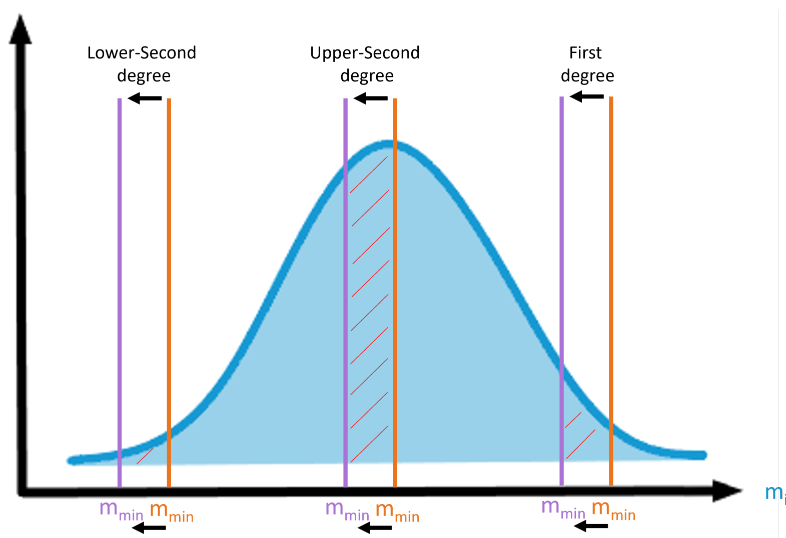
Data: SWC, ASHE, and LFS. Teaching wages are the starting wage predicted via a regression of using DfE data. Outside wages are predicted for the year of entry based on age, region, year, sex, and undergraduate characteristics. The predictions are generated using ASHE data, where jobs are weighted according to their likelihood of belonging to a graduate in the LFS. More details in Appendix.

Figure 9: Model Result 1: Bursary Raise Increases Trainees



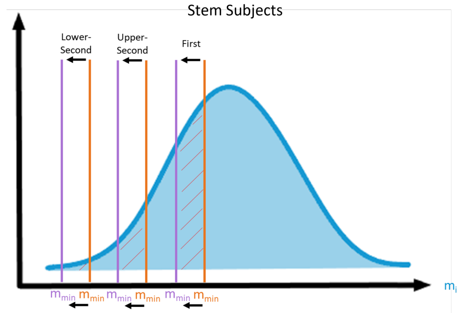
Red shaded area denotes the individuals induced to train due to a marginal increase in the bursary level.

Figure 10: Model Result 3: Selection Effects by Qualification



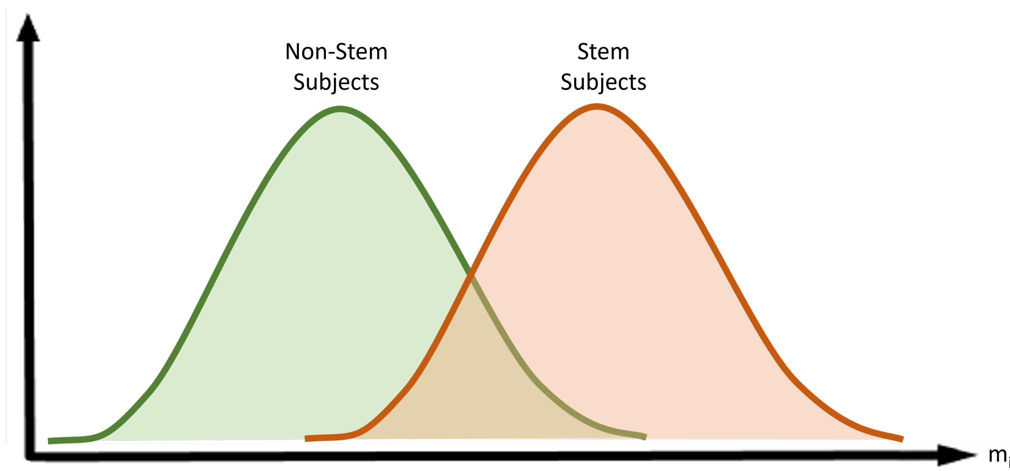
Red shaded area denotes the individuals induced to train due to a marginal increase in the bursary level. Note that the majority of candidates who apply are those who hold an upper-second class degree.

Figure 11: Model Result 3: Separate Selection for STEM and Non-STEM



Red shaded area denotes the individuals induced to train due to a marginal increase in the bursary level. Due to different outside wages, selection differs between two groups.

Figure 12: Model Result 3: Motivation Distribution by Subject



Those who choose to study STEM at undergraduate have a higher average non-pecuniary valuation of teaching than those who study non-STEM for the empirical results to hold.

Figure 13: Simulation Result: Share in Teaching over Time

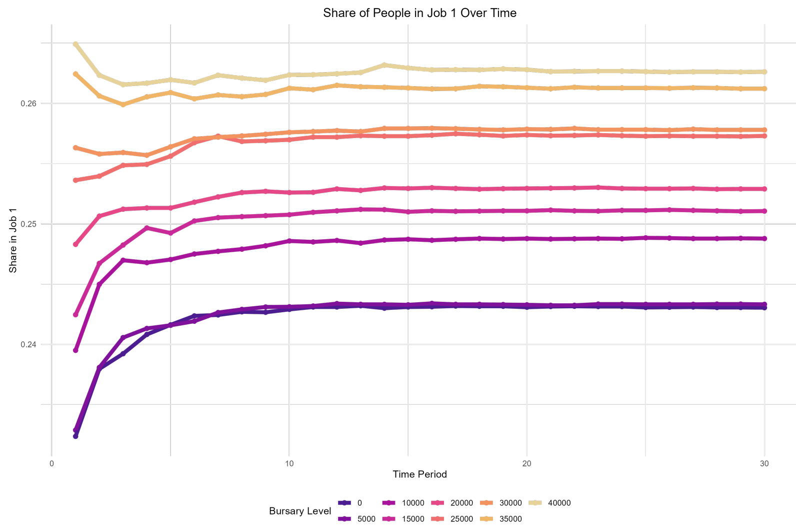


Figure 14: Simulation Result: Cohort Size by Bursary Level

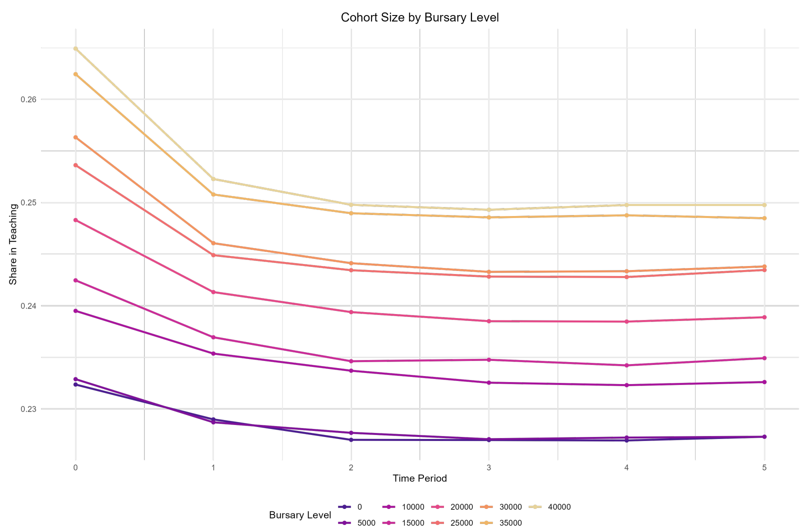


Figure 15: Simulation Result: Cohort Size by Bursary Level and Wage

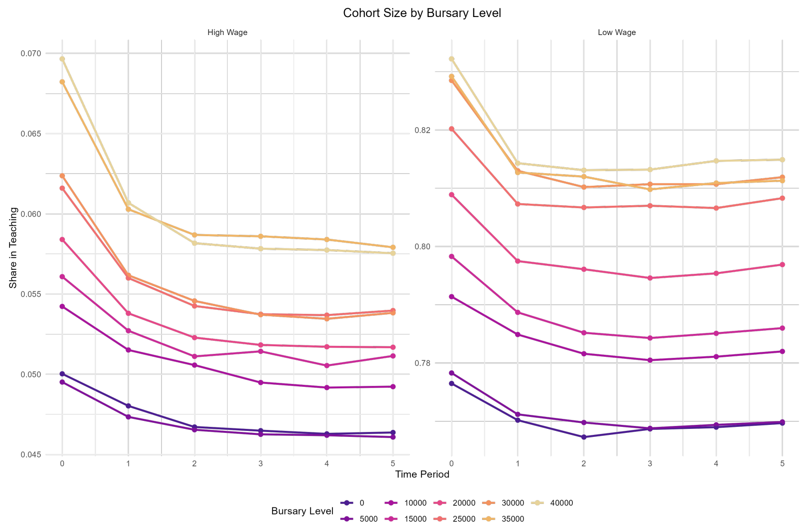


Figure 16: Simulation Result: Retention by Bursary Level

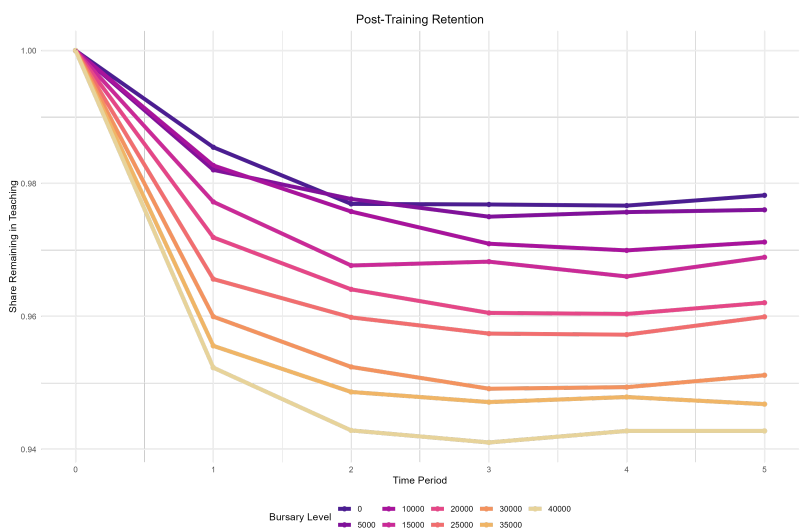


Figure 17: Simulation Result: Retention by Wage

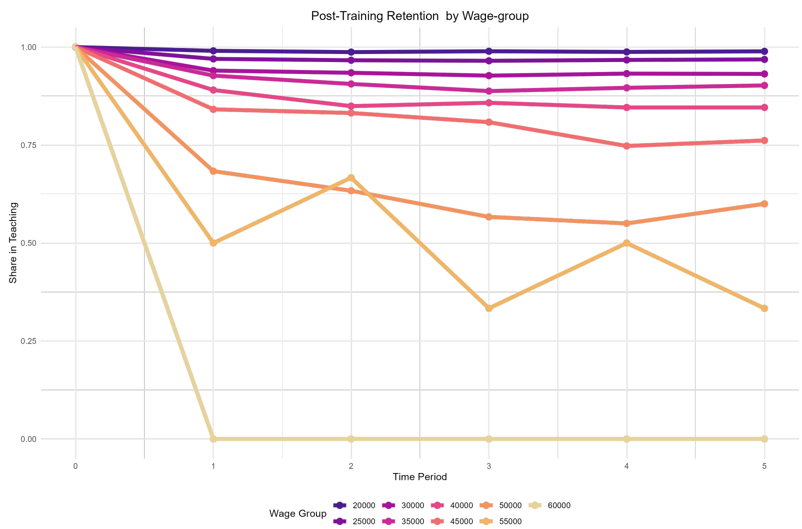


Figure 18: Simulation Result: Retention by Bursary Level and Wage

12 Appendix

A Calculating Wage Gaps

I Estimate the expected wage gaps of teacher trainees using three data sources: the School Workforce Census (SWC), the Labour Force Survey (LFS) and Annual Survey of Hours and Earnings (ASHE). The benefit of the LFS is that it contains detailed information on personal characteristics and qualifications that is not included in ASHE. However, the LFS is a household study with a smaller sample size than the ASHE. Both ASHE and the SWC are completed by the employer and so salary estimates are more comparable. I use a combination of both ASHE and the LFS for the wage gaps used in the main regressions, but for robustness I construct multiple alternative measures as discussed below. In all analysis I restrict my sample to full-time working age UK-born individuals working in England.

I firstly use the LFS to generate job weightings that report the probability of that job being occupied by someone with an undergraduate degree or higher. For each combination of industry division s and 2-digit occupation o, I calculate P_{sodj}: the share of jobs that are held by graduates. I do this separately for each combination of undergraduate classification d and subject j. In other words, for graduates of each subject-classification combination, I estimate the share of jobs held by that group in each occupation-industry. In cases where under 10 observations exist within that occupation-industry combination, I replace this with the calculated share for the larger industry sector and 2 digit occupation group.

I then estimate non-teacher wages using ASHE data with LFS weightings applied. I use the ASHE data in a set of regressions of log weekly non-teacher wages on a set of controls, weighted by the LFS job shares P_{sodj}. The controls include year, age, sex, region, a full set of their interactions, and age squared. I run separate regressions for each subject-classification combination and store the estimated set of coefficients, β_{dj}. These coefficients allow me to predict the wages of an individual using ASHE data whilst still taking into account their qualifications, information that is otherwise missing from ASHE.

Next,to estimate teacher wages, I regress log weekly secondary school teacher wages on a set of controls using the SWC. The controls include teaching tenure, teaching tenure squared, year and region[38]. I use tenure rather than age as I assume that all trainees have no prior experience and will be offered similar wages regardless of age when offered a teaching position.

I predict both outside wages and teacher wages for trainees for each year using the stored regression estimates. The wage gap for person i in year y who trained in year t with age x and tenure $y - t$ is calculated in equation 12 below.

$$\widehat{\text{Wage Gap}}_{i,y,x,t} = \widehat{\text{Teacher Wage}}_{i,y,t} - \widehat{\text{Outside Wage}}_{i,y,x} \tag{12}$$

I also construct alternative measures of outside wages. Firstly, I use the same estimation process but use alternative binary weights that are equal to one if an occupation-sector has over 80 percent graduates. I call this measure the 'graduate' outside wage. The 'basic' wage gap calculates the original job shares for just two groups: STEM and non-STEM graduates. Next, I use weights equal to the share of ex-teachers found within that profession. I exclude any profession with values less than half a percent. The remaining sample of occupation-sector combinations is relatively narrow because teachers tend to exit into a distinct set of jobs. The resulting wage gaps are referred to as the 'exit' wage. Lastly, I generate 'LFS' wage using LFS data with all the controls included in the primary measure, in addition to controls for undergraduate classification, subject, and ethnicity.

The main measure of wage is the median measure of the five methods. Graduate wages predict the highest average, followed by the LFS measure. The basic average wage is second lowest, and the teacher exit method predicts the lowest wages. When controlling for outside wages in tables 4, 5, 6, and 7, the coefficient on outside wages and its significance can vary based on the particular measure. However, the primary measure of wages generally produces the median coefficient estimate, whereas other

[38]coefficients on sex and ethnicity are not included, but were small and not significant when included.

measures can produce more extreme estimates. Most importantly, the coefficient of the bursary level on key outcomes is reliably consistent across these measures.

B Model Simulations

In this section, I discuss how I simulate the effects of the model outlined in section 7. I generate a population of individuals who live 30 time periods and have a discount rate of 5 percent. I enable the relative wage to change by fixing the characteristics of job 1 (teaching) and varying the characteristics of job 2 (the outside option). Teaching has an initial starting wage of £30,000 and return to experience of 3%. Job 2 has 10 possible starting wage values between £20,000 and £65,000, increasing in £5k increments. An individual is offered one of nine different bursary amounts between 0 and £40k, set at £5k increments. Job 2 also has 5 possible values for the return to experience set at 1% increments between 1% and 5%. There are 450 different combinations of these three variables, and for each combination I simulate the decisions of 1,000 agents. Using this set up, I am able to observe the effect of varying bursary amounts on the same population.

Each individual is randomly assigned a non-pecuniary valuation of teaching, or a 'motivation'. Motivation takes one of seven discrete values between -15,000 and 15,000, where the probabilities follow a normal distribution around 0 with a standard deviation of 7,500. In other words, individuals have a 3.66% percent chance of having a motivation of 15,000 and a 27.07% chance of having a motivation of 0. Each agent's wages for jobs 1 and 2 contain a random component that follow an AR(1) error procedure.

At t=0, the agent observed the wage offers for both jobs in period t=1 and evaluates the expected lifetime value of four action sets: job 1 till retirement, job 2 till retirement, job 1 for one period then job 2 till retirement, or job 2 for one period then job 2 till retirement[39]. An individual then selects their job for t=1 in order to

[39]This set of choices is in line with the fact that under a model with no bursaries and positive wage growth in

maximise their expected lifetime utility. In period t=1, the wage offers for period 2 are revealed and the individual selects a job based on the four potential action sets. Given this setup, the individual is able to swap between the two jobs for any number of time periods at any time. Individuals receive the bursary in their first year of teaching rather than their wage offer, but the bursary offer does not expire after t=1 and is constant over time.

I relate the simulations to the empirical context by assuming that the population at t=0 is comprised of all graduates in the workforce. Some workers may have higher wages or higher wage growth due to their undergraduate qualifications, or their years of experience. This is an oversimplification, as in reality older graduates who have higher wages due to existing experience have fewer time periods to evaluate their utility across. However in reality, the shares of individuals in each job are relatively stable by year 10, and so I avoid over-complicating the model in this regard. Time t=0 can be considered as the moment the exogeneously-set bursary is offered, and the share of agents who select teaching at t=1 are considered to be the trainees. We can then observe the occupational choices of these trainees over the remaining time periods.

C Cost-Benefit Analysis

I evaluate the cost-effectiveness three retention and recruitment policies: Bursaries, increasing early-career salaries[40], and raising all secondary school teacher salaries. For each policy I evaluate the total cost of the policy and the additional teaching years gained. I use these figures to generate the average cost per additional teaching year.

The stock of trainees and teachers in their first, second, and third years of teach-

both jobs, the agent would select one occupation and only swap following a shock to wages. If a rational prefers job 2, introducing a bursary would lead to a planned deviation into teaching for at most one year. In reality, these plans are not always followed out due to unexpected wage shocks.

[40] Early career is defined as the first 2 years of teaching.

ing each have their own attrition rates estimated using the data. The remaining 'senior' stock of teachers are grouped together with a single rate of attrition. Teachers enter the workforce at three distinct points: postgraduate trainees enter at the training stage, other newly-qualified teachers (e.g. those with undergraduate teaching degrees) enter at year one, and returning teachers enter directly into the senior teaching stock. I use the 2022/23 academic year as a base case, and extract the wage costs and exact number of teachers, trainees, other entrants, and leavers from publicly available data[41].

Bursaries impact the stock of teachers by increasing the number of trainees recruited and altering the retention rates for this group up to year 3, based on regression estimates. I first estimate how a bursary increase affects the number of trainees recruited by multiplying the base number of trainees by the coefficients in tables 1 and 2. These coefficients report the increase in the cohort size following a £10k increase in the bursary. I then estimate the number of qualified teachers that remain in each teaching year by multiplying the number of trainees by the share that remain as a teacher in each following year. When the bursary increases, the share is multiplied by the figures in table 15. The first sub-column of each year reports the change in probability that a trainee appears as a teacher in years 1-4, without conditioning for personal characteristics.

I estimate the additional costs and teaching years gained when the bursary increase is applied to one cohort. Since bursaries affect retention for many years after, I follow the cohort until the end of their third year of teaching and estimate the total additional teaching years gained across this four year window. I consider the training year to be worth half a teaching year in line with the course's the course's stated balance between practical and academic elements. I calculate the cost-effectiveness of an increase in the bursary level for different trainee groups and different values of

[41]School workforce data: https://explore-education-statistics.service.gov.uk/find-statistics/school-workforce-in-england/2022 [last accessed October 2022] and Trainee data: https://explore-education-statistics.service.gov.uk/find-statistics/initial-teacher-training-census/2022-23dataBlock-85b69273-a803-4985-8c29-8a922ab25491-tables [last accessed October 2022]

the increase.

For early career wages, I assume that a higher payment attracts more applicants to the programme and reduces attrition. Early-career pay increases the number of both trainees and other newly-qualified teachers based on a pay-recruitment elasticity of 2 as estimated in Worth et al. (2022). Early-career pay also reduces attrition until the end of year 2 when pay returns to its base case level. I apply a pay-attrition elasticity of minus 3 from Sims and Benhenda (2022) to estimate the impact on exits from teaching. I estimate the additional number of teachers gained between the training year and the end of teaching year 3 when offering an early-career pay rise to one cohort of teachers.

Raising pay for all teachers has the same impacts as early career pay, however it also increases the number of returning teachers that enter the stock of senior teachers and decreases the exit rate of senior teachers. The same entry and exit used for early-career pay are also applied here. I estimate the additional number of teachers gained in a year from all teaching groups.

For each policy, I set a benchmark of 5,000 additional required teaching years and back out the bursary increase or pay rise required to meet this figure. The additional number of teaching years is the total number of teachers employed minus the number of teachers originally employed in the base case. Total costs include wage costs, bursary costs, and training costs for postgraduate trainees (£23,000)[42].

[42] I calculated the average first-year pay of a teacher in 2019 using the SWC and estimate the ratio of this pay relative to the starting pay of a qualified teacher in 2019. I then calculate the average increase in pay per teaching year until year 4. I apply these ratios to the national teaching starting salary in 2022 to generate the expected teacher salary for each year in 2022 prices. The estimated cost of training is extracted from Allen et al. (2014).

D Tables

Table 14: Marginal Impact on Qualifying as Teacher

	Ever-Qualify			
	All Subjects			
	(1)	(2)	(3)	(4)
Bursary^{+}	-0.003	-0.003	-0.002	-0.002
	(0.002)	(0.002)	(0.003)	(0.003)
Trainee Target^{++}		0.000	0.000	-0.000
		(0.001)	(0.001)	(0.001)
Outside Wage Gap^{++}			-0.022***	0.002
			(0.001)	(0.001)
Course Controls	X	X	X	X
Characteristics	-	-	-	X
Observations	95,397	76,760	76,651	76,651

Standard Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01. $^{+}$ Coefficients and SEs Multiplied by 10k. $^{++}$ Trainee Target Coefficients and SEs Multiplied by 100. Outside wage gap controls for the predicted [outside wage - teacher wage] in the training year.

Table 15: Regression Coefficient of Bursary Level on Remaining in Teaching

All Subjects								
Probability of Retention by Year								
	Year 1		Year 2		Year 3		Year 4	
Bursary†	-0.018***	-0.019***	-0.007	-0.007	-0.007	0.009	-0.005	-0.005*
	(0.007)	(0.007)	(0.008)	(0.009)	(0.009)	(0.010)	(0.015)	(0.015)
Trainee Target††	-0.003*	-0.003*	0.000	0.000	0.002*	0.003*	0.003*	0.003*
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Outside Wage Gap††	-0.046**	0.095***	-0.029***	0.057***	-0.022***	0.064***	-0.033***	0.062***
	(0.020)	(0.023)	(0.010)	(0.012)	(0.008)	(0.011)	(0.009)	(0.013)
Observations	66,973	66,973	57,938	57,938	50,257	50,257	37,146	37,146
Non-Stem Subjects								
	Year 1		Year 2		Year 3		Year 4	
Bursary†	-0.008**	-0.009	-0.006	-0.007	-0.002	-0.001	-0.009	-0.014
	(0.010)	(0.010)	(0.010)	(0.010)	(0.012)	(0.013)	(0.026)	(0.027)
Trainee Target††	-0.008***	-0.008***	-0.003	-0.003	0.000	0.000	0.001	0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Outside Wage Gap††	-0.052**	0.093***	-0.042***	0.025	-0.019	0.050***	-0.023**	0.064***
	(0.026)	(0.027)	(0.015)	(0.017)	(0.012)	(0.015)	(0.012)	(0.016)
Observations	38,702	38,702	33,292	33,292	28,727	28,727	21,389	21,389
Stem Subjects								
	Year 1		Year 2		Year 3		Year 4	
Bursary†	-0.030***	-0.032***	-0.017	-0.015	0.010	-0.013	0.028	0.035
	(0.006)	(0.007)	(0.012)	(0.012)	(0.014)	(0.015)	(0.016)	(0.018)
Trainee Target††	0.006***	0.006***	0.004*	0.004	0.002	0.002	0.000	0.000
	(0.002	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Outside Wage Gap††	-0.035	0.108***	-0.021	0.091***	-0.026**	0.085***	-0.041***	0.062***
	(0.031)	(0.038)	(0.013)	(0.017)	(0.011)	(0.018)	(0.014)	(0.020)
Observations	28,271	28,271	24,646	24,646	21,530	21,530	15,757	15,757 height
Course Controls	X	X	X	X	X	X	X	X
Characteristics	-	X	-	X	-	X	-	X

Standard Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01

Marginal effects evaluated at the average probability. † Units of £10k. †† Units of 100. Y variable is the probability an individual is present in the school workforce census X years after training ends. Outside wage gap controls for the predicted [outside wage - teacher wage] in the reference year.

Table 16: Cohort Size Regression Results: By Degree Classification

All Subjects

	Trainees Recruited			Teachers Employed		
	First	2:1	2:2	First	2:1	2:2
Bursary	0.297	0.229**	0.324**	0.017	0.105	0.205
	[0.104]	[0.046]	[0.016]	[0.892]	[0.152]	[0.106]
Observations	96	96	96	54	54	54

Non-Stem Subjects

	Trainees Recruited			Teachers Employed		
	First	2:1	2:2	First	2:1	2:2
Bursary	0.168	0.156	0.296**	-0.103	0.068	0.169
	[0.664]	[0.160]	[0.034]	[0.736]	[0.532]	[0.132]
Observations	61	61	61	34	34	34

Stem Subjects

	Trainees Recruited			Teachers Employed		
	First	2:1	2:2	First	2:1	2:2
Bursary	0.506*	0.343	0.409	0.349*	0.200*	0.492
	[0.078]	[0.174]	[0.188]	[0.064]	[0.076]	[0.178]
Observations	35	35	35	20	20	20

P-values of Wild-Bootstrapped regressions in brackets. *p <0.10, **p<0.05, ***p<0.01. Bursary is in units of £10k. All regressions control for subject-specific trainee targets, with errors clustered at the subject level. Third class degrees are not included due to insufficient variation in the bursary level.

Table 17: Marginal Impact on Teaching Post-qualification: Other Selected Coefficients

	Appear Post-Qualification		
	All	Non-Stem	Stem
Age	-0.009***	-0.009***	-0.008***
	(0.000)	(0.001)	(0.001)
female	0.098***	0.101***	0.097***
	(0.005)	(0.006)	(0.007)
Upper Second	0.005	-0.004	0.014**
	(0.008)	(0.011)	(0.007)
Lower Second	-0.021**	-0.039***	-0.003
	(0.009)	(0.015)	(0.009)
Other Class.	-0.047**	-0.034	-0.055***
	(0.019)	(0.026)	(0.020)
Indian	-0.006	-0.011	-0.012
	(0.012)	(0.021)	(0.013)
Black Caribbean	0.044**	0.071***	0.014
	(0.018)	(0.026)	(0.026)
Inner London	-0.012	-0.022	-0.008
	(0.014)	(0.020)	(0.019)
Outer London	-0.018**	-0.046***	0.013
	(0.008)	(0.012)	(0.010)
North East	-0.075***	-0.067***	-0.079***
	(0.012)	(0.018)	(0.016)
Course Controls	X	X	X
Trainee Targets	X	X	X
Wage Gaps	X	X	X
Characteristics	X	X	X
Observations	76,651	44,192	32,459

Standard Errors in Parentheses. *p <0.10, **p<0.05, ***p<0.01.
All regressions control for year, region and subject fixed effects. Omitted categories: classification = First class degree, Region = South East, Ethnicity = White.