Essays in Applied Labour Economics

Rebecca Rose

London School of Economics and Political Science
October 1, 2025

A thesis submitted to the Department of Economics
for the degree of Doctor of Philosophy

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science (LSE) is solely my own work other than where I have clearly indicated that it is the work of others.

Chapter 3 is joint work with Tessa Hall (London School of Economics) and Alan Manning (London School of Economics). Chapter 4 is joint work with Rui Costa (London School of Economics), Magdalena Dominguez (Institute for Fiscal Studies) and Matteo Sandi (Catholic University of Milan and London School of Economics).

This thesis consists of **39,101** words.

This work was supported by the Economic and Social Research Council (ESRC) [grant number ES/P000622/1]

Abstract

This thesis consists of four main chapters that focus on the themes of occupational choice, wage gaps, and networks. Chapter 1 assesses 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. Larger bursaries increase both trainee recruitment and teacher cohort size three years post-training. However, the probability of becoming a teacher post-training also falls, which is driven by unobservable selection. Chapter 2 explores the heterogeneity of teacher wage gaps in England. I assess the comparability of teacher wages across sources and explore the robustness of estimated wage gaps for state-funded school teachers. I find substantial variation in wage gaps depending on the data, method, and counterfactual used, in addition to geographic inequality in teacher wage competitiveness. Chapter 3 describes how ethnic and migrant wage gaps vary across the life cycle. By exploiting newly linked UK administrative panel data, we estimate pay gaps on labour market entry and differences in pay growth. We find that the entry pay gaps are large, though they vary between groups. For most groups, pay gaps are largely preserved over the life cycle. For migrants, we find that the extra pay penalty is concentrated in those who arrived in the UK at a later age. Chapter 4 studies gangs in Brazil, an underexplored yet pervasive and volatile setting in organised crime. Using highly detailed information from intelligence, occurrence data, and prison records, we construct a network of gang affiliation and detect gang clusters using Markov stability analysis. We analyse the identified clusters through different network statistics and find a strong correlation between most individual centrality measures, while intercentrality highlights that some "key players" may have been missed in official hierarchical classifications.

Acknowledgements

I am grateful to my supervisors, Stephen Machin and John Van Reenen, for their invaluable guidance and support throughout my PhD. I also thank Alan Manning, whose advice and encouragement since my undergraduate years have helped shape my academic path. I have learned a great deal from my co-authors. I am especially thankful to Rui Costa, Magdalena Dominguez, and Matteo Sandi for their mentorship and to Tessa Hall for her collaboration and support.

I would like to thank the faculty, students, and researchers at the Centre for Economic Performance and the LSE Economics Department for their feedback, conversations, and generosity with their time and ideas.

This research would not have been possible without the financial support of the ESRC and LSE. I am especially grateful to the broader LSE community for shaping my academic and personal development over many years, and above all, to the incredible friends I met here.

Lastly, I would like to thank Michael Rosen for his most excellent book *We're Going on a Bear Hunt*, which provided me with essential advice for surviving my PhD and producing this thesis:

> We can't go over it.
> We can't go under it.
> Oh no!
> We've got to go through it!
> ...
> We're not going on a bear hunt again.

Contents

List of Tables

List of Figures

Chapter 1

Intention to Teach: Incentive Impacts of Bursaries

Teachers are an essential input of the education production function. However, at a time of global teacher shortages, there is limited research on the effectiveness of policies designed to attract more individuals into the profession. I assess the impact of front-loaded 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, a £10k increase in bursary leads to negative retention outcomes. Trainees are 2.4% less likely to appear as a teacher post-training, which is driven primarily by unobservable 'motivation' rather than observable personal characteristics. Results are primarily driven by stem trainees and are robust to controlling for an individual's outside option wage. Attracting an additional 5,000 FTE teachers by increasing bursaries costs approximately the same as raising all teacher pay. Raising training bursaries is a flexible tool to address teacher shortages, but leads to compositional effects that can impact the long-term motivation and occupation decisions of the resulting teacher workforce.

1.1 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 current 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, Hanushek, and Kain (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, Ladd, and Vigdor (2006), Rockoff (2004)). If there is an 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 chapter, 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 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

[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]

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 retention, 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. Disaggregating 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

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

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. 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 chapter is as follows: In section 1.1.1 I discuss the key contributions of this chapter. I then provide the context of the policy and discuss the data in sections 1.2 and 1.3 respectively. In section 1.4, I describe the empirical strategy. Section 1.5 presents the empirical form results and section 1.7 offers an explanation of these results using a simple model of occupational choice. I discuss robustness in section 1.8, policy implications in section 1.8 and conclude in section 1.9.

1.1.1 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 in 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

[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.

a one-time recruitment payment similar to the one studied in this chapter. 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. (2019, 2023) 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, Hattemer, and Sierra Vásquez (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, Tang, and Worth, 2023).

The relevance of bursaries as a recruitment tool cannot be understated in the presence of myopic or financially-constrained workers. Christian, Ronfeldt, and Zafar (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 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.

This chapter 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

[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.

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, Finan, and Rossi (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 had 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, Caria, and Ortiz-Ospina (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*a*) 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, Piopiunik, and West (2020) find that teachers who enter during a recession are more effective at raising student test scores. This suggests that a higher bursary 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.

1.2 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 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

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

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 instalments, 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 1.1b 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 waiting an additional year to train.

Figure 1.2 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. 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.

[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).

[8]Raised from £9,000 in 2017

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

1.3 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 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 a primary school teacher differs in many respects to secondary school teaching, and the

[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.

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.

1.3.1 Descriptive Results

I include 95,397 trainees across seven years. Table 1.1 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 1.3 shows the 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 1.4 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 1.4 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

[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].

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

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 1.5 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.

1.4 Empirical Strategy

1.4.1 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 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 DfE 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.1. 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

bursaries can.

[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-entries and exits, the fulfilment of previous year's targets, and changes in student populations

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 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.

1.4.2 Cohort-level Characteristics

Firstly, I explore the impact that financial incentives have on the size and characteristics of trainee cohorts. Equation 1.1 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

[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.

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 chapter, I discuss alternative methods of estimation to correct for potential biases.

$$log(N_{jt}) = \beta\overline{B}_{jt} + \alpha_j + \alpha_t + \gamma TraineeTarget_{jt} + \epsilon_{jt} \tag{1.1}$$

I also examine how bursary levels impact the observable characteristics of trainee cohorts. Equation 1.2 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.

$$C_{idjt} = \beta B_{djt} + \alpha_d + \alpha_j + \alpha_t + \gamma TraineeTarget_{jt} + \epsilon_{idjt} \tag{1.2}$$

1.4.3 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 1.3, 1.4 and 1.6, where the error terms are distributed logit. I estimate 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.

$$Y_{idjt} = \beta B_{djt} + \alpha_d + \alpha_j + \alpha_r + \alpha_t + \epsilon_{idjt} \tag{1.3}$$

$$\begin{aligned} Y_{idjt} =& \beta B_{djt} + \alpha_d + \alpha_j + \alpha_r + \alpha_t \\ &+ \gamma_1 TraineeTarget_{jt} + \gamma_2 \widehat{WageGap}_{idjt} + \epsilon_{idjt} \end{aligned} \tag{1.4}$$

$$\widehat{WageGap}_{idjt} = \hat{w}_{out,idjt} - \hat{w}_{teach,idjt} \tag{1.5}$$

Regression 1.4 estimates the overall effect of the bursary level on retention. In equation 1.4 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.1. 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].

$$\begin{aligned} Y_{idjt} =& \beta B_{djt} + \alpha_d + \alpha_j + \alpha_r + \alpha_t \\ &+ \gamma_1 TraineeTarget_{jt} + \gamma_2 \widehat{WageGap}_{idjt} + \\ &+ \gamma_3 Characteristics_{idjt} + \epsilon_{idjt} \end{aligned} \tag{1.6}$$

Regression 1.4 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 1.6 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 1.4 will identify both effects, whereas the regression1.6 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

[19] In robustness checks, the addition of the current wage does not have an impact on the bursary level coefficient. In section 1.6 I also discuss how the use of different wage estimations impacts my results.

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 determine 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.

1.5 Empirical Results

1.5.1 Cohort Characteristics

Table 1.2 shows the impacts of the average bursary level on trainee cohort size, where column 1 estimates equation 1.1. 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. In STEM, a bursary uplift of £10k 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 1.2 can be seen in table 1.3. 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

[20]Calculated as the exponent of the regression coefficient

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.

1.5.2 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 1.3, 1.4 and 1.6 in tables 1.4 to 1.6. 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 1.4 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 when 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 1.5 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

[21] Note that controlling for outside wages decreases the coefficient on wages 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 1.4 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.

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 1.3 and A.6, 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.

I also estimate the impact of financial support on the likelihood of appearing as a teacher at each year post-training in table 1.6. 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 A.2 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 1.6c 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.

1.5.3 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. In Chapter 2.11 I analyse the dynamics of teacher wages versus other graduate wages in detail, and construct a measure of the teaching wage gap. A brief description of how I apply this measure to my trainee teacher sample is also included in the appendix section A.1.

Figure 1.6 represents the predicted 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. 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].

I control for the wage gap upon entry into training as defined in equation 1.5 on the probability of appearing post-qualification in the fourth column of tables 1.5 and 1.6. If weekly teaching wages become £100 more competitive (around one standard-deviation) during the training year, trainees are 4.5 percentage points more 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

[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, 2025).

[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

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 1.6). 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 1.7, I investigate the regional impact of the bursary level on appearing post-qualification by running regression 1.6 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 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 1.4, 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 1.7 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.

1.6 Robustness

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 1.7 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.

$$log(N_{djt}) = \beta_d B_{djt} + \alpha_{dj} + \alpha_{dt} + \gamma_d TraineeTarget_{jt} + \epsilon_{djt} \tag{1.7}$$

The results of equation 1.7, which disaggregates the impact of a bursary increase at the subject-classification level, are shown in table A.3[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 A.3b and A.3c, 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 1.2.

Using an average bursary level as the independent variable in regression 1.1 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 1.7 with a pooled sample of all subject-classes to estimate the average impact of the bursary across all classifications. The results, in table A.4, suggest that the impact remains significantly positive for total trainee cohort size, but the magnitude of estimates are reduced. The impact on the number of STEM teachers is also no longer significant.

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

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

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 1.7. Table A.5 shows that the statistically significant coefficients decrease in magnitude, but remain significant. The coefficient for the number of teachers remaining in all subjects also becomes significant at 14.6%.

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.

Cohort Characteristics

Table 1.3 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 1.1 and 1.7, where N_{jt} is the count of individuals that possess 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.

[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.

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 1.4, 1.5 and 1.6 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 1.5.

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.

1.7 Modelling Occupational Choice

I develop a simple model to interpret the results presented in section 1.5 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 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.

$$U_{it}^j = w_{it}^j + d_{it}^j m_i \tag{1.8}$$

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[26]. For simplicity, labour demand is not modelled and the parameters of the wage function are considered to be exogenous.

$$w_{it}^j = f_j(q_i, e_{it}^j) + \epsilon_{it}^j \tag{1.9}$$

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

$$V_{it}^j = U_{it}^j + \beta E_{max}\left[\{V_{it+1} | d_{it}^j\}\right] \tag{1.10}$$

$$d_{it} = \begin{cases} 1 & \text{if } V_{it}^T > V_{it}^O \\ 0 & \text{if } V_{it}^T < V_{it}^O \end{cases} \tag{1.11}$$

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[27].

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

[27] 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.

Teacher wages vary less with respect to undergraduate classification compared to other occupations, and better qualifications are rewarded more in occupations other than teaching[28]. 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.

1.7.1 Model impact of Bursaries

Teacher training is considered as the first year of teaching. Therefore, the bursary amount $B_{i,t}$ 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 when the bursary level is varied. I summarise the key results of this exercise below, and more details can be found in the appendix A.2.

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 for the individual to 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 may 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'.

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

Model results

Due to the dynamic structure of the model, I produce simulations to observe the decisions of 450,000 individuals. I use the outcomes of these simulations to describe six results of the model below. The simulation process is described in more detail in appendix section A.2.

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 1.7, and aligns with the results in section 1.5. Figure 1.10 also shows that a larger bursary attracts a larger share of the population to train in year t=1. Figure 1.11 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 1.10, 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. Figure 1.7 assumes a theoretical normal distribution of motivation. 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, then increases till the motivation threshold is at the mean level of motivation, and then decreases on the left hand side of the distribution. By observing the gap between each line in figure 1.11, we see that the marginal impact of the bursary in the model simulations is largest for the median bursary amounts.

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[29]. In simulation figure 1.10, we see that the shares of individuals in teaching stabilises for all bursary levels by around year 10. Figures 1.10, 1.11 and 1.12 demonstrate that the share of the population engaged in teaching is higher for larger bursary levels any number of years later.

Result 3: Composition effects based on qualifications

[29]Though not always significant.

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, Walker, and Waltmann, 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 1.8 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 1.12 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, Walker, and Waltmann, 2022). Section 1.6 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[30]. The model simulations can also reflect this by assuming that STEM graduates have a higher proportion of trainees with high starting wages.

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.

Figure 1.13 replicates the survival rates in the data shown in figure 1.4 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-

[30]Note that these coefficients were significant, but not statistically significant from the estimates for other classifications.

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 1.10 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%.

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 1.9 confirms that the average motivation for all trainees, as well as for converts only, is lower at higher bursary amounts.

Result 6: Retention varies by outside option

Retention of trainees depends on their likelihood of being a dropout 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 1.11 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[31]. 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 1.15 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 1.5, 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

[31]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.

higher classifications in their undergraduate degrees are generally less likely to remain in teaching due to higher outside wages[32]. Whilst this is in line with most findings in the literature[33], this is not reflected in the empirical results (see figure A.6). 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 1.15 as the gap between two neighbouring bursary levels increases as the wage level increases.

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 depend on the outside wages and underlying distribution of motivation of the potential trainees. In light of these findings, in section 1.8 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 [34]. The wage schedule for

[32] This can be seen by comparing the retention rate in figure 1.12 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.

[33] For example, see Podgursky, Monroe, and Watson (2004), Stinebrickner (1998)

[34] 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.

This exercise has presented one potential explanation for the empirical results observed, however there are other possible mechanisms that could generate similar outcomes. The most important mechanism to consider is whether the change in recruitment and retention is generated by alleviating financial constraints. If individuals are unable to forego a year of income to train, a bursary will enable an increase in cohort size. However, this would imply that the marginal recruit is highly motivated and financially constrained. In this case, we would not anticipate an increase in attrition since these individuals would be more incentivised to remain employed in teaching following completion of the training.

1.8 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, Tang, and Galvis (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 1.12 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 training 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 that they can be easily targeted at specific subgroups of trainees, whereas varying teacher pay by subject is more difficult to implement politically. Table 1.13 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[35]. 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 possible 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

[35] 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.

also are subject to diminishing returns[36]. 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.

1.9 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 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

[36] See previous footnote.

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.

1.10 Tables

Table 1.1: 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 1.2: Cohort Size Pooled 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 1.3: 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 1.4: 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 1.5: 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 1.6: 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 1.7: 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 1.8: 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 1.9: 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 1.10: 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 1.11: 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 1.12: 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.

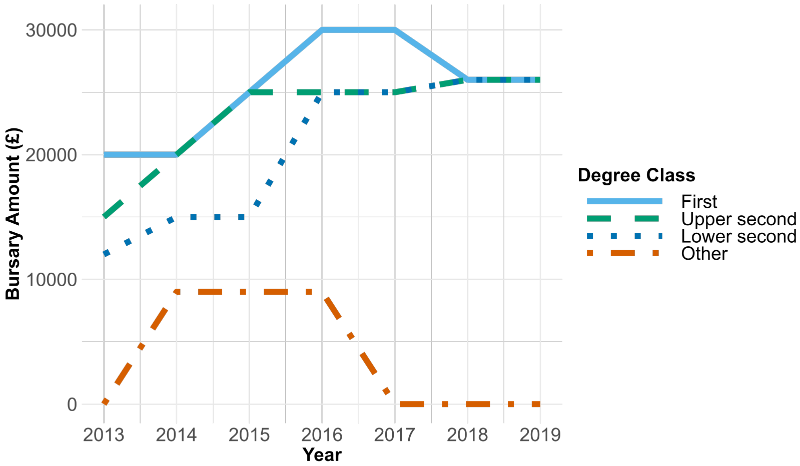
Table 1.13: Cost-Benefit of a 10 Percent Increase in Trainees

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

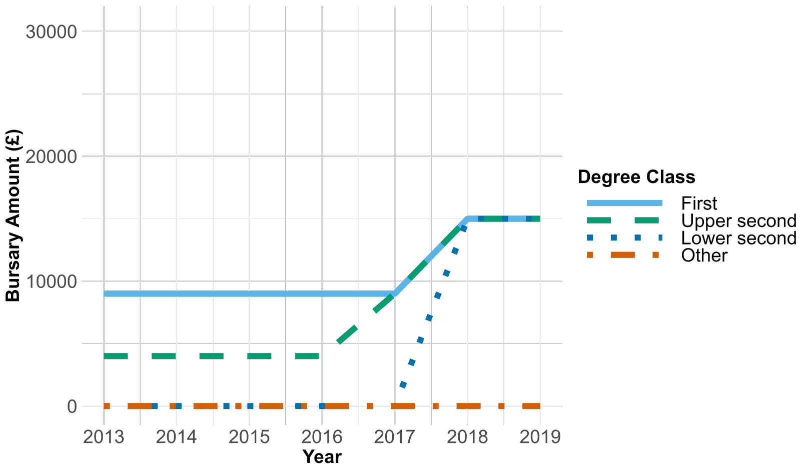
Figures are evaluated for a 10% increase of the specified cohort size from the existing average bursary level for that group in the 2022/2023 training cycle. The group refers to the specific training cohort that would be given a bursary uplift. Best case estimate uses the upper bound of coefficients,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

1.11 Figures

Figure 1.1: 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 1.2: Training Timeline

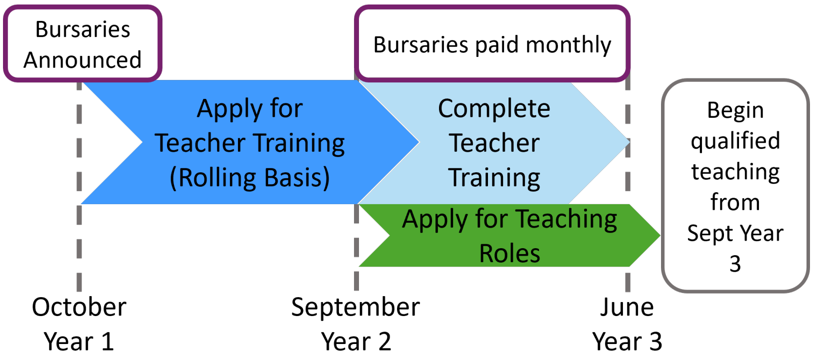
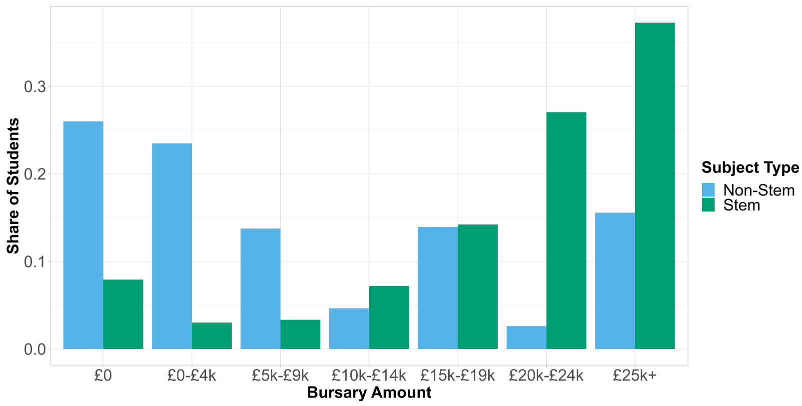
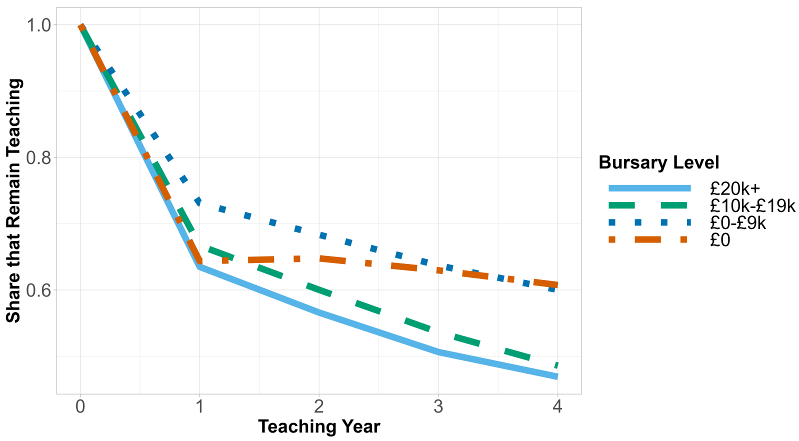


Figure 1.3: Distribution of Bursary Awards



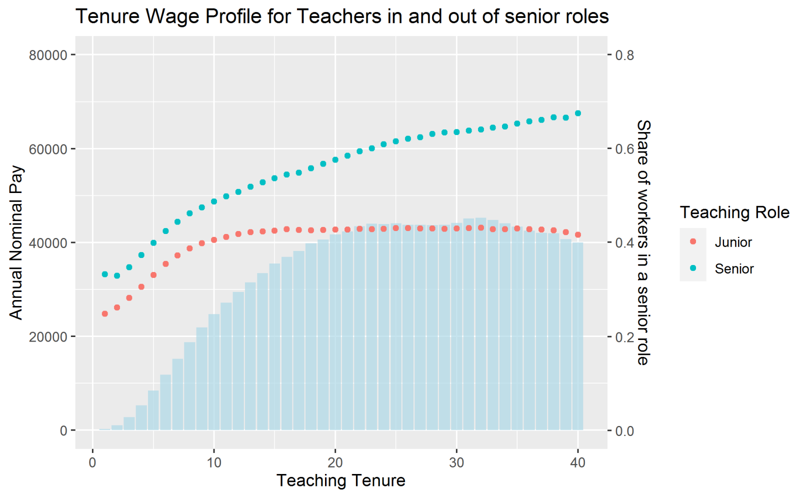
Initial Teacher Training (ITT) Data, 2013-2020, DfE. Bursary level is matched to trainees based on their characteristics and subject chosen.

Figure 1.4: Retention Rate by Level of Bursary Award



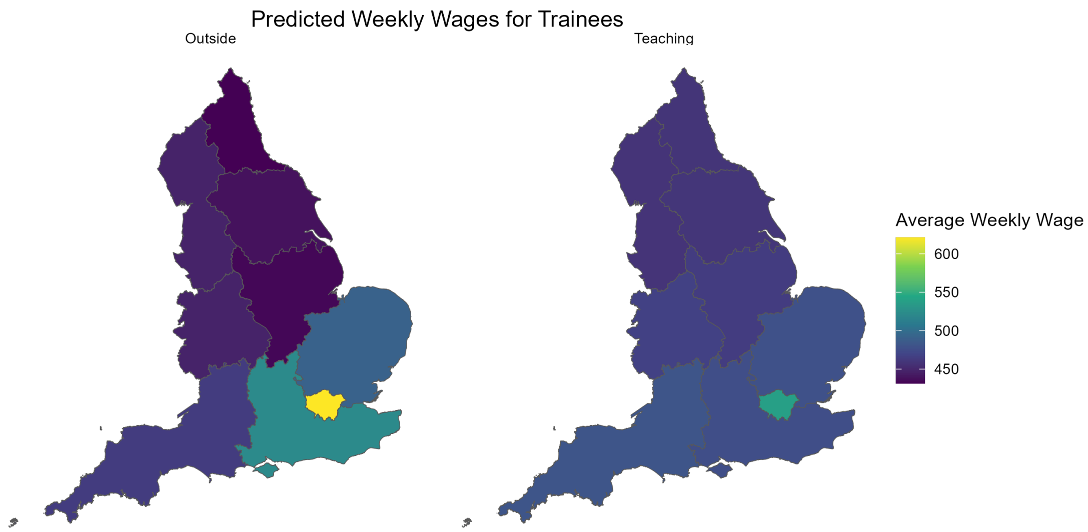
Data: School Workforce Census (SWC) and Initial Teacher Training (ITT) Data, 2013-2020, DfE. An is teaching year in x if they are present in the school workforce census x years after their training.

Figure 1.5: Average Teacher Salaries over Teaching Tenure



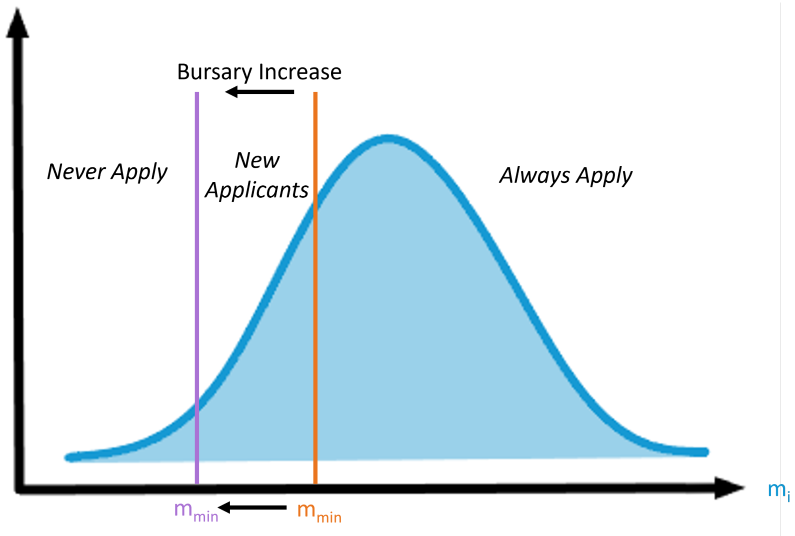
Data: School Workforce Census (SWC) 2013-2020, DfE. Blue bars represent the share of the teaching workforce at each tenure level that holds a senior role. Teaching tenure is calculated as years since qualified teacher status was attained. Senior staff roles are heads of department, heads of year, deputy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

Figure 1.6: Wages for Teachers and their Outside Options



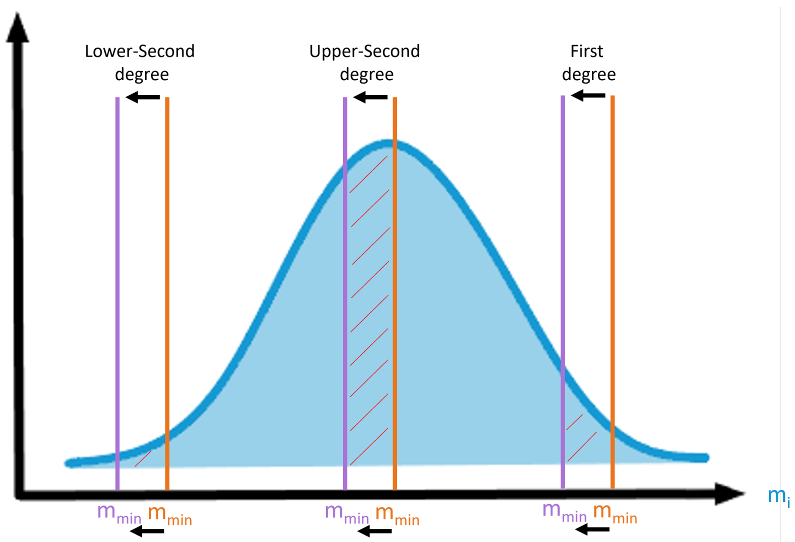
Data: SWC, ASHE, and LFS. Teaching wages are the starting wage predicted via a regression of using DfE data. Outside wages are predicted for the year of entry based on age, region, year, sex, and undergraduate characteristics. The predictions are generated using ASHE data, where jobs are weighted according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

Figure 1.7: Model Result 1: Bursary Raise Increases Trainees



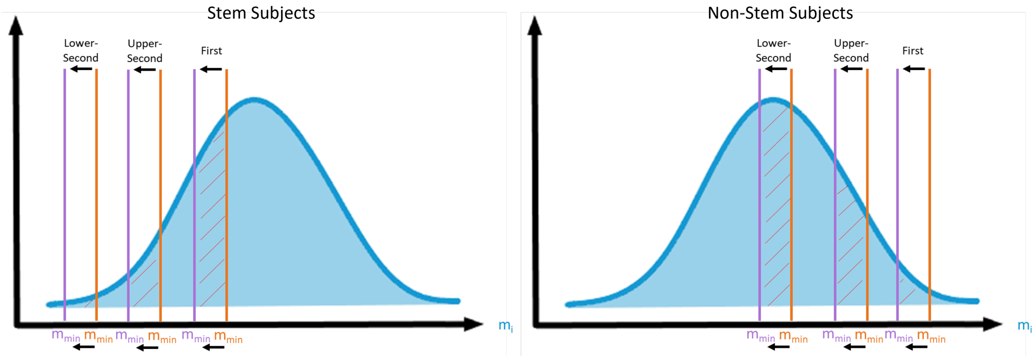
Red shaded area denotes the individuals induced to train due to a marginal increase in the bursary level.

Figure 1.8: 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 1.9: 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 1.10: Simulation Result: Share in Teaching over Time

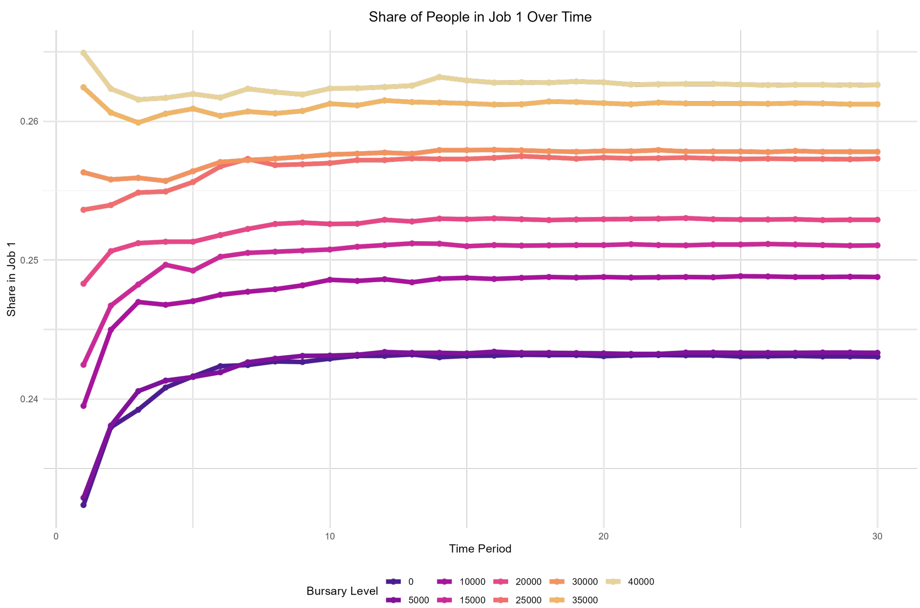


Figure displays the simulated share of individuals engaged in teaching each year over the 30 year time period. Trends for groups with different bursary offers are displayed with different colours.

Figure 1.11: Simulation Result: Remaining Cohort Size over Tenure by Bursary Level

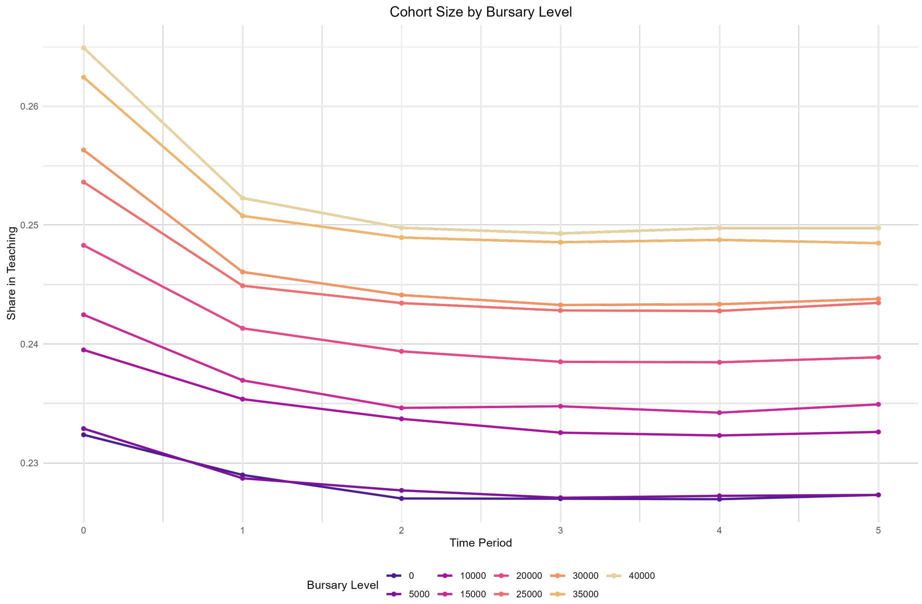


Figure displays the simulated share of individuals remaining in teaching each year after training. Time period here represents the number of years since an individual entered teacher training. Individuals who never trained are included in the denominator. Trends for groups with different bursary offers are displayed with different colours.

Figure 1.12: Simulation Result: Cohort Size by Bursary Level and Wage

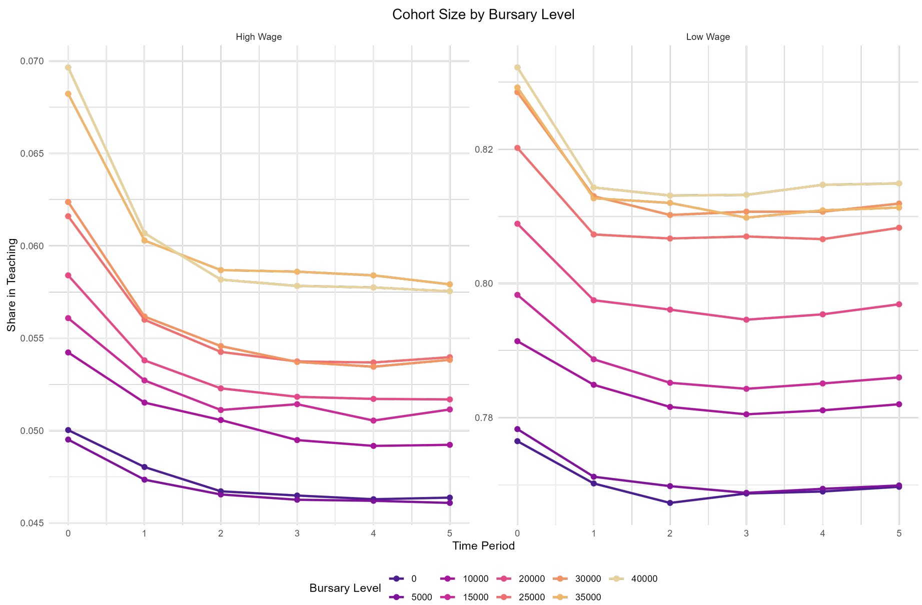


Figure displays the simulated share of individuals remaining in teaching each year after training. The two graphs represent those whose starting outside wage is higher or lower than the starting wage in teaching. Time period here represents the number of years since an individual entered teacher training. Individuals who never trained are included in the denominator. Trends for groups with different bursary offers are displayed with different colours.

Figure 1.13: Simulation Result: Retention by Bursary Level

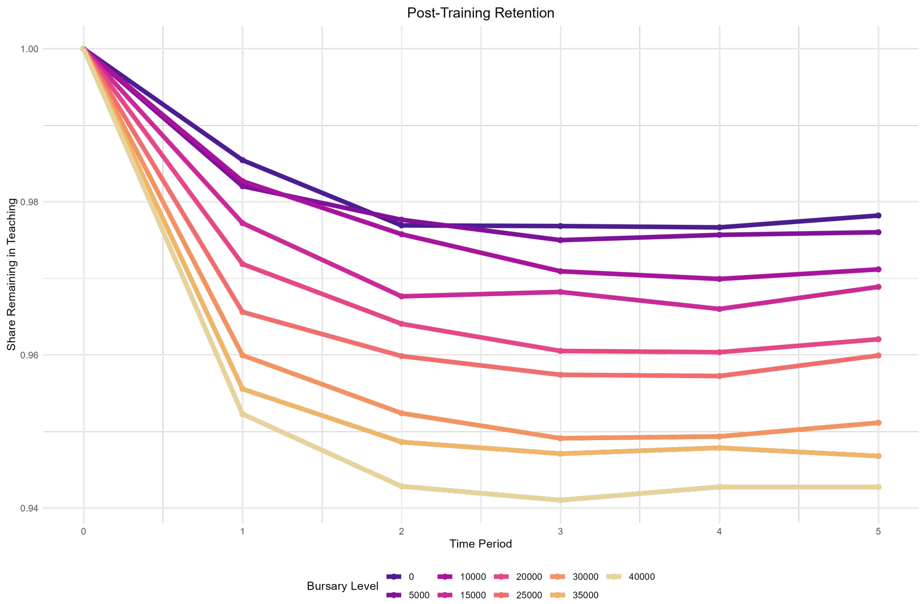


Figure outlines the share of individuals remaining in teaching each year after training. Time period here represents the number of years since an individual entered teacher training. Trends for groups with different bursary offers are displayed with different colours.

Figure 1.14: Simulation Result: Retention by Wage

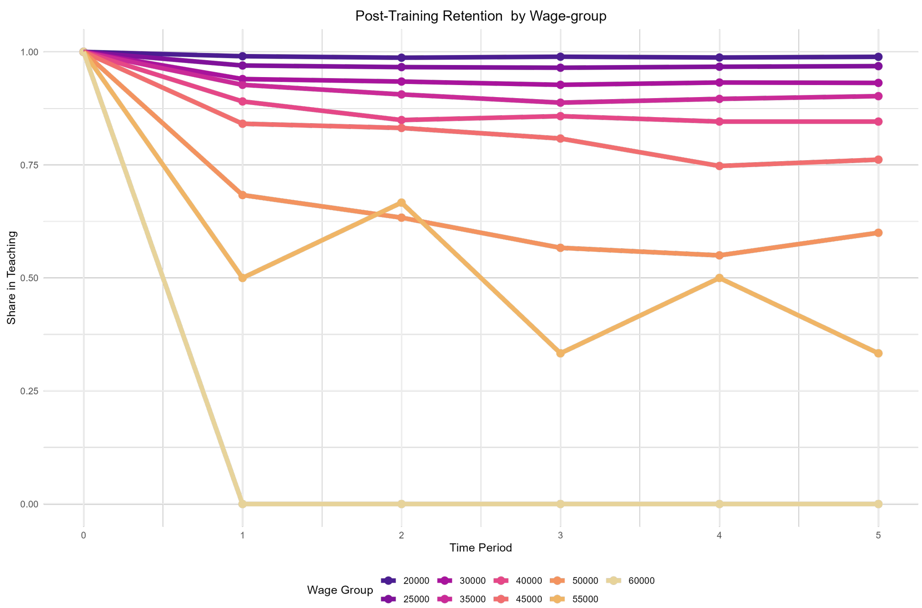


Figure outlines the share of individuals remaining in teaching each year after training. Time period here represents the number of years since an individual entered teacher training. Trends for groups with different starting outside wage offers (excluding random components) are displayed with different colours.

Figure 1.15: Simulation Result: Retention by Bursary Level and Wage

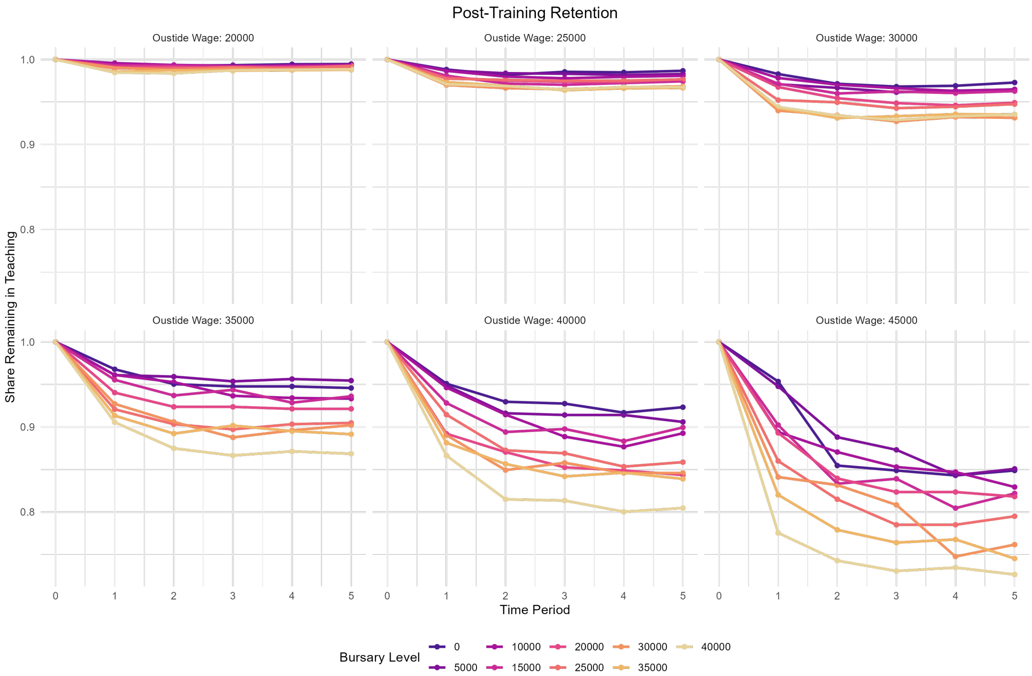


Figure outlines the share of individuals remaining in teaching each year after training. Time period here represents the number of years since an individual entered teacher training. Trends for groups with different bursary offers are displayed with different colours. Trends for groups with different starting outside wage offers (excluding random components) are displayed using different sub-charts.

Chapter 2

Estimating Teacher Wage Gaps

Understanding teacher wage gaps is critical in order to address England's ongoing teacher recruitment and retention challenges. Yet estimates of these wage gaps vary considerably depending on how the counterfactual is defined and which data source is used. This chapter compares teacher wages to outside options using three major UK datasets: the Labour Force Survey (LFS), the Annual Survey of Hours and Earnings (ASHE), and the School Workforce Census (SWC). I assess the comparability of teacher wages across sources and explore the robustness in estimated wage gaps for state-funded school teachers between 2013 and 2019. I find substantial variation in estimated wage gaps depending on the data, method, and counterfactual used. Teachers appear to earn a statistically insignificant 5.6% wage premium in the LFS relative to similar graduates, while ASHE-based estimates range from a 2.3% penalty to a 33.5% premium, depending on the comparison group. Wage gaps are least competitive in regions with large cities, particularly in the South East, though age-related trends are less consistent. These findings underscore the importance of carefully defining counterfactual wages in policy discussions around teacher pay.

2.1 Introduction

Understanding teacher wage gaps is critical to addressing England's teacher retention and recruitment crisis. But how underpaid are teachers compared to their outside options? The answer depends not only on how the outside option is defined, but also the data used to estimate it. Studies of teacher's wages across the years have generated markedly different conclusions on the size and sign of these wage gaps. This chapter estimates teacher wage gaps using three different UK data sources and explores the robustness and extent of variation in the results.

This work has three main contributions. Firstly, it produces a variation of estimates of teacher wages using different counterfactual groups. Having an understanding of why wage gaps vary, both within teachers and across datasets and methods, is essential for tackling issues of retention within the current teaching workforce. Secondly, I exploit two separate and rich sources of employment and wage data in England, the Annual Survey of Hours and Earnings (ASHE) and the Labour Force Survey (LFS), and assess their comparability with the wages reported in the School Workforce Census (SWC). The SWC itself is a relatively new data source containing detailed information on the full population of state-funded school teachers in England, offering exciting new research opportunities. It is therefore important to understand how it compares to other major data sources. Lastly, this exercise provides a foundation for estimating the wage gaps of trainee teachers in chapter one.

The analysis unfolds in three stages. I first compare the consistency of weekly wages for English state-school teachers across the LFS, ASHE, and SWC data sources for the years 2013 to 2019, and explore their variation across ages and regions. Next, I use conditional wage regressions in the LFS and ASHE to generate a benchmark of the wage gap for teachers. I then establish four detailed measures of counterfactual wages for teachers, and predict wage gaps for each individual within the teaching workforce. These four counterfactuals represent graduates in the LFS, two separate estimates of graduates in ASHE, and ex-teachers in ASHE.

Reported weekly teacher wages in the LFS and ASHE are very similar in terms of average and distribution, and both are a close approximation of the weekly wages in the SWC. However, average wages are slightly higher in the SWC than in the LFS and ASHE, particularly for ages 35 and over. ASHE continues to roughly approximate the SWC when examining hourly wages, but the LFS reports much lower hourly wages since hours are self-reported by the employees.

Varying the data source and method used to calculate the log wage gap produces very different estimates. The LFS predicts that teachers experience a positive but statistically insignificant wage premium of 5.6% compared to graduates of the same age in the same year and region. Using ASHE data, teachers experience a negative but statistically insignificant wage penalty of 2.3% compared to similar individuals employed in jobs where the majority (over 80%) are graduates. However, ASHE predicts a positive, statistically significant wage premium of 14.3% when the counterfactual is occupations weighted by the share of graduates employed in that industry-occupation. When using occupations that teachers typically move into after teaching as a counterfactual, ASHE predicts a teaching wage premium of 33.5%. London and the South East are generally the least competitive across each measure of the wage gap. This is consistent with teaching wages being relatively more uniform across space. Trends across age groups are less consistent across data sources. There were also no major differences in wage gaps between teachers who eventually left teaching and those who remained as teachers for the entirety of the data period.

Each measure of wage gaps outlined in the paper is relevant for a different purpose. Using the jobs of ex-teachers as a counterfactual is particularly relevant for those whose undergraduate degrees specialise in education, or for more senior teachers whose relative human capital in other occupations is lower. In contrast, the graduate counterfactuals are much more relevant when considering the occupation choices of junior teachers, or prospective teachers, who have degrees in other subjects. In any case it is important to consider that the pay differential between teachers and other graduates is not the only element that determines the competitiveness of teaching. Green (2021) notes that many elements of job quality in teaching have declined. When taking this into account, pay parity between teaching and other occupations may not be sufficient to attract and retain good quality teachers within schools.

The rest of the chapter proceeds as follows: in section 2.2 I review the relevant literature related to teaching wages and wage gap estimation. In section 2.3 I discuss each of the data sources and compare the distribution of teacher wages reported by each of them. Section 2.4 outlines three methods of identifying the counterfactual wages of teachers within ASHE, and compares these with the graduate wages estimated in the LFS. In section 2.5 I estimate teacher wage gaps using conditional wage regressions and in 2.6 I outline my method for estimating more detailed wage gaps that are representative for the population of teachers in the SWC. Finally, section 2.7 describes the results and 2.9 concludes with a discussion.

2.2 Literature Review

Chapter one established that retention for UK state-school teachers is an ongoing issue, with Fullard and Zuccollo (2021) demonstrating that teacher retention has also been worsening over time. To understand these patterns, we can consider both the monetary compensation of teachers and their experience in the workplace. Green (2021) notes that in the last 30 years, teachers have reported increased work intensity and decreased control over their work. In a recent teacher's union wellbeing survey, 86% of teachers reported that their job had an adverse impact on their mental health in the last 12 months, and 68% claimed it had an adverse affect on their physical health over the same time period (The National Association of Schoolmasters Union of Women Teachers, 2024).

Declining job conditions in teaching imply that wages need to become increasingly competitive in order to continue to attract and retain teachers in the occupation. However, relative teaching wages in England have generally declined over time. As outlined in chapter one, Sibieta (2023) found that real wages for (in)experienced teachers fell (5) 13 percent in real terms between 2010 and 2022. This is not unique to England: Allegretto (2022) found that real weekly wages of teachers in the US have also been relatively flat since 1996. Whilst not directly evaluated in this chapter, pensions are a key form of compensation that differ substantially between the private and public sector. Disney, Emmerson, and Tetlow (2010) establish that teacher pensions remain competitive over a typical private sector pension, but that pension changes in 2007 reduced the value of these pensions. Dolton, Samek, and She (2019) found that further changes in 2015/16 reduced the total reward of teaching in the UK over the life-cycle.

Research demonstrates that wages are an effective tool to influence both prospective and current teachers. In a comparison across countries, Hanushek, Piopiunik, and Wiederhold (2019) found that higher teacher wage premiums were associated with greater cognitive skills of teachers. Falch (2017) observes that in Norway, a teacher pay premium of 10 percent increased the recruitment rate by 6 to 7 percentage points. Biasi, Fu, and Stromme (2021) also note that pay also affects the allocation of the existing teacher stock: higher-performing teachers moved to higher-paying schools when Wisconsin districts were able to vary pay. Dolton and Van der Klaauw (1999) find that within the UK, the relative wages of teachers are also an important determinant of the turnover of teachers. Chevalier, Dolton, and McIntosh (2007) analysed long run recruitment trends in the UK from 1960 to 2002 and found that the decline in the relative wages of teachers was associated with a decline in the relative attainment of graduates entering into the occupation. Finally, Nickell and Quintini (2002) observed that a decline in the relative pay of male teachers in the 70s and 80s reduced the relative test scores of teachers.

To estimate relative teacher wages, it is necessary to establish a valid counterfactual of alternative outside wages. A traditional method is to use a Blinder-Oaxaca that directly compares teachers with other graduates of similar characteristics (e.g. in Allegretto (2022)). However, teachers may not be directly comparable to the average graduate who chose an alternative occupation. For example, Podgursky and Tongrut (2006) note that wage gaps narrow when accounting for the fact that teachers work for irregular periods across the year. Another issue is the question of selection bias: whether individuals who choose to be teachers differ to those who select other occupations. Dolton (1990) and Chevalier, Dolton, and McIntosh (2007) develop a method to address such selection bias using instruments that affect the likelihood of becoming a teacher but are orthogonal to teacher pay, and therefore compare individuals with similar propensities to become a teacher. Another method involves observing the wages of teachers, or those qualified to teach, who leave for alternative jobs (Stinebrickner (2002), Scafidi, Sjoquist, and Stinebrickner (2006), Worth and McLean (2022), Goldhaber et al. (2024)). Building on this method, Tsao (2025) uses a fuzzy regression kink method to identify the difference in wages for those who narrowly missed out on teaching certification in Kentucky.

Measurements of teacher wage gaps vary a great deal by context and method. When comparing the pay of teachers to the pay of ex-teachers, numerous studies find teacher premiums, whether in the US (Stinebrickner (2002), Scafidi, Sjoquist, and Stinebrickner (2006) or the UK (Worth and McLean, 2022). Tsao (2025) estimates teacher pay in Kentucky to be 33–40% higher compared to a teacher's next-best job, and Goldhaber et al. (2024) similarly find that teachers in Washington are paid more than those qualified to teach who chose alternative occupations. In England, Chevalier, Dolton, and McIntosh (2007) calculate a wage premium of 10-15% for female teachers but a wage penalty of 11-19% for male teachers in teaching cohorts that graduated between 1960 and 1995, and that the competitiveness of teacher wages had increased over time. Studies that compare the wages of similar teacher and non-teacher graduates more directly typically find teacher wages to be less competitive: A recent study in the US by Allegretto (2022) estimates a pay penalty of 23.5%, whereas in the UK the penalty has been estimated to be 6.8% between 1993 and 2019 (Fullard, 2021*b*), 30% between 1996 and 2006 (Britton and Propper, 2016), and 18-25% in 2013 (OECD, 2015).

2.3 Data

I use three data sources in this analysis: The Labour Force Survey (LFS), The Annual Survey of Hours and Earnings (ASHE), and The School Workforce Census (SWC). In the following section, I explain each of these sources and outline why they are beneficial for estimating the teaching wage gap. For each data source, I have access to data between 2013 and 2019. I also limit my analysis to workers in England, to keep pay scales experiences by teachers consistent. This is also consistent with the analysis in chapter one.

The Labour Force Survey (LFS) is a study of employment outcomes across the UK. It collects quarterly data across households and records a wide range of information on personal characteristics, training and qualifications, employment, and job characteristics. This is particularly useful for analysing the counterfactual wages of teachers since the majority of teachers have degrees[1]. I also observe the region of work and residence. The LFS is a valuable dataset when analysing pay and personal characteristics, but is vulnerable to the typical biases of survey data. Of particular relevance to this study is response bias, and the fact that higher incomes are top-coded for anonymity. I limit my sample to UK nationals to maintain consistency with chapter one.

[1]As of 2014, near the start of the time window of the data, 96% of teachers held a degreeDepartment for Education (2014).

The Annual Survey of Hours and Earnings (ASHE) is a panel containing a sample of 1 percent of all employed individuals in the UK. It includes wage and employment data extracted from HM Revenue & Customs Pay-as-you-earn records, reported by employers each April. ASHE is a useful data source due to its larger sample size[2], and reduced reporting bias compared to the LFS. As ASHE is a panel, it also allows us to observe the occupations of ex-teachers. However, it excludes information on qualifications and many personal characteristics (such as nationality and ethnicity), which limits the ability to compare like-for-like. Most crucially, observations must be weighted within ASHE to account for the absence of qualifications.

Lastly, I use The School Workforce Census (SWC), which contains data provided by schools each November. It outlines the hours worked, income, and details about their employment for all individuals who are directly employed by, or predominantly work within, state-funded schools in the UK. This includes all teachers, teaching assistants, and administrative staff employed in any publicly funded nursery, primary school, or secondary school. This expands the sample size of teachers to a huge degree and makes it fully representative: As can be seen in table 2.1, the number of teachers observed in the census over the time period is over 100 times larger than the sample in ASHE and approximately 95 times larger than the sample in ASHE. The census also records the main roles of each individual, their teaching qualifications, and the subjects they teach. Individuals are assigned a unique identifier and can be followed across time and schools. My data gives me access to the age of the individual and the location of their school at the ward level, but I do not observe their sex, residence, or ethnicity[3]. For a subsample of teachers, I infer whether they are a STEM specialist (Science, Technology, Engineering and Maths) based on the subject they teach for the majority of their time. This data is predominantly available for secondary school teachers. Lastly, the census records the date that each individual attained their Postgraduate certificate of education (PGCE), so I define teaching tenure as the time elapsed since this point.

Table 2.1 describes the total observations in each source. When comparing wages within and outside teaching, I focus on full-time workers between the ages of 21 and 59 inclusive. The age of 21 is considered a minimum as this is typically the earliest age an individual will have finished an undergraduate degree. I compare the salaries of the universe of teachers based on age, year, STEM status, and location. I am unable to compare based on sex as I do not observe it in my SWC data. For both sources, I compare wages at the smallest possible common spatial unit. For the LFS, I compare wages at the region level, whereas for ASHE I compare wages at the NUTS3 level, which is roughly equivalent to sub-regions.

[2]Note that in table 2.1, the LFS has more observations, but this includes both employed and unemployed individuals,

[3]Sex and ethnicity are recorded in the SWC, but were not provided for this study.

2.3.1 Teacher wages

As a first step to estimating the wage gap of teachers, it is important to select a reliable measure of wages. The SWC reports the annual earnings of teachers (both gross pay and base pay), the weeks worked per year, and the hours worked per week. ASHE reports their derived annual, weekly, and hourly pay (both gross and base), and LFS reports derived annual and hourly pay (both gross and base). Both LFS and ASHE report base hours and overtime worked per week. For each data source, I estimate weekly and hourly gross wages[4]. I use gross wages as individuals are likely to consider total income when comparing occupations, and overtime and bonuses may matter more for some jobs than others.

Figure 2.1 shows the distribution of weekly teaching wages across the three sources using kernel density estimation. They show that both ASHE and LFS reasonably approximate the SWC for weekly wages, particularly in the shape of the right hand side tail. However there are some key differences: the left tail for the SWC is much shorter, and there is a higher level of bunching around certain wage values which generates spikes in the plot. The reason for this bimodal distribution could be surmised based on observation of teacher pay scales and figure 1.5; Most teachers start on the low end of the pay scale, and then their salary progresses over time. As senior teaching roles are limited, those remaining in junior teaching may be stuck at the higher spine points of the junior teaching pay scale, whereas senior teachers continue to see their pay increase, generating the long right tail.

The distribution of hourly wages is shown in figure B.1. The ASHE continues to mirror the SWC, but the LFS has a distinctly different distribution and is shifted to the left of the SWC. This suggests that the hours worked reported by the individual in teaching systematically differ from those reported by the employer, in this case teachers may report working more hours of unpaid overtime. For this reason, I use weekly wages in all the remaining analysis to keep the LFS comparable to the other two sources, and consider unpaid overtime as a non-pecuniary element of work. Table 2.2 reports the average wages of teachers in each sample. LFS and ASHE estimate average teacher wages to be £712 and £713 respectively, however the SWC estimates the average teacher wage to be £733. When comparing teacher wages to other graduate wages, using the SWC could overestimate the relative pay of teaching. Since the distributions of wages are otherwise similar across sources, I will calculate wage gaps separately using just ASHE and just LFS data. This ensures that wages were collected in a comparable way for both teachers and other graduates.

I disaggregate the wages of teachers across age and region in figures 2.3 and 2.4. Teacher wages tend to progress steadily for the first ten to fifteen years and then are relatively stable. The standard deviation of wages is also comparable across sources,

[4] Weekly wages are calculated as weekly wages divided by weeks worked per year. Hourly wages, where missing, are calculated as weekly wages divided by hours. For each sample, the highest and lowest 0.5% of values are dropped. This method is chosen over windsorising as the majority of these extreme observations are deemed to be entry or reporting errors rather than representing true wages.

despite the SWC having a much larger sample size. ASHE wages mirror SWC wages particularly well for early-career teachers, however both sources underestimate wages later in the career. Spatially, teacher wages are relatively uniform across the country, with the exception of inner and outer London that are subject to a higher pay scale. ASHE and LFS show greater levels of variation across regions, but are not consistent with each other. For example, whilst both maps show Liverpool to be a high wage region, ASHE predicts Bristol and the south east of England to have lower wages, but the LFS estimates relatively higher wages. None of the three maps suggest that wages are systematically higher in the south of England, which runs contrary to graduate wages.

2.4 Graduate wages

In the above section, I explored the dynamics of teacher wages. However, before I can replicate the same statistics for graduates, I need to identify likely graduates within ASHE because it does not include information on qualifications. I introduce three measures that are potential counterfactual groups for teachers: The binary graduate measure, the continuous graduate measure, and the leavers measure. The first three measures leverage the LFS to identify which occupation-sector pairs graduates are found in. I 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_{so}: the share of jobs that are held by graduates. 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 use these shares to assign weights to each job in ASHE with the same industry-occupation as follows:

- **The continuous graduate measure**: weight $w_{so} = P_{so}$, the share of graduates in occupation-sector so.
- **The binary graduate measure**: weight $w_{so} = 1$ if graduate share in occupation-sector so is greater than 80% ($P_{so} > 0.8$), and weight $w_{so} = 0$ otherwise.

The final measure, the 'leavers' measure, takes advantage of ASHE's panel structure to identify occupation-sectors that teachers are most likely to enter after leaving teaching. I identify a sample of 1,829 observations for 738 ex-teachers. I calculate the share of ex-teachers found in each occupation-sector, where each individual is given an equal weighting. These shares serve as the final set of ASHE weights, with any share that is less than half a percent set to zero. Table 2.3 reports the eleven most common occupation-industry combinations[5]. We can see that most ex-teachers stay

[5]The top eleven are chose as these are the only combinations with a share or one percent or greater.

in a job close to school teaching: 24% remain in the same general occupation (Teaching and other educational professionals), and 37% remain within the education sector. The only occupation-industry combination that houses over 1% of ex-teachers and is unrelated to teaching is Caring personal service occupations in the human health and social work activities sector. Lastly, note that the second most popular occupation-industry pair is teachers and other educational professionals in the education sector, who house 13% of ex-teachers. This includes, but is not limited to, teachers in private schools.

Having identified four sets of weights, I can apply these to the ASHE data to estimate the average graduate outside wage. As the LFS contains data on qualifications, I can observe wages of graduates directly and apply LFS survey sampling weights. Figure 2.5 summarises the distribution of graduate wages using LFS and the estimated distribution of graduate wages in ASHE using the continuous and binary graduate weights[6]. Each distribution varies, with the ASHE binary measure being more negatively skewed than the continuous measure. The LFS measure is roughly between the two, but has a mass at around £2000 which is caused by the top-coding of annual wages. According to the LFS, the average full-time graduate earns £757 a week, whereas the ASHE continuous measure estimates average earnings of £703 a week.

Figure 2.6 displays wages by age for the LFS and ASHE continuous weighted samples. Both average wages begin at a similar level at age 21, however the ASHE sample experiences lower wage growth over time, causing the wages to diverge around age 30 and the difference to increase over time. Compared to the comparable figure for teaching wages, graduate wages have a larger variance, start lower and progress faster during the early career years. However where teaching wages stabilise mid-career, graduate wages exhibit small levels of decline. Figure 2.7 displays the geographic variation of wages for the same two measures. Graduate wages exhibit much more geographic variation than teaching wages. Both ASHE and LFS show that wages are higher in the south east of England, and are particularly high within inner London. However the ASHE continuous weights measure of graduate wages displays generally lower wages and much less wage variation, with average wages across the remaining regions of the UK being broadly similar.

2.5 Conditional wage gap regressions

Before estimating the wage gaps faced by the teaching population, I estimate the wage gap using an OLS regression of weekly wages. For person i in year t in region r, I estimate wages as:

[6] The distribution of leavers weighting is not summarised here for simplicity, but is summarised in the main analysis.

$$\ln(\text{WeeklyWage}_{irt}) = \beta_0 + \beta_1\text{Teacher}_{irt} + \beta_2\text{Public Sector}_{irt} + \gamma X_{irt} + \varepsilon_{irt} \quad (2.1)$$

Where β_1 identifies the average difference in pay between teachers and other graduates in the public sector. In other words, this estimates the wage gap premium of teachers, controlling for the fact that teachers are part of the public sector. The full gap between teachers and private sector graduates is the sum of β_1 and β_2. I disaggregate wages in this way as public sector employees are often considered to have additional non-pecuniary benefits (better pensions, more job stability), and so the inclusion of this dummy acts as a proxy.

The controls include age, age squared, sex, and region of work and year fixed effects. For comparability between the LFS and ASHE, I estimate the regression once using controls common to both data sets with errors clustered at the region-year level. The common controls are age, age squared, a female dummy, and a public sector dummy. LFS includes only observations for those with a graduate degree and observations are weighted according to their sampling weights. Continuous graduate weights are applied to ASHE observations[7]. I also estimate the regression a second time including more detailed controls unique to each source. In this case, I include ethnicity, degree classification and a dummy for being born in the UK to the LFS specification. For ASHE, I include more granular spatial fixed effects and cluster at the individual level.

Table 2.4 summarises the results for equation 2.1. Even when controlling for public sector benefits, teachers experience a pay premium compared to similar graduates. Compared to private sector graduates, teachers are paid 2.4% more according to the LFS, or 15.6% more according to ASHE[8]. Compared to other public sector employees, the differences are 9.4% and 16.3% respectively. When including additional controls, the LFS teacher pay premium is largely unchanged, and remains statistically significant from private sector pay at the 5% level. When including more granular spatial fixed effects in ASHE, teacher pay premium is also minimally affected at 15.7% compared to private sector graduates, but the pay penalty for public sector workers (including teachers) becomes significant at 1.7%. Overall, the LFS and ASHE both predict a pay premium but vary dramatically in magnitude.

2.6 Estimation of teacher's wage gaps

Section 2.5 estimated the teaching wage gap by comparing teachers with other graduates, but this estimate depends on the relative sample of teachers in each data source. Teaching wage gaps will also vary by region, age, or year. One way to address this is to

[7]For simplicity I only summarise the results for continuous weights as in the full analysis they are found to be the median measure.

[8]Both these differences are statistically significant at the 1% level.

include additional terms that interact all controls with the teacher dummy. However, an alternative method is to predict both teaching and graduate wages for the universe of teachers in the SWC. Estimating the average difference between these two wages generates the equivalent of a fully interacted regression which is correctly weighted to the full population of state-funded teachers. This method, outlined below, also allows me to generate the individual wage predictions used in chapter 1.

In this section, I outline my method for estimating the teacher-graduate wage gap. I will estimate the gap separately using ASHE and LFS data. This method involves the following steps:

1. Generating observation weights for counterfactual graduates within ASHE
2. Estimating teaching wages in ASHE and LFS
3. Estimating counterfactual graduate wages in ASHE and LFS
4. Predicting both wages for all teachers in the SWC
5. Constructing and summarising the wage gap

The four different sets of weights outlined in section 2.4 fulfils the first step of this method. Applying each set of weights to the ASHE data generates four wage distributions for the counterfactual wages of teachers.

In step two I estimate teacher wages by fitting equation 2.2 on the sample of teachers in ASHE and LFS separately. For individual i, wages are a function of their age, region of work r, and year t. I also include additional fixed effects for the interaction of age and region, and of year and region in X_{irt}. Sex is omitted from this equation as it is not present in the SWC dataset[9]. Errors are clustered at the year-region level. To explore variation for STEM and Non-STEM specialist teachers, I can also estimate the equation separately for STEM and non-STEM graduates within the LFS[10].

$$\ln(\text{WeeklyWage}_{irt}) = \beta_0 + \beta_1 \text{age}_{irt} + \beta_2 \text{age}^2_{irt} + \gamma_r + \delta_t + \zeta X_{irt} + \varepsilon_{irt} \tag{2.2}$$

To estimate graduate wages in LFS, I estimate the same regression on the sample of graduates, excluding state-school teachers. Similar to teaching wages, I can estimate the regression separately for STEM and Non-STEM graduates. In ASHE, I include

[9]As both ASHE and LFS were good approximations of teaching wages in the SWC, they are likely to have a similar ratio of women in their sample of teachers and so estimates of teaching wages are unlikely to exhibit a large amount of bias. However, given that women typically face a wage penalty in the labour market and a majority share of teachers are female, this will bias the estimates of counterfactual graduate wages upwards. Therefore, the omission of sex from these regressions will likely bias teaching wage gaps downwards.

[10]For the ASHE sample, I do not estimate wages for STEM and non-STEM teachers separately.

additional terms as the sample size is much larger. I replace region fixed effects with NUTS3 fixed effects, and add an additional set of year-region fixed effects. I estimate equation 2.2 four separate times using the three separate weightings outlined in section 2.4.

Next, I predict both counterfactual graduate wages and teacher wages for each teacher in each year they are observed using the stored regression estimates. I predict teaching wages rather than using actual teaching wages to ensure that both types of wage data are collected using the same method and to avoid bias that may stem from comparing wages across two separate data sources. The wage gap for teacher i working in region r in year y is calculated to be the difference in the predicted log wages:

$$\widehat{\text{Wage Gap}}_{irt} = \ln(\widehat{\text{Teacher Wage}}_{irt}) - \ln(\widehat{\text{Counterfactual Graduate Wage}}_{irt}) \quad (2.3)$$

In summary, this generates four main different types of teacher wage gaps, depending on the counterfactual group:

1. LFS Graduates
2. LFS Graduates, STEM
3. ASHE Graduates, continuous weights
4. ASHE Graduates, binary weights
5. ASHE leavers

In the results section, I report the weighted mean and standard error of the estimated wage gaps for various subgroups. The means are weighted so that each individual is given equal importance, for example if an individual appears in the data N times, each of their observations is given a weighting of $1/N$. Since the mean is constructed from two fitted variables, I employ the delta method to account for this additional uncertainty, which is discussed in the appendix.

2.7 Results

2.8 Wage Gaps

Table 2.5 reports the average predicted wage gap for each estimation method. For the LFS, teachers earn a non-significant wage premium of 5.6% compared to other full time graduates. Using the continuous weights method in ASHE, this wage gap increases to 14.3% and becomes statistically significant. However, when considering only jobs where a large majority are graduates, the binary method, the wage gap becomes small, negative, and insignificant from zero. The largest estimated wage gap is generated by comparing teachers to workers in jobs that teachers commonly move into after they finish teaching. Teacher salaries are on average 33.5% higher than comparable individuals in these jobs. For a subsample of teachers, primarily those working in secondary schools, I am able to take their specialist subject as a proxy for whether they are a STEM or non-STEM graduate. When taking this into account, the average LFS wage gap remains insignificant at 1.8%. Comparing the average wage gaps with the predicted wage gaps from the conditional wage regressions in section 2.5, the magnitude of both the ASHE and LFS average wage gaps has decreased 2-4 percentage points (12% and 42% respectively).

Figure 2.8 reports the average wage gap across four age groups[11]. For most of the estimation methods, the gap remains relatively stable over time and for all methods, the gaps across age groups are not statistically different from one another. The only discernible trend is that there is a steady increase in the average ASHE leavers gap as age increases, and a U-shaped trend for LFS wages. A U-shape can stem from the underlying wage progression for the graduate and teaching samples in the LFS: For new graduates, teaching offers a higher starting salary. Graduate wages grow at a faster pace in the early part of the career, however in the last part of the career, teaching wages are more constant whereas graduate wages partially decline. This may be a result of differential retirement (or other workforce exit) patterns between teachers and other graduates. A steady increase in ASHE wage gaps across age can be explained by the fact that early career graduate wages exhibit a lower growth rate. Disaggregating gaps by subject (using the LFS STEM measure) and by teaching tenure showed no major discernible patterns.

The geographic variation in wage gaps is displayed in figure 2.9[12]. The highest wage premiums for teachers tend to occur in the northern parts of England. Teacher wages are least competitive in the South East and inner London, but are not substantially less competitive than other areas in the South. In outer London the teaching wage premium is relatively higher than inner London, but it is important to consider that wages are estimated based on the location of work. The real wage premium is likely less as individuals are likely to be geographically mobile across inner and outer London.

[11]Note that the first age group has a span of five years, the second and third group a span of 10, and the last group a span of 14 years. This is done in order to reflect the fact that wages rise quickly during the early part of a career and stabilise towards the end of a career. A line chart displaying the full shape of the trends can be found in figure appendix B

[12]The corresponding map for the ASHE leaver measure can be found in figure B.3 in appendix B. As the wage gaps are considerably higher, the heat map uses a different legend and colour scale to maintain sufficient contrast in colours for each separate map.

2.8.1 Teacher Retention

Next, I investigate whether these estimated wage gaps are related to the retention of teachers. As the SWC is a panel, I label teachers who are not present for at least one year of the census (after having first observed them as a teacher) as a leaver. In figure 2.10, I display the average wage gaps for those that will leave and never leave by age group. The averages are not statistically different from each other, and the means of those that will leave are actually slightly higher than those that remain[13].

To further explore the relationship between wages and retention, I predict the wage gap for the individuals in the SWC for the years where they are not teaching. I expand my panel of teachers to predict the wage gap for each teacher in years that they are not observed in the census[14] Figure 2.11 shows that the average wage premiums for non-teachers is slightly higher, but not significant.

Using predicted wage gaps, which are a function of personal characteristics, to analyse retention is sub-optimal as they exclude the random components of both the teaching and outside option wage. To explore exit intentions and still make use of random variation in wages, I regress the probability that an individual will leave teaching (the dummy constructed in figure 2.10) on their log weekly wages. I employ a linear probability model where errors are clustered at the individual level, and each observation is weighted inversely to the number of times that teacher appears in the SWC. Table 2.6 column 1 reports the coefficient of interest in a regression with no controls, and column 2 includes controls for age, teaching year, teaching years^2, year and region. A one standard-deviation increase in the log weekly wage significantly decreases the probability of leaving teaching by 3.2 percentage points. The magnitude increases to 9.9 percentage points when including controls.

2.9 Discussion

Each method outlined in this chapter produced a markedly different estimation of the average teacher wage gap. However, none of the methods predicted a statistically significant teacher wage penalty. All gaps were either insignificant from zero, or estimated that teachers are paid significantly more than comparable graduates employed in the same area. Whilst this may be unexpected given common narratives that teachers are underpaid, it is not a unique conclusion within the literature. It also does not suggest that teachers are overpaid, as the non-pecuniary benefits of alternative jobs could compensate for this wage imbalance.

It is important to consider that each measure might be most applicable to different

[13]This difference is unlikely to stem from teacher with higher tenure and higher salaries in each age group being more likely to leave via retirement. The difference in means is stable, and even increases slightly, when controlling for age or teaching tenure.

[14]I do not predict wages for any years before I first observe the individual as a teacher. I assume that the teacher's region is the same as the region they were most recently observed teaching in.

types of teachers, or different contexts. For example, the ASHE leaver measure was constructed based on the most likely jobs that ex-teachers occupy - and these were mostly teaching-adjacent jobs. This makes the measure a relevant counterfactual to estimate the outside options of individuals who have spent enough time in teaching to develop human capital specific to this occupation or industry, but whose human capital in other fields is limited. One such group is teachers who completed an undergraduate degree in education, but the leaver measure is also particularly relevant for senior teachers considering a different career path. For those who are considering becoming a teacher or have just completed postgraduate initial teacher training, the most relevant counterfactual is one of the graduate measures. Considering these wage gaps may inform this group of which career they should choose. These estimates are particularly relevant for the sample in chapter one, where I analyse the retention of trainee teachers. Given that the wage gap measure for graduate jobs is less than the leavers measure, this can reflect the fact that in order to attract more specialist teachers into secondary schools, they require a higher wage.

One surprising finding is the difference in the wage gap of the ASHE graduate measures that apply continuous and binary weights. Statistically, the binary measure contains a much smaller subsample of the continuous measure: the number of observations used in the binary regressions is 75,000, whereas the continuous measure includes 600,000 observations (with a weighted count of 252,000). It is therefore to be expected that including jobs with a lower share of graduates employed in the regression will decrease the fitted wage. However, the continuous measure is still a very relevant statistic given that 42% of graduates employed outside of London work in a job that does not require a degree(Xu, 2023). An oversupply of graduates in the labour market means that many may find a job in a graduate role, and this is particularly true for those who are not sufficiently mobile to relocate to areas with industries that have a higher demand for graduate labour. Alternatively, teaching roles are distributed more evenly across the country and there are relatively high levels of vacancies. For graduates located away from London and the South East, the continuous measure may be more appropriate.

This chapter was also an exercise in comparing wage estimates produced using LFS and ASHE data. As the LFS is able to directly identify graduates, it is not surprising that its wage gap estimate lies in-between the ASHE binary measure (that excludes graduates in typically lower paid jobs) and the continuous measure (that may disproportionately take into account non-graduate wages[15]). ASHE provided additional benefits over the LFS in that it more closely matched teaching wages

[15]The regressions used to predict graduate wages essentially take a weighted average of the wage all job types included in the sample, adjusting for controls. In the continuous measure, the regression will include jobs that have a low share of graduates (albeit with a low weight) to calculate the average. Let us call a hypothetical type of job with a low share of graduates job A. If graduates tend to have particularly high wages relative to non-graduates in job A, then the average wage underestimates the wage of graduates in this job. Therefore, the predicted wage of the continuous measure will be likely biased downwards.

recorded in the SWC, has less response bias, a larger sample size of workers (which allowed more granular regression analysis), and no censored earnings data. So, whilst the LFS generates a more conservative estimate of the wage gap, the ASHE wage gaps still bring valuable insight when analysing the dynamics of these gaps. Indeed, ASHE demonstrates different wage growth and spatial patterns.

It is important to consider sources of bias in these estimates. Firstly, the gaps are constructed using weekly wages which ignore differences in hours worked per week. As shown by the comparison of weekly wages between the three sources, teachers report working more hours in the LFS than those recorded by employers in ASHE and SWC. This biases the competitiveness of teacher pay in this study upwards. There are a number of reasons to also suggest that the wage gaps presented in this chapter are an underestimate. Firstly, the wage regressions used to produce the predicted graduate and teacher wages did not control for sex. Given that the wage penalty for females in the general labour market is less than the female penalty for teachers in the ASHE and LFS samples, the teacher pay premium is likely underestimated. Secondly, teaching wages in ASHE and LFS were slightly lower than those reported in the SWC. The wage gaps reported are likely to be more accurate for early career teachers, as wages across sources are more similar and wage variation is less.

This analysis primarily compared average earnings across groups. However as can be seen from figures 2.1 2.5, the distribution of weekly earnings are not similar across teaching and other graduate jobs. The wages of graduates are more negatively skewed, with a much longer right tail compared to teachers, where wages are relatively more equal. Comparison of wages at different quantiles of the distribution may show very different results.

There is scope to improve the accuracy of these estimates by taking advantage of improved detail in each of the datasets. Accessing additional information collected in the SWC on sex and ethnicity will allow improved comparability of wages, and new data-linking projects have enabled information on additional personal characteristics collected via the UK census to be matched to observations in ASHE. Lastly, the SWC contains a large amount of detailed data that hold great potential for both the economic literature related to teachers and for labour economics in general. It is a detailed panel that allows the tracking of individuals across jobs, details of their workload and additional responsibilities. It can also be linked to data on the performance and quality of schools. This makes it a rich dataset for measuring preferences in the labour market and measuring pay inequality.

In conclusion, this chapter has demonstrated that there is considerable heterogeneity in the teacher-graduate wage gap, depending on the method of estimation. For some of these measures, I find evidence of a teacher pay premium, and some measures estimate no significant difference in pay. I find mixed evidence of how pay gaps progress over time, but that regions in the south tend to have lower magnitude pay gaps. It is notable however, that I do not find a statistically significant negative pay gap using any of the pay methods, for any subgroup of teachers. Whilst this suggests that teacher pay is relatively competitive, it is important to remember that teaching may not remain competitive when taking account of the non-pecuniary elements of the job. Workload, student behaviour, and the general mental wellbeing of teachers needs to be considered. For example, it is important to understand whether the rise in flexible work arrangements, particularly homeworking, is disproportionately available to other graduates relative to teachers. Given that these initiatives may also disproportionately affect women, who are over-represented in teaching, only estimating pay gaps does not fully inform us of the competitiveness of teaching as an occupation.

2.10 Tables

Table 2.1: Total observations by data source

	LFS	ASHE	SWC
Total Observations	931,055	909,181	3,084,394
Employed Graduates	259,972	-	-
Teachers	31,572	29,284	2,390,265
State-funded School Teachers	25,071	18,721	2,390,265
Years Observed	2013-2019	2013-2019	2013-2019

Table summarises the observations used in each of the three key data sources for this chapter.

Table 2.2: Mean and standard deviation of weekly wages by data source

	LFS	ASHE	SWC
Full Population	£540	£505	-
(sd)	(372)	(351)	-
Employed Graduates	£757	-	-
(sd)	(400)	-	-
Teachers	£704	£744	-
(sd)	(256)	(255)	-
State-funded School Teachers	£712	£713	£733
(sd)	(246)	(225)	(227)

Mean wages and their standard deviations are reported for different populations across the three data sources. The wages in the final row differ from the second-to-last row as the sample has been limited to only teachers that work in the education sector, and only those who are employed by the public sector.

Table 2.3: Most common occupation-sector combinations for ex-teachers

Occupation	Industry	Share of Ex-Teachers
Caring personal service occupations	Education	0.14
Teaching and other educational professionals	Education	0.13
Teaching and other educational professionals	Administrative and support service activities	0.05
Teaching and other educational professionals	Human health and social work activities	0.04
Administrative occupations	Education	0.03
Corporate managers and directors	Education	0.03
Business and public service associate professionals	Education	0.02
Teaching and other educational professionals	Professional, scientific and technical activities	0.02
Business, media and public service professionals	Education	0.01
Caring personal service occupations	Human health and social work activities	0.01
Health and social care associate professionals	Education	0.01
Total Observations		1,829

Table reports the share of ex-teachers found in full-time employment in a given occupation occupation-sector (determined by the 2 digit SOC and Industrial section). The shares are weighted shares, where each individual is given equal weight, as individuals appear multiple times across the panel. Source: ASHE.

Table 2.4: Regression Results: Estimated Log wage gap by data source

	LFS		ASHE	
	(1)	(2)	(3)	(4)
Teacher dummy	0.094***	0.092***	0.163***	0.174***
	(0.006)	(0.006)	(0.004)	(0.004)
Public Sector	-0.070***	-0.073***	-0.007	-0.017***
	(0.008)	(0.007)	(0.010)	(0.003)
Age	0.104***	0.108***	0.085***	0.084***
	(0.002)	(0.002)	(0.002)	(0.001)
Age^2	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.188***	-0.191***	-0.179***	-0.176***
	(0.005)	(0.005)	(0.004)	(0.003)
Additional controls				
Year	X	X	X	X
Region of work	X	X	X	
Region interactions (age, year)	X	X	X	X
NUTS3 geographic unit				X
Ethnicity		X		
Degree classification		X		
Born in UK		X		
Total Observations	58,628	34,765	617,280	617,280

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors in parentheses are clustered at the region-year level in columns 1-3, and at the individual level in column 4. The table reports the coefficients in a linear regression of the log weekly wage of full-time graduates aged 21-59 in England. In ASHE, each obervation is weighted by its probability of belonging to a graduate, as explained in detail in section 2.4.

Table 2.5: Average predicted wage gaps by source

Data source	Method	Mean	Std. Error (delta)	Total Observations
LFS	Grads	0.056	0.108	2,577,277
LFS	Grads, stem	0.018	0.136	1,371,583
ASHE	Grads, Continuous weights	0.143**	0.068	2,577,277
ASHE	Grads, Binary weights	-0.023	0.078	2,577,277
ASHE	Leaver weights	0.335***	0.083	2,577,277

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors were calculated using the delta method. Table reports the average predicted wage gap,
ln(Weekly Teacher Wage) $-$ ln(Weekly Counterfactual Wage) for the sample of teachers in the SWC using different alternative wage counterfactuals. The teacher wage is estimated using full-time state school teachers.

Table 2.6: Regression results: Probability of ever leaving teaching - Actual teaching wage

	(1)	(2)
Log weekly wage	-0.1097***	-0.3403***
	(0.0021)	(0.0137)
Age		0.0015***
		(0.0001)
Teaching year		0.0196***
		(0.0003)
Teaching year^2		-0.0005***
		(0.0000)
Year & Region controls		X
	2,273,526	2,273,526

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors in parentheses are clustered at the individual level. The table reports the coefficients in a linear regression of the probability of ever leaving teaching during the data period (2013-2019), using the sample of teachers in the SWC.

2.11 Figures

Figure 2.1: Distribution of full-time teacher weekly wages by data source

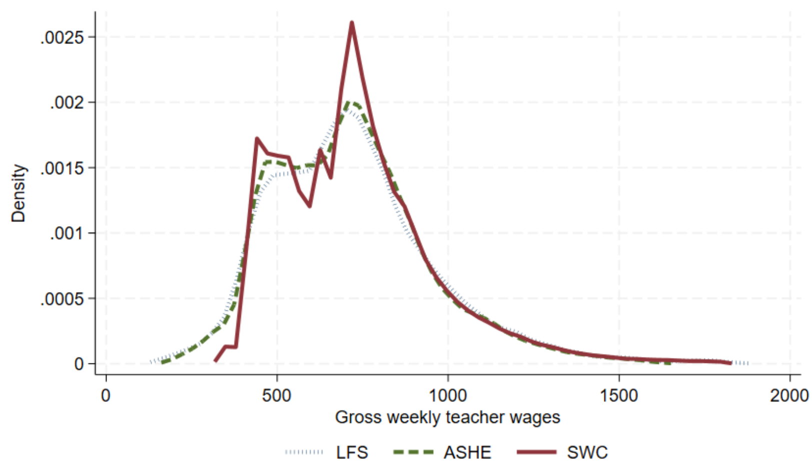


Figure outlines the kernel density of weekly full-time pay for state-funded school teachers in England for each of the data sources.

Figure 2.2: Distribution of full-time teacher hourly wages by data source

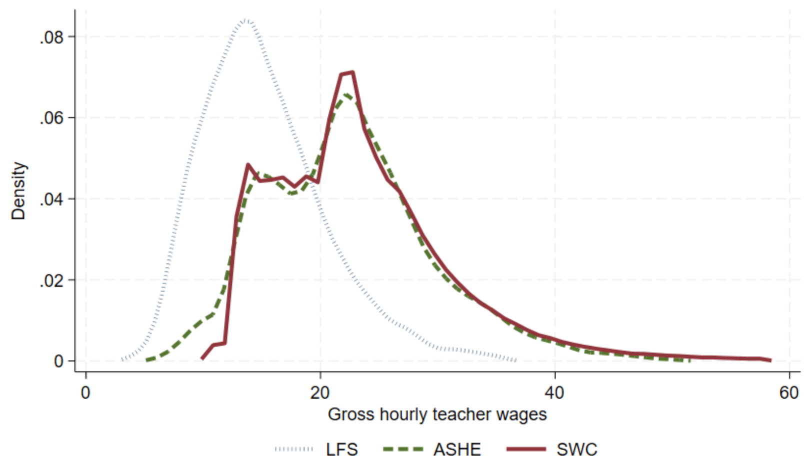


Figure outlines the kernel density of hourly full-time pay for public school teachers in England for each of the data sources.

Figure 2.3: Full-time teacher weekly wages by age and data source

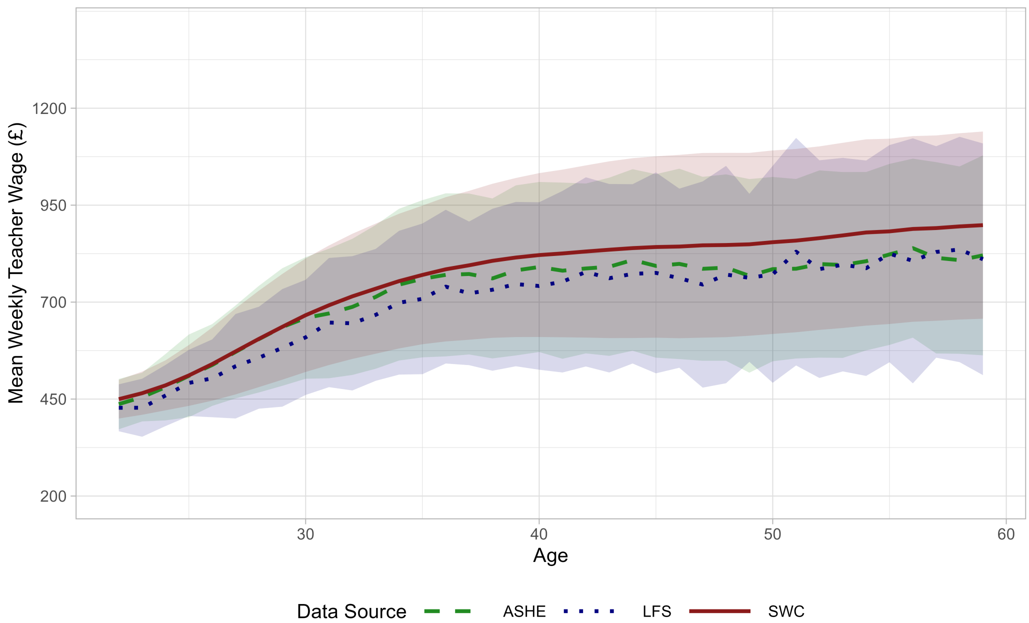


Figure outlines the average weekly wages for graduates across ages 22-59. The shaded area represents the standard deviation of the average. The axis are scaled to match the corresponding image for graduate wages.

Figure 2.4: Full-time teacher weekly wages by region and data source

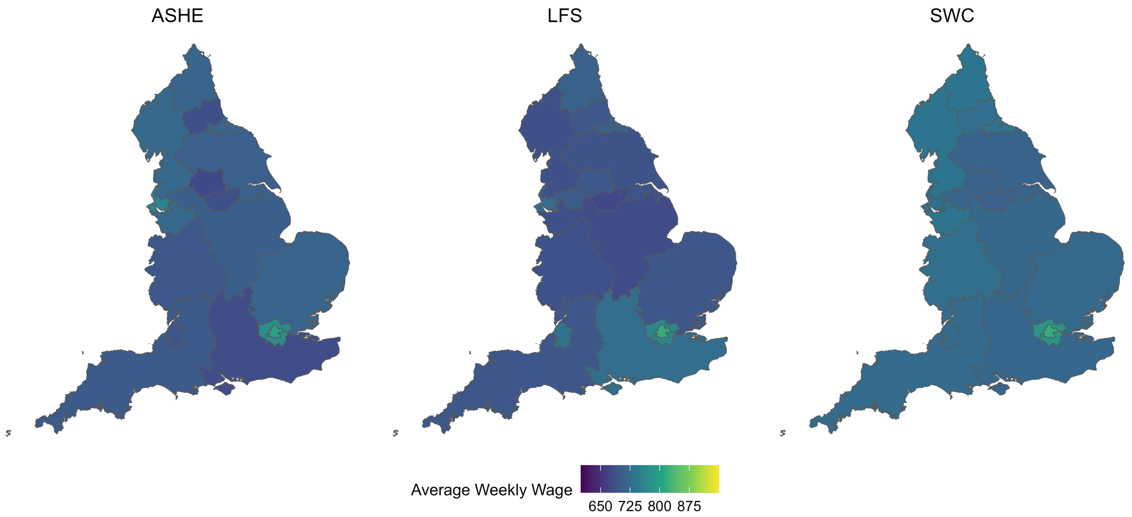


Figure outlines the average weekly wages for state-funded school teachers across the regions of England. Major subregions, such as the metropolitan areas of Bristol, Manchester, and Tyne and Wear are considered separately. The colour legend is scaled to match the corresponding image for graduate wages.

Figure 2.5: Distribution of full-time graduate weekly wages by data source

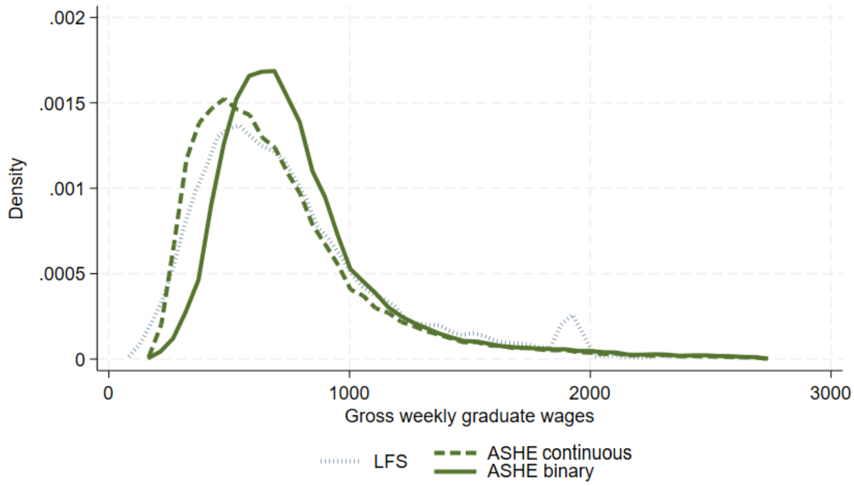


Figure outlines the kernel density of weekly full-time pay for graduates in England. The ASHE measures are constructed using two alternative sets of weights. The weighted measure weights each industry-occupation pair by the share of graduates employed in that job type in the LFS. The binary measure employs binary weights that are equal to one if an occupation-sector has over 80 percent graduates in the LFS. The mass in the right tail of the LFS sample is due to the censoring of annual wages in the LFS.

Figure 2.6: Average full-time weekly graduate wages by age and data source

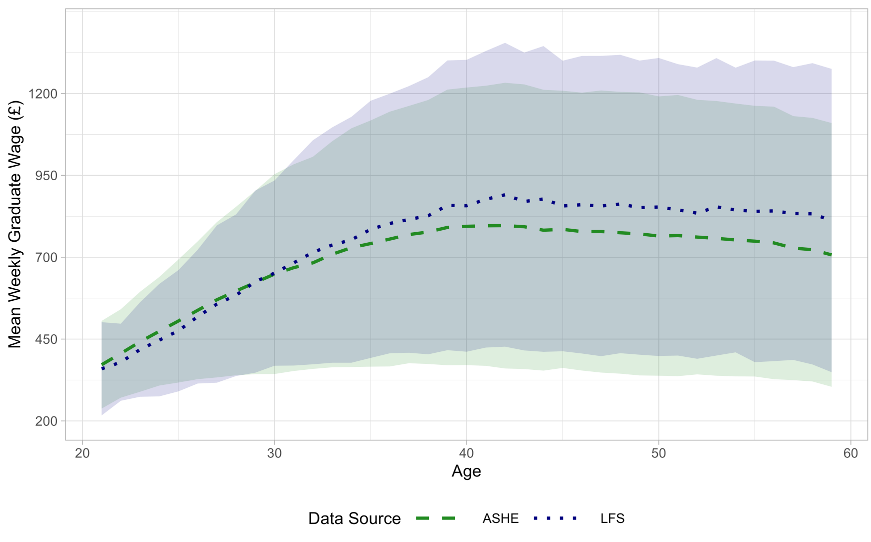


Figure outlines the average weekly wages for graduates across ages. The ASHE measure weights each industry-occupation pair by the share of graduates employed in that job type in the LFS. The shaded area represents the standard deviation of the mean.

Figure 2.7: Full-time graduate weekly wages by region and data source

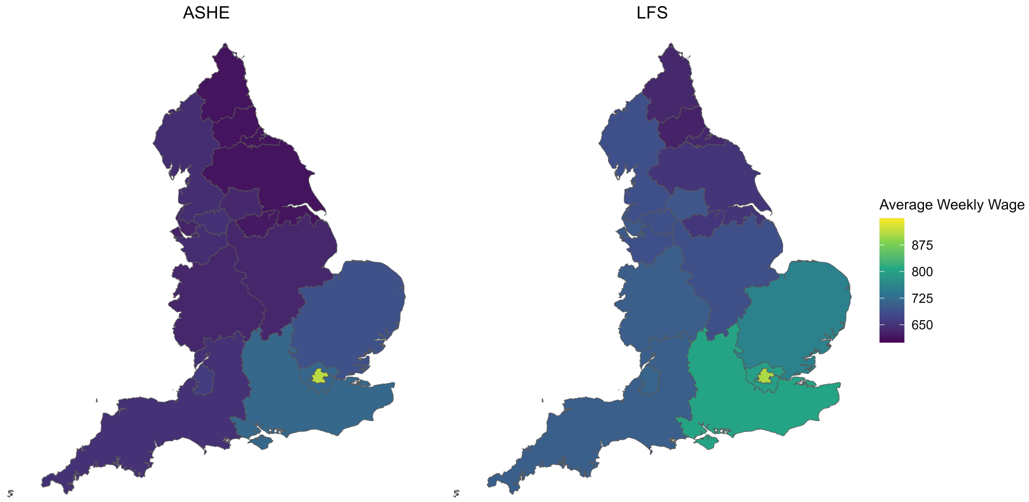


Figure outlines the average weekly wages for graduates across different regions. The ASHE measure weights each industry-occupation pair by the share of graduates employed in that job type in the LFS. Major subregions, such as the metropolitan areas of Bristol, Manchester, and Tyne and Wear are considered separately. The colour legend is scaled to match the corresponding graph for teacher pay.

Figure 2.8: Estimated average teacher wage gap by age group

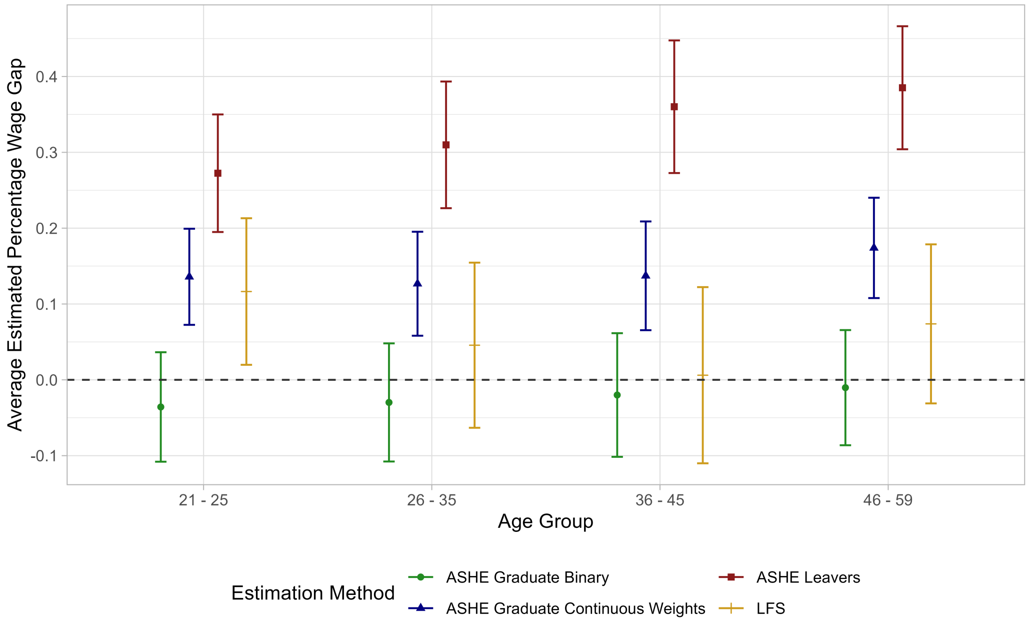


Figure estimates the average predicted wage gap, $ln(\text{Weekly Teacher Wage}) - ln(\text{Weekly Counterfactual Wage})$ for different age groups using the four key estimation methods. Note that the age bands are not uniformly distributed to reflect the relative difference in growth across the career. An equivalent line chart can be found in the appendix. Standard errors, estimated using the delta method, are also displayed.

Figure 2.9: Estimated average teacher wage gap by region and data source

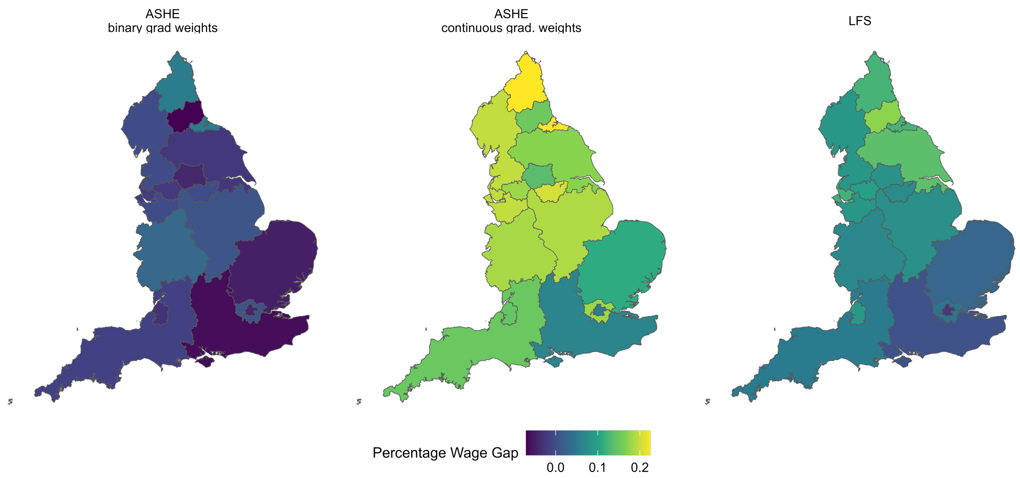


Figure estimates the average predicted wage gap, ln(Weekly Teacher Wage) − ln(Weekly Counterfactual Wage) for different regions in England. Major subregions, such as the metropolitan areas of Bristol, Manchester, and Tyne and Wear are considered separately. An equivalent map for the ASHE leaver measure is included in the appendix using a different colour scale.

Figure 2.10: Estimated average teacher wage gap by age and retention

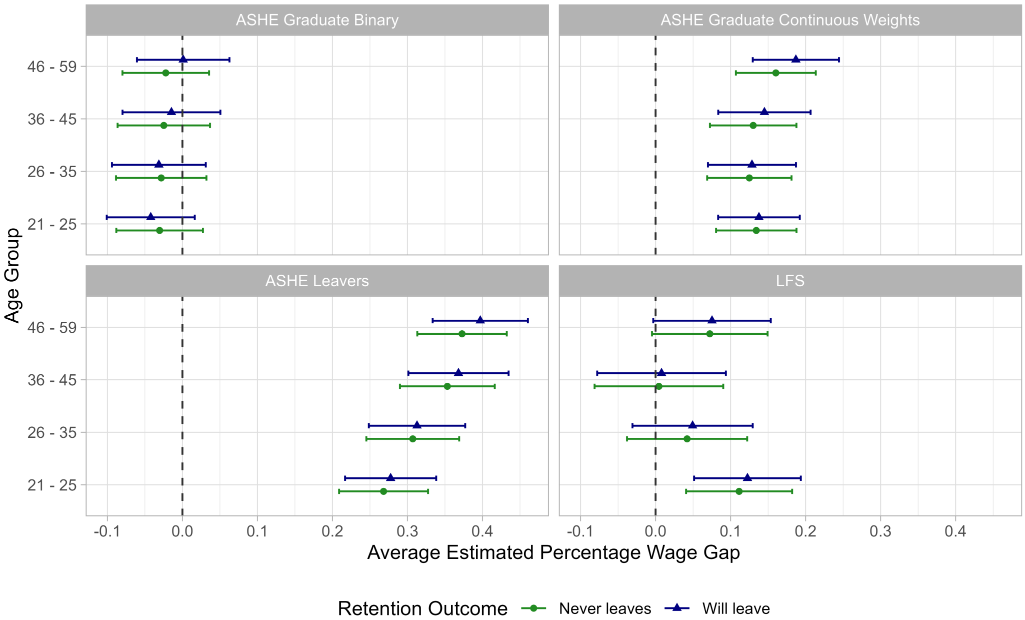


Figure estimates the average predicted wage gap, $ln(\text{Weekly Teacher Wage}) - ln(\text{Weekly Counterfactual Wage})$ for teachers in different age groups using the four key estimation methods. For each age, the chart compares those that are observed to leave teaching during the data interval, and those who remain as a teacher for all available census years after first observing them. Note that the age bands are not uniformly distributed to reflect the relative difference in growth across the career. Standard errors, estimated using the delta method, are also displayed.

Figure 2.11: Estimated average teacher wage gap by job type

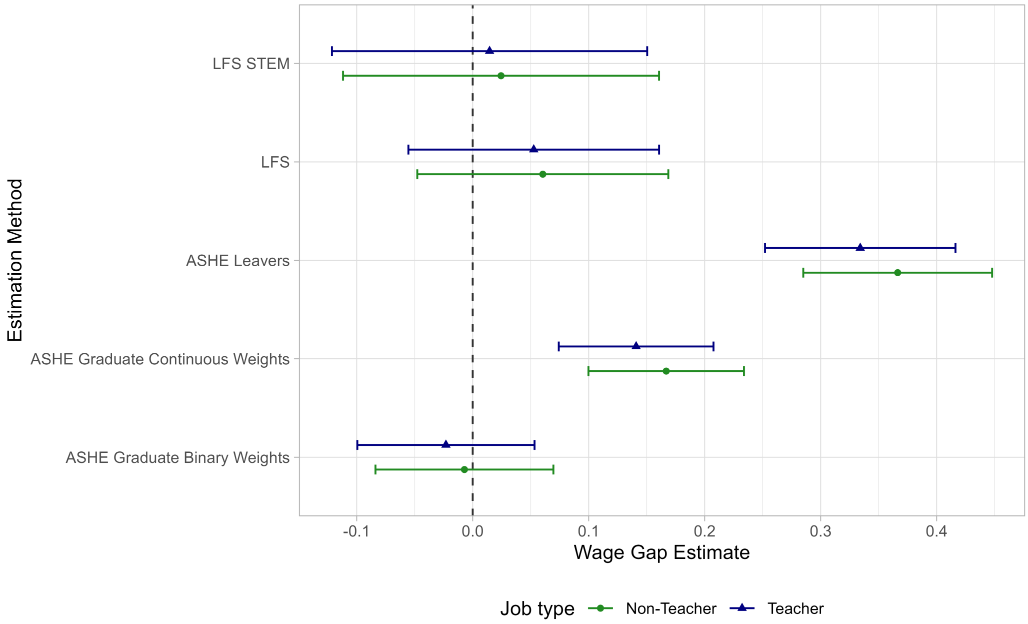


Figure estimates the average predicted wage gap, $ln(\text{Weekly Teacher Wage}) - ln(\text{Weekly Counterfactual Wage})$ for the sample of teachers in the SWC using four key estimation methods. The chart compares the predicted wages of the sample for years that they are teaching, versus years that they are not observed to be teaching in the census (after they have first been observed). The key difference in this data summary is that wage gaps have been predicted even for years where individuals are not teaching. Standard errors, estimated using the delta method, are also displayed.

Chapter 3

Ethnic Minority and Migrant Pay Gaps Over the Life-Cycle

Joint work with Tessa Hall and Alan Manning[1]

It is well-known that ethnic minority and migrant workers have lower average pay than the white UK-born workforce. However, we know much less about how these gaps vary over the life-cycle because of data limitations. We use new data that combines a 1999-2018 panel from the Annual Survey of Hours and Earnings (ASHE) with individual characteristics from the 2011 Census in England and Wales. We investigate pay gaps on labour market entry and differences in pay growth. We find that differences in entry pay gaps are more important than differences in pay growth. The entry pay gaps are large, though they vary across groups. The pay penalties on labour market entry can, to a considerable degree, be explained by over-representation in lower-paying firms and, within firms, in lower-paying occupations. For most groups, the pay gaps at entry seem to be largely preserved over the life cycle neither narrowing nor widening. For migrants, we find that the extra pay penalty is concentrated almost exclusively in those who arrived in the UK at later ages.

[1]This work contains statistical data from ONS which is Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. This analysis was carried out in the Secure Research Service, part of the Office for National Statistics. All errors are my own.

3.1 Introduction

It is well-known that ethnic minorities are paid less than similar white British workers (see Clark and Shankley (2020) for a recent review) though the magnitude of the pay gaps varies a lot by ethnicity. Differences in personal characteristics such as age, education, or region of work cannot fully explain these pay gaps (Brynin and Güveli, 2012). There are also large differences in unemployment rates (Clark and Shankley, 2020). Field experiments, where researchers send sets of fictitious job applications to employers which have the same level of education and skill but differ in the ethnicity of the applicant, find evidence of direct discrimination against ethnic minorities in hiring (Heath and Stasio (2019); Wood et al. (2009)).

There is also not much evidence that these pay gaps have declined, despite policy initiatives aimed at improving the situation of minority workers in the labour market. Manning and Rose (2021) find that pay gaps between black, Pakistani, and Bangladeshi groups and the white majority in the UK have widened in the past decade.

Existing research on ethnic pay penalties primarily uses cross-sectional data and estimates a single pay gap between ethnic groups[2]. There is little research on how differences in career progression contribute to observed pay gaps[3]. This is surprising considering that divergent wage progression plays a crucial role in explaining the gender pay gap (Blau and Kahn (2017); Manning and Swaffield (2008)). Men and women enter the labour market with similar wages, but women average slower pay growth in their twenties and thirties, mainly due to the responsibilities that come with raising children. As a result, there has been a growing interest in understanding the role of career progression, or the lack thereof, in the discussion of labour market inequalities.

The study of ethnic differences in pay progression by ethnicity has been hampered by poor availability of data. The Labour Force Survey (LFS), commonly used to analyse labour market disparities, contains at most two observations on individual earnings, spaced one year apart. The Annual Survey of Hours and Earnings (ASHE) alone follows individuals for their entire career but lacks much information on individual characteristics, notably ethnicity. Other longitudinal data sets have either too small a sample of ethnic minorities for analysis (e.g., the British Household Panel Survey) or only a short sample period (e.g., the UK Household Longitudinal Survey). As a result of these data deficiencies the best that could be done with available data is to estimate how earnings gaps vary over the life-cycle using repeated cross-sections.

[2]Though Clark and Nolan (2021) investigate how the gaps vary across the pay distribution. They find that black men face a glass-ceiling barring access to high paid jobs, which has worsened over time, driving an increase in their wage gap. At the bottom of the pay distribution, the introduction of national minimum wages in 1998 helped to reduce the pay gap between ethnic minorities and white workers.

[3]In surveys workers from minority backgrounds report that this hinders their opportunities for career progression (McGregor-Smith, 2017).

However, these estimates may be contaminated with cohort effects and selection into employment that varies with age.

This chapter utilizes new panel data to provide timely evidence on the distinct career dynamics experienced by ethnic minority and migrant groups. The Annual Survey of Hours and Earnings (ASHE) is the most reliable source of wage information in the UK and follows the same individuals over a long period. Our sample covers the years from 1999 to 2018. This has been linked to a rich set of individual characteristics from the 2011 Census in England and Wales, including ethnicity, education, country of birth, and year of entry to the UK. We use this data to understand how pay gaps by ethnicity evolve over the life cycle. Using the ASHE - 2011 Census panel data, we can fully account for individual characteristics and thus accurately isolate the effects of differential wage growth from cohort effects.

The chapter investigates pay gaps on labour market entry and differences in pay growth to explore intersectional gaps in career dynamics. We find that differences in entry pay gaps are more important than differences in pay growth. The entry pay gaps are large; after accounting for region of work and educational level, ethnic minority groups face an average wage penalty at entry compared to the white UK-born of 0.23 log points ($\tilde{2}3\%$) from men and 0.17 log points ($\tilde{1}7\%$) for women. This entry gap varies across groups, being widest for black African migrant men who face a penalty of 0.41 log points. For ethnic minority women, fixed wage gaps are smaller, and again largest for black African migrant women who face a 0.32 log point penalty. The pay penalties on labour market entry can, to a considerable degree, be explained by over-representation in lower-paying firms and, within firms, in lower-paying occupations (as in Phan et al. (2022)). Without considering the role of firms, most of these wage differences would have been attributed to individual characteristics like education and location.

For most groups, the sizeable pay gaps at entry seem to be largely preserved over the life cycle neither narrowing nor widening. UK-born ethnic minority men also generally experience slower wage growth throughout their career but, this contributes less to the overall wage gap than the pay gap at entry. The largest growth penalty is for black Caribbean migrant men: at age 45 around two-thirds of the pay penalty they experience is due to slower wage growth. On the other hand, UK-born women exhibit more pay convergence throughout their career. None of the UK-born women ethnic minority groups experience slower wage growth than whites. Black African, Indian, Pakistani and Bangladeshi women experience faster wage growth.

We find that migrants face an extra pay penalty on top of ethnicity but that this is concentrated almost exclusively in those who arrived in the UK at later ages. We argue this is because migrant status will often be invisible to employers while ethnicity rarely is.

The UK is actively engaged in policy discussions aimed at addressing the challenges faced by ethnic minority workers and the disadvantages they encounter. Our research provides policymakers with a more comprehensive understanding of the evolving inequalities within the labour market for different groups in the UK. By identifying the specific groups that experience significant disparities and pinpointing the specific stages in their careers when these disparities occur, we can develop more targeted and effective policies.

The remainder of the chapter is structured as follows: Section II describes the data and provides descriptive evidence of ethnic minority and migrant pay gaps across the career; Section III presents our empirical strategy to decompose fixed and dynamic pay gaps; Section IV outlines results; and in Section V we conclude and discuss policy implications.

3.2 Data and Descriptive Evidence

Our analysis primarily relies on the Annual Survey of Hours and Earnings (ASHE) matched to the 2011 Census in England and Wales. The ASHE dataset is derived from a 1% sample of employee jobs, extracted randomly from the Pay as You Earn (PAYE) register. This survey encompasses both the public and private sectors but excludes the self-employed who make up on average 12% of the UK workforce in our sample period. As shown in Table 3.1, Pakistani and Bangladeshi groups have considerably higher rates of self-employment than the white majority, especially among migrant men for which around 30% are self-employed. Much of the solo self-employment undertaken by this group in the UK is low-pay work with little opportunity for career progression; in 2018 around half of self-employed Pakistanis in the UK were taxi drivers. And the solo self-employed in the UK on average earn less than employees (Giupponi and Xu, 2020). Clark and Drinkwater (2000) suggest that the high rates of self-employment for Pakistani and Bangladeshi men in the UK are partly due to these workers being deterred from paid employment by discrimination. Boeri et al. (2020) highlight that for many minority groups solo self-employment is often a transitory state between unemployment and paid work. Although self-employment plays an important role in the career dynamics of ethnic minority and migrant workers, due to the nature of our data, the remainder of this chapter is focused on pay gaps over the career for paid employees.

Each year, ASHE provides information on approximately 140,000 to 180,000 employees. As workers are tracked throughout their entire careers based on their NINO (National Insurance Number), multiple years can be combined to create a panel dataset. The ASHE study has been extensively utilized for research on inequality and wage rigidities, thanks to its comprehensive earnings data and long panel (Bell, Bloom, and Blundell (2022); Elsby, Shin, and Solon (2016)). Additionally, the inclusion of firm identifiers in the dataset has allowed for investigations into within and between-firm inequality (Schaefer and Singleton, 2020).

The ASHE dataset provides valuable information on employees' hourly earnings, paid hours, occupation as well as a limited number of personal characteristics: gender, age, and location of work. To include a richer set of personal characteristics, the

ASHE dataset has been merged with the 2011 Census in England and Wales. The merged dataset includes additional personal characteristics such as educational and vocational qualifications, health status, migration status, country of birth, year of arrival to the UK, and ethnicity.

The ASHE dataset provides valuable information on employees' hourly earnings, paid hours, occupation as well as a limited number of personal characteristics: gender, age, and location of work. To include a richer set of personal characteristics, the ASHE dataset has been merged with the 2011 Census in England and Wales. The merged dataset includes additional personal characteristics such as educational and vocational qualifications, health status, migration status, country of birth, year of arrival to the UK, and ethnicity.

The dataset linking was done by identifying individuals in ASHE from either 2010, 2011, or 2012 in the 2011 Census by matching on a combination of name, sex, age, and residential postcode. Approximately 62% of eligible ASHE records (2010 – 2012) were matched in this way. The linked panel follows this subset of ASHE individuals with a successful Census match in the period from 1999 to 2018. Hence, only individuals who appeared in the 2011 Census can appear in this panel. We further restrict our sample to those aged 20 – 50[4]. Figure 3.1 shows the sample size of each year. As expected, this is highest in 2011 when the linkage was performed.

The quality of linkage between ASHE records (2010 – 2012) and the 2011 Census varies by employee, job, and employer characteristics. Linkage rates are lower for older and younger workers, greater for male employees, and lower for those living in high density areas like London (Forth et al., 2022). Since linkage rates are lower in London, ethnic minorities and migrants are under-represented in the ASHE – 2011 Census sample (see Appendix). This non-random sampling has the potential to bias our pay gap estimates if, for example, linkage quality by ethnicity also correlates with hourly pay. However, in the Appendix we show that the relationship between linkage quality and hourly pay is weak with the inclusion of region and age controls (included in all our main regression), which absorb the variation in linkage quality. Furthermore, our results are robust to applying sample weights designed by the ONS to make the ASHE-2011 Census sample representative of all jobs held by employees in England and Wales in 2011.

Also, in the Appendix we repeat our analysis using the Quarterly Labour Force Survey (QLFS) to check the representativeness of the ASHE – 2011 Census sample. Population characteristics and cross-sectional pay gaps in the QLFS are similar to those in the ASHE – 2011 Census sample, providing further evidence that the sample

[4]This is done because there are fewer than 40 observations for black African and Pakistani/Bangladeshi migrants past the age of 50, since these are relatively recent migrant cohorts. Similarly, there are insufficient observations for UK-born teenagers (age 16 – 19) in most ethnic minority groups. Since our main regression specification involves separately estimating ethnic and migrant pay gaps at each age group with individual fixed effects, we drop age ranges where there is insufficient sample size.

selection issue does not massively skew our results. The study of ethnic pay gaps in the UK has made significant progress by acknowledging the diverse nature of Britain's ethnic minority population, resulting in varying average earnings outcomes compared to the white population. Indian men and women tend to have lower wage gaps, while those with black or Pakistani/Bangladeshi heritage face the largest pay gaps. Ethnic minority women, on average, face smaller pay penalties compared to white women. However, it is important to note that white women already face significant penalties compared to white men. Migrant workers face an additional pay penalty when compared to UK-born of the same ethnicity. In light of this, we examine career wage gaps from an intersectional perspective, considering how gender, migrant status, and ethnicity interact in the labour market.

In our analysis, we study the following ethnic groups: Black Caribbean, Black African, Indian and Pakistani/Bangladeshi. We chose these groupings based on a combination of previous research documenting differences in pay gaps across ethnic groups and the need to have enough observations for reliable analysis. Some groups (Chinese, Arab, and those of other or mixed heritage) are excluded because of small sample sizes. Even the groups we use are themselves heterogeneous; Black African groups together Nigerians and Somalis, countries separated by 6 thousand kilometres. Figure 3.2 breaks down the 2011 maximum sample into each ethnic group by migrant status and gender. This figure makes it clear why we've decided to separate black African and black Caribbean groups. They display very different characteristics in terms of migrant composition: black African are a much more recent cohort whist black Caribbean are mostly UK-born (e.g., children of the Windrush generation).

A drawback of our data is that while the matched sample covers up to 19 years of an individual's work history (from 1999 to 2018), key demographic variables are only observed at one point in time (in 2011). However, the variables from the 2011 Census which we use in our analysis (ethnicity, country of birth, year of entry to the UK, qualification level) are likely to be fixed after entering the labour market.

3.3 Descriptive Evidence

Table 3.2 shows the characteristics of ethnic minority and migrant groups in the ASHE-2011 Census sample, pooling the years 1999 – 2018. Table 3.2 illustrates the dangers of simply looking at headline pay gaps. For UK-born men, Black Africans and Indians earn more, on average, than white men and for UK-born women average hourly earnings are lowest for the white group. But this is not comparing like with like. All ethnic minority and migrant groups are more likely to live in London, where average wages in the UK are highest, and hold at least a bachelor's degree compared to the white UK born majority. This is starkest for the black African UK born, of which 75% of men and 78% of women in our sample hold a degree[5]. Women all earn less than men of the same ethnicity, and the spread of their average wages by ethnicity is smaller. The most recent migrant cohort are black Africans while the oldest are Black Caribbeans.

These differences highlight the necessity to control for region of work and qualification level when measuring ethnic and migrant pay gaps. Whilst some minority groups may earn more than the white UK born on average - the highest earning men are white and Indian migrants and women are black African UK born - they experience a pay penalty when we control for educational level and region of work.

Table 3.3 presents cross-sectional estimates of the average difference in log hourly wage between each ethnicity / migrant status group and the white UK born majority of the same gender. Even when we look at the pay gap within occupations, firms, regions, education levels, and age, the results show that penalties persist. These estimates are in line with the literature (Clark and Shankley, 2020).

Raw gaps with only region, year, and age controls (column 1) are largest for black African migrants for both men and women. These gaps generally increase when we control for education, reflecting the fact that minority groups tend to be better educated than the white UK born majority. Further controlling for occupation and firm effects, gaps decrease. However, there are still significant pay gaps within the same firm and occupation for all groups (apart from white migrants).

The pay gaps reported in Table 3.3 are averages across the life cycle. Whilst these cross-sectional estimates give a good overview of average pay gaps in the UK, they do not tell us at which point in the career pay penalties arise. The main interest in this chapter is whether these gaps vary with age. As a first look at this, Figure 3.3 shows estimated earnings profiles by ethnicity, migrant status, and gender, controlling for region of work and year effects.

The main takeaway is that most of the pay gaps are on labour market entry and do not noticeably widen or narrow over the career. For UK-born men, all ethnic minorities enter the labour market at a statistically significant lower wage than whites. The entry gap is largest for black minorities and smallest for Indian minorities. On the whole, these gaps don't close over the career cycle, except for Indian UK-born men who earn the same as white UK-born in the same region between the ages of 40 – 50. For UK-born women, the pay gaps at entry are smaller. Women display more convergence in pay throughout their careers. All migrant men and women from ethnic minority groups face an additional penalty compared to UK-born of the same ethnicity. This is highest for black African male migrant minorities. The Appendix shows that the estimated effects are similar to those found in repeated cross-sections of the LFS, providing reassurance that the unusual nature of the ASHE-2011 Census matched data does not lead to very different conclusions.

The estimates in Figure 3.3 might also reflect changing cohort characteristics or

[5]This is a bit higher than appears in other sources e.g. the Census and the LFS. But those other sources also find that the UK-born with Black African ethnicity are very highly educated.

changing selection into work over the life cycle. Our main estimates exploit our panel data to account for these possibilities; we now turn to this.

3.4 Estimation Methodology

First, we use our panel data to disaggregate observed wage gaps at each age into the contribution of a fixed pay gap at labour market entry and pay growth gaps. To do this, we regress log basic hourly pay from ASHE on a set of dummies $\beta^{g,A}$ that represent each 5-year age category A for a distinct group g, where g is defined by the combination of ethnicity, migrant status, and sex. This set of dummies allows pay to vary flexibly across age groups for each unique ethnicity-migrant-sex group. As we are using panel data, we can follow workers over time and can control for individual fixed effects α_i; this controls for changing selection into work over the life cycle. For individual i belonging to group g in year t and age category A, log hourly wages w_{it} are given by:

$$w_{it} = \beta^{(g,A)} + \alpha_i + \epsilon_{it} \tag{3.1}$$

All the regression specifications also control for region of work, time trends and regional time trends but omit them from the notation in the interests of simplicity. We chose to use dummy coefficients $\beta^{g,A}$ for each age category instead of specifying a functional form in age (e.g., quadratic or quintic). This more flexible specification bypasses the debate over which functional form is best when analysing how wages evolve over the career (Murphy and Welch, 1990). Because of the inclusion of individual fixed effects, we can set $\beta^{g,entry} = 0$ without any loss of generality.

Pay on entry is captured by the individual fixed effects α_i. To investigate how these vary by ethnicity, we take the estimated fixed effects and regress them on ethnicity. We further look at how much of the fixed effect is explained by education by adding controls for highest qualification level (at a detailed 16-level description from the 2011 Census).

3.5 Results

3.5.1 Entry Pay and Pay Growth Gaps

The first two columns of Table 3.4 show estimates of the entry pay gaps for men (relative to white UK-born), while the fifth and sixth columns show the entry gap for women. Except for the white migrant groups, all the estimated entry pay gaps are negative, implying that all non-white groups earn less than the white UK-born population. For all non-white groups, the entry pay gaps for men are over 10 log points; for women they are generally smaller. Some of the entry pay gaps are very large (41 log points for Black African migrants). Controlling for an individual's highest qualification level widens the pay gap at labour market entry for most groups. This is especially true for black African UK-born minorities for which the pay gap increased by 0.18 log points for women and 0.14 log points for men with the addition of educational controls. This is in line with Clark and Nolan (2021), who also observe that all ethnic groups experience lower earnings compared to the white majority than they would if their qualifications were equally valued and rewarded. For migrants, this might be because foreign qualifications are not equally regarded in the UK. For ethnic minorities, entry pay gaps among those with the same level of education might reflect differences in human capital developed in the labour market (Tomaskovic-Devey, Thomas, and Johnson, 2005) or unmeasured differences in educational quality, like university and subject prestige (Gaddis, 2015). Another possibility, which is supported by field experiments which find evidence of direct labour market discrimination (Heath and Stasio (2019); Wood et al. (2009)), is that ethnic minorities with the same human capital are treated unequally by employers.

Table 3.4 also shows the pay growth gaps in the early career (between entry and 35), and mid-career (between 35 and 45). The estimates for men are in columns 3 and 4 while for women they are in columns 7 and 8. These pay growth gaps should be added to the entry growth gaps to give the total pay gap at age 35 or 45. In contrast to the entry pay gaps, the pay growth gaps are not always significantly different from zero and are not always negative. For example, Indian UK-born men are estimated to have pay growth 0.056 log points ($\tilde{5}.6\%$) higher than white UK-born men. The main conclusion is that most of the pay gaps are on labour market entry and persist through the life cycle.

3.5.2 The Role of Firms and Occupation

Segregation by ethnicity into lower-paying occupations or firms could also account for some of the ethnic and migrant pay gap. Evidence for firm-specific wage effects is found by Phan et al. (2022) who, using the same ASHE – 2011 Census linked sample as we do, find that the concentration of ethnic minorities into lower-paying firms account for sizeable parts of estimated wage gaps. Zwysen and Demireva (2020) find that UK-born ethnic minorities are less likely to work in the highest-paying occupations, but the type of disadvantage differs strongly between groups. To explore labour market entry and growth gaps within and between firms and occupations, we add firm (establishment identifier) fixed effects $\eta_{k(i,t)}$ and occupation (4-digit SIC 2010 codes) fixed effects $\phi_{j(i,t)}$ to our regression.

$$w_{it} = \beta^{(g,A)} + \alpha_i + \phi_{j(i,t)} + \eta_{k(i,t)} + \epsilon_{it} \tag{3.2}$$

This specification is used to decompose the fixed entry wage gap into the contribution of:

(i) fixed entry gap within firm and occupation (differences in the individual fixed effects),

(ii) over-representation of minorities in low-paying occupations at labour market entry,

(iii) over-representation of minorities in low-paying firms at labour market entry.

Likewise, the dynamic growth gap can be decomposed into the contribution of:

(iv) minority groups facing differential growth within firm and occupation,

(v) differential occupational upgrading for minority groups,

(vi) differential firm switching behaviour for minority groups.

For simplicity we do not control for education in reporting the average gap in fixed effects. Table 3.5 reports the results of equation 3.2 and decomposes the fixed gap and growth penalty gap (from entry to age 45) for men into the within firm-occupation effects and the between firm-occupation effects.

One systematic finding is that, at labour market entry, all non-white groups are over-represented in low-wage firms and, within those firms, low-wage occupations. The magnitude of this effect is quite similar across groups. One consequence of this is that controlling for firm and occupation fixed effects reduces the gaps in the average fixed effects and most are no longer significantly different from zero. This means that a large part of the entry gaps come from minority groups being concentrated in firms and occupations which pay lower wages at labour market entry. A notable exception are black African migrants, who still experience large entry gaps within firm and occupation.

For the growth gaps, most, but not all, of the groups seem to have modest firm and occupation upgrading over their careers but not enough to surmount the initial entry gaps faced. Pay gaps remain very substantial at age 45.

Table 3.6 shows the same results for women. For women, the patterns are similar to those for men; over-representation of non-white groups into low-paying firms and occupations on labour market entry, differences that are partially undone over the course of the career. The overall conclusion is that the pay penalties experienced by non-white groups are present on labour market entry and largely persist through the life cycle without either narrowing or widening.

Migrant Wage Penalties

Finally, we explore further the pay penalties experienced by migrants. In particular, we investigate different pay gaps by age of arrival to the UK. Migrant workers who arrived at older ages may face extra challenges such as limited English proficiency, qualifications that may not be universally recognized by employers, and unfamiliarity with the cultural norms of the UK. These factors may directly impact their earnings. Our empirical specification differs from the way the migrant 'assimilation' effect was initially explored by Chiswick (1978) who revealed a positive correlation between the length of time migrants spent in a host country and their wages[6].

Length of time in the UK can be inferred from the difference between age and age at arrival. Conditioning on age of arrival has the advantage that it allows a natural comparison with the UK-born as we might expect migrants who arrived at very young ages to be treated similarly to those born in the UK.

Table 3.7 investigates the impact of age at arrival. The reported migrant wage penalties are in addition to the ethnic wage gaps reported in the first row. Migrants face an average additional penalty of 0.02 for women and 0.08 log points for men compared to UK-born of the same ethnicity (given region, year, age and education controls). Decomposing the migrant penalty by age of arrival group, we find that it is mostly migrants who arrive at older ages that face the largest penalties. For migrants who arrived as children (before age 10), no pay penalty is experienced, and in some cases a pay advantage. This suggests assimilation effects: migrants who arrived older have less UK labour market-specific knowledge or qualifications and hence experience pay disadvantages, those who arrived young don't face these barriers. This is not that surprising; applications for jobs typically do not ask for country of birth so a migrant who has been in the UK almost their whole life will seem indistinguishable from someone born in the UK. Their ethnicity will, however, remain visible.

3.6 Conclusion and Policy Implications

The UK has large pay gaps between ethnic minorities and the white population. Non-white migrants typically face an additional pay penalty. This chapter explores these pay penalties over the life cycle by gender, ethnicity and migrant status for the UK using a new data set that links the longitudinal Annual Survey of Hours and Earnings with the 2011 Census. This combines high-quality longitudinal earnings information with individual characteristics that are often missing from employer-employee data sets.

The chapter investigates the disparity in wages when individuals enter the labour market, as well as the differences in wage growth. We find that differences in entry

[6]Borjas (1985) pointed out the limitation of the pioneering cross-sectional regression analysis in distinguishing between the influence of time spent in the host country and the different characteristics among different migrant cohorts. Put simply, the presence of strong assimilation in a cross-section could potentially be attributed to the fact that previous migrant cohorts were more skilled. Later studies (Borjas, 1995, 2015) use longitudinal data to address this issue. Selective outmigration of less successful migrants can also bias cross-sectional estimates of migrant assimilation (see Dustmann and Görlach (2015) for a review). Lubotsky (2007) uses longitudinal earnings data to show that selective emigration leads to an overestimation of wage growth for migrants who stay. The size and direction of this bias has been recently debated (Akee and Jones, 2019; Rho and Sanders, 2021).

pay gaps are more important than differences in pay growth. The entry pay gaps are large, though vary across groups. For most groups, the pay gaps at entry seem to be preserved over the life cycle neither narrowing nor widening. For migrants, we find that the extra pay penalty is primarily concentrated among those who arrived in the UK at a later age. We have argued that this is because for migrants who arrived as children migrant status will often be invisible to employers while ethnicity rarely is.

The pay penalties on labour market entry can, to a considerable degree, be explained by over-representation in lower-paying firms and, within firms, in lower-paying occupations. A significant body of literature explores how wage-setting behaviour by firms contributes to inequality in the labour market (Card, Heining, and Kline (2013); Song et al. (2019)). Differences in wage levels among employers may arise due to labour market frictions, monopolistic behaviour on the part of employers, and variations in rent sharing. Ethnic differences can arise, for example, if ethnic minority and migrant groups have lower ability to extract rent (Card et al., 2018), or if minority groups have lower reservation wages, strengthening the monopsony power of firms to suppress wages (Amior and Stuhler, 2024).

Occupational segregation of ethnic minorities and migrants into lower paying jobs, within the same firm and education level, can arise as a result of discriminatory practices in hiring (Heath and Stasio, 2019) or historical and cultural ties between ethnic groups and certain occupations (Engstrom, 1997). The influence of social networks can also contribute to occupational segregation (Calvó-Armengol and Jackson (2004); Ioannides and Datcher Loury (2004)). If workers are more likely to refer people from their own ethnic group, networks of individuals of the same ethnicity can increase the chances of finding a job through informal referrals. However, if these networks consist mostly of people in low-paying occupations, they can exacerbate occupational segregation.

The pay gaps we have estimated may also be influenced by geographical factors we're unable to capture in our data. Many ethnic minorities live in London, and we have accounted for this by including regional controls. However, labour markets may be more local than broad region (Manning and Petrongolo, 2017) and ethnic minorities are often overrepresented in more deprived urban areas (Clark and Drinkwater, 2002) where job opportunities and average wages are lower.

There are limitations to our research caused by the limitation of the data to employees, omitting those who are self-employed or unemployed. Black Caribbean and African men in particular experience large unemployment gaps, even after controlling for age, region, education, and marriage status (Clark and Shankley, 2020). Pakistani and Bangladeshi men, especially migrants, are more likely to be self-employed (see Table 3.1). Self-employment often serves as a middle ground between unemployment and traditional employment. In 2019, approximately 25% of newly self-employed individuals were previously unemployed, while an additional 31% were inactive (Giupponi and Xu, 2020). Career dynamics in self-employment and unemployment may be important for a comprehensive understanding of career dynamics but are beyond the scope of this chapter.

Moreover, ethnic minorities are disproportionately represented in precarious work arrangements involving 'gig' economy jobs and zero-hours contracts which have rapidly increased over the past decades. Bowyer and Henderson (2020) analyse data from the Next Steps, a longitudinal study of the millennial generation in England, and find that millennials from black and Asian minority ethnic backgrounds are 47% more likely to be on a zero-hours contract. This suggests that hourly earnings may mask additional pay gaps on the intensive margin, particularly on labour market entry. The scope for promotion and pay progression in these precarious jobs is also limited, therefore pay growth gaps could also widen in the future as these cohorts age.

Our findings, which suggest that there remains widespread disadvantage among non-white groups in the UK labour market, are consistent with evidence from audit studies of employer discrimination (Wood et al. (2009); Heath and Stasio (2019)). This contrasts with the more optimistic conclusion of the Sewell Report on Race and Ethnic Disparities (Commission on Race and Ethnic Disparities, 2021) which implies that gaps are small and falling. There is an urgent need to develop policies to address this injustice. The finding that most ethnic and migrant pay gaps are incurred at labour market entry suggests that this is a key time for policy intervention. Possible initiatives include providing better career information in schools, improving student-university matching, and ensuring equal access to vocational training for ethnic minorities. It is also essential to address the issues on the employers' side. This involves analysing recruitment procedures and closely examining hiring discrimination, especially in specific occupations where ethnic minorities are under-represented.

3.7 Tables

Table 3.1: Self-employment rates in England and Wales by migrant status and ethnicity, aged 20 – 50.

	UK-born		Migrant	
Men	Self-employed (%)	Of which solo (%)	Self-employed (%)	Of which solo (%)
White	15.30	77.89	17.06	83.23
Black Caribbean	15.05	91.65	16.35	85.32
Black African	13.20	83.51	11.83	84.76
Indian	14.89	63.95	15.20	67.96
Pakistani/Bangladeshi	21.70	74.54	30.89	78.39
Women	Self-employed (%)	Of which solo (%)	Self-employed (%)	Of which solo (%)
White	7.26	81.85	10.99	88.77
Black Caribbean	4.56	89.20	4.82	91.18
Black African	5.49	89.89	4.95	85.65
Indian	6.35	73.59	8.34	69.34
Pakistani/Bangladeshi	6.55	74.29	10.57	75.75

Source: LFS – 1999-2018.

Table 3.2: Sample characteristics by sex, ethnicity, and migrant status.

UK Born	**Men**					**Women**				
	Hourly wage	Age	Share with Degree	Share In London	N	Hourly wage	Age	Share with Degree	Share In London	N
White	14.34	39.16	36%	11%	513,668	11.33	40.27	36%	9%	518,318
	(10.79)	(9.99)				(7.63)	(9.96)			
Black	12.57	39.72	33%	49%	3,888	12.46	39.58	38%	59%	5,674
Caribbean	(6.64)	(8.91)				(7.18)	(9.19)			
Black	15.19	37.28	75%	76%	793	13.60	37.42	78%	69%	1,209
African	(8.57)	(9.20)				(7.54)	(9.22)			
Indian	14.58	32.96	57%	33%	5,793	13.13	33.42	63%	33%	5,861
	(11.83)	(7.77)				(8.94)	(8.05)			
Pakistani/	12.24	30.92	48%	27%	3,758	11.91	31.26	51%	30%	3,473
Bangladeshi	(8.43)	(6.99)				(7.14)	(7.38)			
Total	14.32	39.03	36%	12%	527,900	11.37	40.12	37%	10%	534,535
	(10.76)	(9.98)				(7.65)	(9.97)			

Migrant	**Men**						**Women**					
	Hourly wage	Age	Degree (%)	In London (%)	Time in UK	N	Hourly wage	Age	Degree (%)	In London (%)	Time in UK	N
White	18.16	40.46	51%	23%	29.66	12,672	13.34	41.39	49%	20%	29.78	14,253
	(16.06)	(9.38)			(14.23)		(8.76)	(9.75)			(14.32)	
Black	12.26	44.06	27%	53%	25.73	1,349	10.91	45.27	34%	57%	27.66	2,719
Caribbean	(7.15)	(9.76)			(15.66)		(5.43)	(9.34)			(14.68)	
Black	12.32	40.51	55%	52%	14.08	4,599	11.62	40.67	57%	56%	14.64	5,040
African	(8.02)	(9.12)			(8.18)		(7.26)	(9.48)			(8.47)	
Indian	15.93	41.74	53%	37%	22.20	8,779	11.93	42.64	45%	42%	23.76	8,957
	(13.49)	(9.05)			(14.60)		(8.82)	(8.89)			(13.63)	
Pakistani/	13.46	37.47	50%	40%	20.35	4,482	10.98	37.15	42%	36%	24.54	1,907
Bangladeshi	(12.40)	(9.19)			(12.48)		(6.65)	(9.18)			(11.39)	
Total	15.79	40.55	51%	35%	23.88	31,881	12.36	41.69	48%	36%	25.34	32,876
	(13.82)	(9.36)			(14.52)		(8.27)	(9.56)			(14.25)	

Means reported and standard deviations in parentheses. Source: ONS ASHE – 2011 Census.

Table 3.3: Cross-sectional regression table of log hourly wage gap with white UK born, various controls.

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Men				**Women**			
White	0.0910***	0.0351***	0.0225***	0.00319	0.0955***	0.0367***	0.0143***	0.00908**
Migrant	(0.00485)	(0.00427)	(0.00476)	(0.00384)	(0.00396)	(0.00345)	(0.00388)	(0.00309)
Black Caribbean	-0.329***	-0.229***	-0.145***	-0.0875***	-0.160***	-0.153***	-0.122***	-0.0731***
Migrant	(0.0113)	(0.0112)	(0.0116)	(0.00874)	(0.00768)	(0.00686)	(0.00702)	(0.00556)
Black Caribbean	-0.259***	-0.217***	-0.185***	-0.113***	-0.0506***	-0.0497***	-0.0544***	-0.0237***
UK born	(0.00680)	(0.00638)	(0.00655)	(0.00532)	(0.00570)	(0.00527)	(0.00569)	(0.00434)
Black African	-0.378***	-0.434***	-0.302***	-0.156***	-0.176***	-0.251***	-0.193***	-0.101***
Migrant	(0.00724)	(0.00671)	(0.00652)	(0.00478)	(0.00627)	(0.00554)	(0.00576)	(0.00450)
Black African	-0.169***	-0.301***	-0.193***	-0.112***	-0.0314**	-0.201***	-0.183***	-0.0911***
UK born	(0.0167)	(0.0160)	(0.0154)	(0.0126)	(0.0117)	(0.0119)	(0.0125)	(0.00944)
Indian	-0.113***	-0.155***	-0.0989***	-0.0831***	-0.117***	-0.134***	-0.105***	-0.0662***
Migrant	(0.00640)	(0.00552)	(0.00590)	(0.00424)	(0.00537)	(0.00467)	(0.00492)	(0.00359)
Indian	-0.0478***	-0.117***	-0.131***	-0.0794***	0.0541***	-0.0475***	-0.0818***	-0.0388***
UK born	(0.00634)	(0.00576)	(0.00585)	(0.00458)	(0.00599)	(0.00547)	(0.00622)	(0.00473)
Pakistani/Bangladeshi	-0.254***	-0.257***	-0.165***	-0.110***	-0.150***	-0.144***	-0.126***	-0.0565***
Migrant	(0.00838)	(0.00724)	(0.00753)	(0.00563)	(0.0106)	(0.00914)	(0.0104)	(0.00722)
Pakistani/Bangladeshi	-0.137***	-0.164***	-0.154***	-0.102***	-0.0346***	-0.0848***	-0.0866***	-0.0457***
UK born	(0.00731)	(0.00663)	(0.00659)	(0.00530)	(0.00756)	(0.00682)	(0.00735)	(0.00558)
Cons	2.510***	2.513***	2.556***	2.552***	2.306***	2.310***	2.354***	2.351***
	(0.000661)	(0.000577)	(0.000491)	(0.000395)	(0.000625)	(0.000533)	(0.000475)	(0.000374)
N	555925	555848	483829	483828	557545	557510	490130	490124
R-squared	0.253	0.433	0.688	0.796	0.211	0.431	0.634	0.772
Controls								
Region - year	X	X	X	X	X	X	X	X
Age	X	X	X	X	X	X	X	X
Educ - year		X	X	X		X	X	X
Firm			X	X			X	X
Occupation				X				X

Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: ONS ASHE – 2011 Census.

Table 3.4: Decomposition entry and growth gaps in log hourly wages

	Men				Women			
	Entry gap		Growth gap		Entry gap		Growth gap	
	(1) *Entry Gap*	(2) *Control for Education*	(3) *Growth gap: Entryto35*	(4) *Growth gap: 35to45*	(5) *Entry Gap*	(6) *Control for Education*	(7) *Growth gap: Entryto35*	(8) *Growth gap: 35to45*
White Migrant	0.0292 (0.0207)	-0.0254 (0.0207)	0.0527*** (0.0147)	0.0445*** (0.00839)	0.0382 (0.0203)	-0.0168 (0.0203)	0.0579*** (0.0151)	0.0191* (0.00934)
Black Caribbean Migrant	-0.103 (0.0547)	-0.0126 (0.0547)	-0.137** (0.0419)	-0.0498 (0.0377)	-0.165*** (0.0434)	-0.138** (0.0434)	-0.0416 (0.0362)	0.0937*** (0.0234)
Black Caribbean UK born	-0.136*** (0.0321)	-0.0931** (0.0321)	-0.0775*** (0.0201)	-0.0471*** (0.0136)	-0.0870** (0.0306)	-0.0862** (0.0306)	0.0325 (0.0205)	0.0237 (0.0123)
Black African Migrant	-0.413*** (0.0366)	-0.471*** (0.0366)	-0.00557 (0.0255)	0.0548*** (0.0164)	-0.322*** (0.0356)	-0.400*** (0.0356)	0.0568* (0.0278)	0.150*** (0.0175)
Black African UK born	-0.293*** (0.0690)	-0.428*** (0.0690)	0.136** (0.0511)	-0.0349 (0.0335)	-0.197 (0.107)	-0.372*** (0.107)	0.166*** (0.0419)	0.00880 (0.0296)
Indian Migrant	-0.141*** (0.0284)	-0.190*** (0.0284)	0.0391 (0.0234)	0.0555*** (0.0107)	-0.132*** (0.0309)	-0.150*** (0.0309)	0.0302 (0.0278)	0.0152 (0.0111)
Indian UK born	-0.191* (0.0816)	-0.270*** (0.0816)	0.0560*** (0.0142)	-0.00929 (0.0141)	-0.0626 (0.0510)	-0.185*** (0.0510)	0.109*** (0.0151)	0.00233 (0.0134)
Pakistani/Bangladeshi Migrant	-0.270*** (0.0401)	-0.293*** (0.0401)	0.00297 (0.0183)	-0.00941 (0.0142)	-0.0925* (0.0438)	-0.107* (0.0438)	-0.0788** (0.0281)	-0.0349 (0.0248)
Pakistani/Bangladeshi UK born	-0.292 (0.215)	-0.336 (0.215)	0.0108 (0.0142)	0.0193 (0.0186)	-0.105 (0.0809)	-0.182* (0.0809)	0.000517 (0.0162)	0.0802*** (0.0201)
N	555,492				557,154			
R-squared	0.8644				0.8267			

Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: ONS ASHE – 2011 Census.

Table 3.5: Decomposition entry and growth gaps in log hourly wages. Within and between firm and occupation effects: Men.

Men	Entry gaps			Growth gaps		
	Individual FE gap	Firm FE gap	Occupation FE gap	Within firm and occ wage growth to 45	Firm switching wage growth to 45	Occupation upgrading wage growth to 45
White Migrant	0.0192 (0.0214)	-0.0125 (0.00955)	0.0108** (0.00349)	0.0848*** (0.0180)	0.0209* (0.0106)	0.0116** (0.00406)
Black Caribbean Migrant	-0.0303 (0.0545)	-0.0629** (0.0242)	-0.0615*** (0.0102)	-0.215*** (0.0483)	0.0373 (0.0270)	-0.00748 (0.0135)
Black Caribbean UK born	-0.0494 (0.0320)	-0.0336** (0.0122)	-0.0468*** (0.00433)	-0.175*** (0.0225)	0.0305* (0.0140)	0.0166** (0.00541)
Black African Migrant	-0.256*** (0.0356)	-0.0331** (0.0107)	-0.0552*** (0.00400)	-0.0202 (0.0291)	-0.00112 (0.0123)	-0.0182*** (0.00523)
Black African UK born	-0.124 (0.0660)	-0.0438* (0.0222)	-0.0273** (0.00914)	-0.0283 (0.0582)	-0.0143 (0.0285)	0.00651 (0.0115)
Indian Migrant	-0.0691* (0.0297)	-0.0497*** (0.0115)	-0.0370*** (0.00594)	0.0323 (0.0272)	0.0475*** (0.0129)	0.0267*** (0.00663)
Indian UK born	-0.143 (0.0958)	-0.0178** (0.00622)	-0.0169*** (0.00240)	0.0388* (0.0173)	0.0320** (0.0106)	0.0107* (0.00486)
Pakistani/Bangladeshi Migrant	-0.221*** (0.0421)	-0.0442*** (0.00839)	-0.0354*** (0.00434)	0.0113 (0.0242)	0.0259* (0.0114)	-0.00272 (0.00627)
Pakistani/Bangladeshi UK born	-0.173 (0.233)	-0.0292*** (0.00574)	-0.0343*** (0.00274)	-0.0748*** (0.0201)	0.0245* (0.0101)	0.0218** (0.00690)

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: ONS ASHE – 2011 Census.

Table 3.6: Decomposition entry and growth gaps in log hourly wages. Within and between firm and occupation effects: Women.

Women	Entry gaps			Growth gaps		
	Individual FE gap	Firm FE gap	Occupation FE gap	Within firm and occ wage growth to 45	Firm switching wage growth to 45	Occupation upgrading wage growth to 45
White Migrant	0.0248 (0.0208)	-0.00693 (0.00842)	0.00781* (0.00380)	0.0431* (0.0173)	0.0246** (0.00916)	0.0116** (0.00406)
Black Caribbean Migrant	-0.122** (0.0455)	-0.0796*** (0.0130)	-0.0309** (0.00946)	0.00765 (0.0428)	0.0857*** (0.0159)	-0.00748 (0.0135)
Black Caribbean UK born	-0.0462 (0.0315)	-0.0196* (0.00951)	-0.0332*** (0.00354)	0.00836 (0.0247)	0.0258* (0.0107)	0.0166** (0.00541)
Black African Migrant	-0.201*** (0.0350)	-0.0332*** (0.00828)	-0.0316*** (0.00566)	0.0988** (0.0311)	0.0212* (0.0101)	-0.0182*** (0.00523)
Black African UK born	-0.0917 (0.0727)	-0.0498** (0.0174)	-0.0471*** (0.00557)	0.0355 (0.0452)	0.0479* (0.0220)	0.00651 (0.0115)
Indian Migrant	-0.108*** (0.0284)	-0.0278* (0.0126)	-0.0220*** (0.00651)	0.0464 (0.0258)	0.0142 (0.0136)	0.0267*** (0.00663)
Indian UK born	-0.0236 (0.0365)	-0.00579 (0.00619)	-0.00368 (0.00306)	0.0571** (0.0182)	0.00646 (0.00867)	0.0107* (0.00486)
Pakistani/Bangladeshi Migrant	-0.0556 (0.0430)	-0.0431*** (0.00937)	-0.0370*** (0.00690)	-0.102** (0.0321)	0.0480*** (0.0131)	-0.00272 (0.00627)
Pakistani/Bangladeshi UK born	-0.0619 (0.0687)	0.00717 (0.00594)	-0.0109** (0.00348)	0.0199 (0.0240)	0.00877 (0.0131)	0.0218** (0.00690)

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: ONS ASHE – 2011 Census.

Table 3.7: Decomposition of the migrant gap in log hourly wages by age of arrival

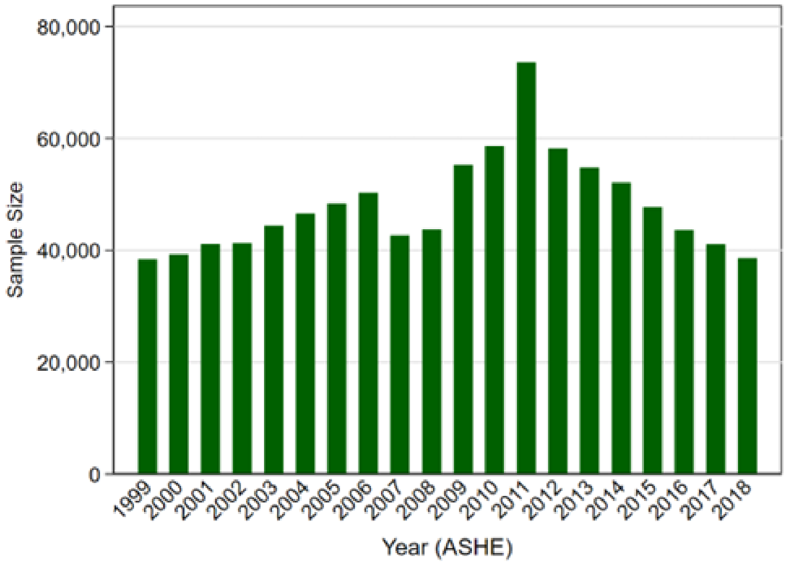
Men	White	Black Caribbean	Black African	Indian	Pakistani/ Bangladeshi
Ethnic gap	0	-0.217***	-0.300***	-0.116***	-0.164***
	(.)	(0.00638)	(0.0160)	(0.00576)	(0.00663)
Migrant Penalty					
Before age 10	0.0225***	0.0190	0.0581	0.0299**	-0.0000344
	(0.00526)	(0.0207)	(0.0341)	(0.0114)	(0.0133)
Age 10 - 18	-0.0141	-0.120***	-0.0335	-0.0380***	-0.0593**
	(0.0133)	(0.0249)	(0.0237)	(0.0112)	(0.0184)
Age 19 - 30	0.0532***	-0.0104	-0.0703***	-0.118***	-0.168***
	(0.0111)	(0.0253)	(0.0212)	(0.0119)	(0.0142)
Age 31 – 40	0.133***	0.0311	-0.173***	0.0399*	-0.156***
	(0.0174)	(0.0462)	(0.0207)	(0.0160)	(0.0192)
After age 40	0.0821***	-0.0979***	-0.215***	-0.124***	-0.176***
	(0.0178)	(0.0255)	(0.0191)	(0.0176)	(0.0301)
N	555848	555848	555848	555848	555848
r2	0.434	0.434	0.434	0.434	0.434

Women	White	Black Caribbean	Black African	Indian	Pakistani/ Bangladeshi
Ethnic gap	0	-0.0493***	-0.201***	-0.0476***	-0.0850***
	(.)	(0.00527)	(0.0119)	(0.00547)	(0.00682)
Migrant Penalty					
Before age 10	0.0296***	0.0201	0.0310	0.0499***	-0.00460
	(0.00449)	(0.0138)	(0.0295)	(0.0105)	(0.0139)
Age 10 - 18	0.0249**	-0.0991***	-0.00370	-0.0557***	-0.0948***
	(0.00812)	(0.0122)	(0.0188)	(0.00978)	(0.0222)
Age 19 - 30	0.0735***	-0.105***	-0.0888***	-0.202***	-0.107***
	(0.00893)	(0.0203)	(0.0155)	(0.00957)	(0.0221)
Age 31 - 40	0.00336	-0.171***	-0.0284	-0.0494**	-0.0851**
	(0.0163)	(0.0270)	(0.0170)	(0.0167)	(0.0328)
After age 40	0.0477***	-0.208***	-0.0632***	-0.141***	-
	(0.0128)	(0.0158)	(0.0148)	(0.0151)	-
N	557510	557510	557510	557510	557510
R-squared	0.432	0.432	0.432	0.432	0.432

Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Fixed-effects controls for: Age, Region-Year, Arrival cohort, and Highest qualification. Source: ONS ASHE – 2011 Census.

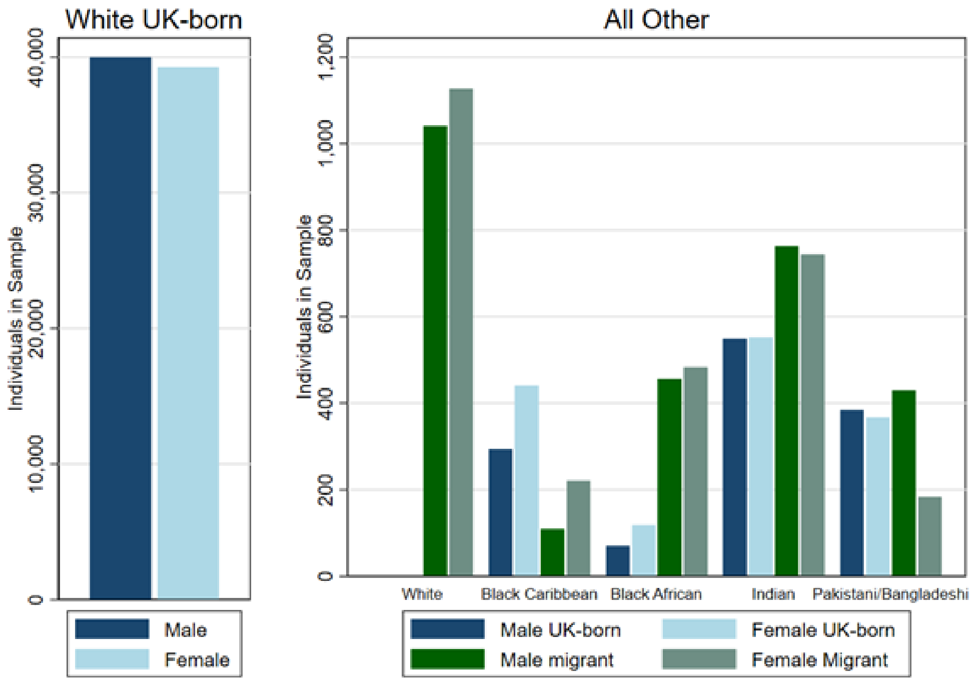
3.8 Figures

Figure 3.1: Sample size by ASHE year



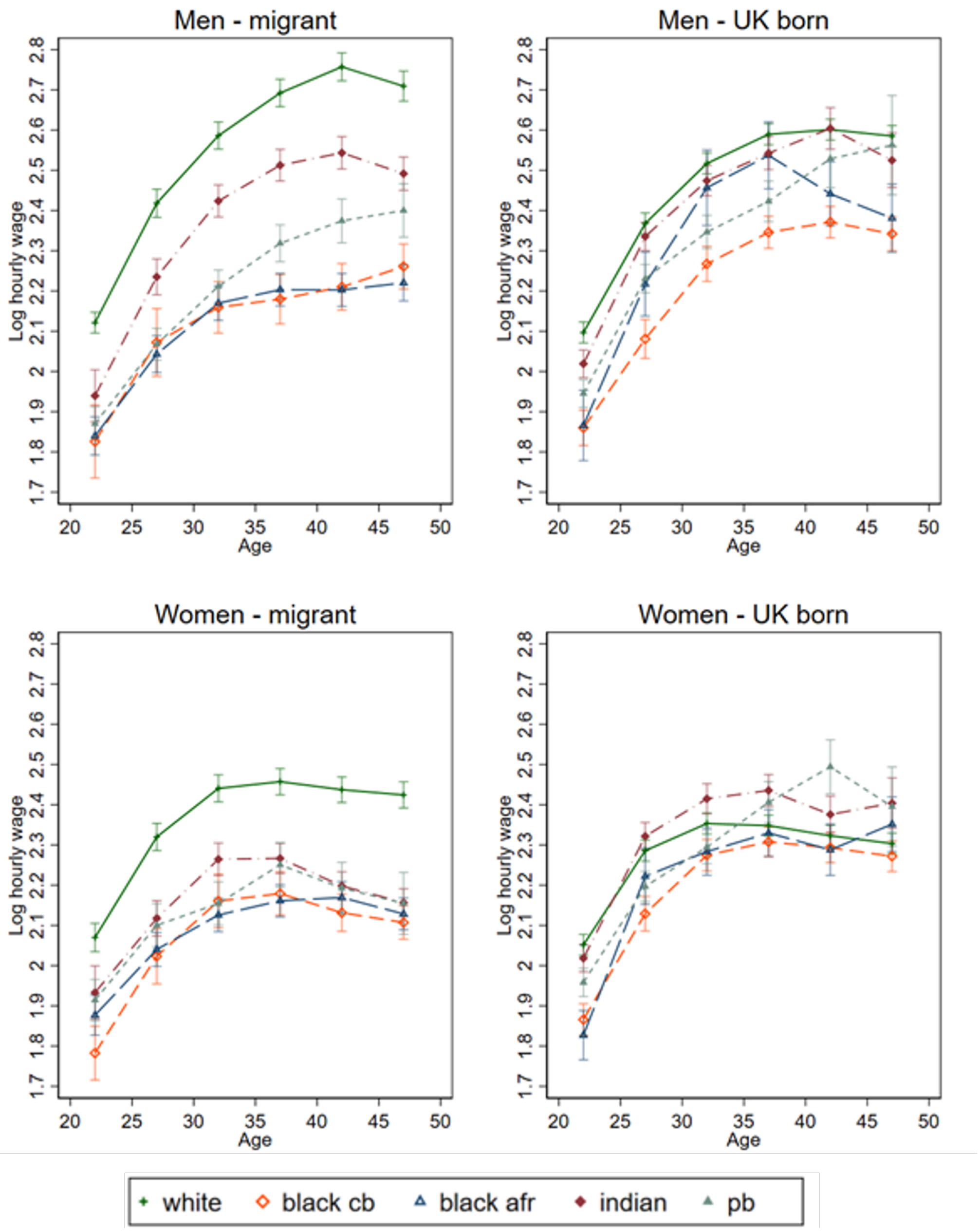
Source: ONS ASHE – 2011 Census

Figure 3.2: The Distribution of the Sample by Ethnicity, Gender, and Migration Status



Source: ONS ASHE – 2011 Census

Figure 3.3: Log hourly wages by ethnic group and age group



Year-region controls. Source: ONS ASHE – 2011 Census

Chapter 4

No Country for Hidden Men: Centrality and Control in Gang Networks

Joint work with Rui Costa, Magdalena Dominguez, and Matteo Sandi[1]

This chapter studies gangs in Brazil, an underexplored yet pervasive and volatile setting in organised crime. Using information from the intelligence police, occurrences and prison records, we identify and label gang communities by optimising the Markov stability of community structure. We estimate a network peer effect model that enables us to identify "key players" using the concept of intercentrality in social networks. We find a strong correlation between individual standard centrality measures, while model-based intercentrality highlights that some "key players" may have been missed in official hierarchical classifications and other centrality measures. The findings of this study will help assess and outline optimal disruption strategies in this setting.

[1]We are grateful to Antonio Padilha, Rafael Bernardini, Gisele Ferreira, and the RS Seguro team at the Cabinet of the Governor of Rio Grande do Sul, as well as the team at PROCERGS, for their invaluable support, assistance, and collaboration, which made this project possible. We also extend our sincere thanks to the many members of the Polícia Civil, Brigada Militar, Polícia Penal, and Instituto Geral de Perícias, whose insights significantly deepened our understanding of the local context and the data. The conclusions presented in this chapter do not reflect the official positions of the security agencies of Rio Grande do Sul. Any errors are solely our responsibility. This chapter received no financial support

4.1 Introduction

Organised crime groups, from street gangs to transnational cartels, constitute a pervasive global threat with immense economic and human costs (United Nations Office on Drugs and Crime, 2010). Across dozens of countries, these networks undermine governance, security, and development, making organised crime one of the most pressing political challenges worldwide (United Nations Office on Drugs and Crime, 2010). Illegal profits of transnational organised crime may exceed \$800 billion per year (on the order of several percent of global GDP), fuelling corruption and siphoning resources from legitimate economies (United Nations Office on Drugs and Crime, 2011). Whilst over 400,000 people are murdered annually around the world, roughly one-fifth of these homicides are attributable to gang or cartel violence (United Nations Office on Drugs and Crime, 2019). Since the early 2000s, organised crime and drug violence have killed as many people globally as armed conflicts, destabilizing countries and eroding the rule of law (United Nations Office on Drugs and Crime, 2019).

In Latin America, which accounts for about one-third of the world's homicides despite having less than a tenth of its population (United Nations Office on Drugs and Crime, 2019), homicide rates reach five to eight times the global average (World Bank, 2025), with crime and violence costing Latin American economies on the order of 3–5% of GDP (United Nations Office on Drugs and Crime, 2011; Inter-American Development Bank, 2017; Bisca et al., 2024). By corrupting institutions, driving away investment, and inflicting vast social harm, organised crime has emerged not just as a law-enforcement issue but as a core development and economic concern (Pinotti, 2015; World Bank, 2025). Due to the sheer scale and pervasive impact of gangs and organised crime worldwide, understanding their network structure and ultimately how these criminal networks may be disrupted is a first-order research question in economics.

This chapter provides novel empirical evidence on the network structure of Brazilian gangs. While the economics literature has made great strides in recent years in understanding the effects of organised crime on society, remarkably little evidence exists on the interactions within and between these organizations. While most studies focus on the outputs of criminal groups (e.g., crime, violence and corruption) rather than their internal workings, generating an understanding around the internal structure, hierarchy, and key players of these organizations is crucial for both research and policy. It has the potential to help us understand how these groups are actually organised and to inform strategies to target their key actors and combat their activities.

Drawing on exceptionally detailed administrative and intelligence data from the Security of State of Rio Grande do Sul, Brazil's southernmost state, this paper develops a network analysis of gangs operating within the state. In doing so, we present new empirical insights into gang structures and provide a novel contribution to a remarkably underexplored domain in the economics of crime. The unprecedented richness and detail of the data allow us to advance the existing literature in this area, which has previously lacked access to such granular administrative microdata. Moreover, Brazil is an apt and urgent setting to conduct this analysis, as its prison gangs and favela militias are powerful organised crime entities that have been relatively understudied (Misse, 2019).

Using an integration of intelligence data, prison stay data, occurrence data, and detailed data on individual characteristics, we apply a network community-detection model (Markov Stability) to identify communities (network subgraphs) within Brazilian gangs. We then analyse these communities with various network statistics to discern the potential hierarchy and roles of individuals within these criminal structures. By comparing these network-derived insights with official designations, such as known leadership roles, we provide evidence that the conventional wisdom about these gangs' hierarchy is incomplete. Only 9.8% of the top bosses identified by the intelligence data are identified as key players. Although some of these individuals hold relatively low rank in the network, their extensive connections across subgroups could position them to be key liaisons. Removing them from the criminal organization may be particularly effective.

These findings contribute to the economics literature studying organised crime, as they refine our understanding of gangs' interactions in both Brazil and other countries by pinpointing key players and network substructures that could be targeted to significantly disrupt criminal networks. By delving into the internal anatomy of organised crime groups, the contribution of our approach is twofold: firstly, it incorporates methods from network analysis to study gang structures and hierarchies; secondly, and more practically, it provides actionable intelligence on how law enforcement might remove key individuals from existing gangs more effectively by targeting those who hold the network together (even if they are not obvious kingpins). By shedding light on the internal workings of Brazilian organised crime, this study advances the broader agenda of understanding how these illicit organizations function and how they can be destabilized. In sections 4.2 and 4.3 we further outline how our work complements the existing literature that has documented why these organisations are so harmful to economies and societies.

The rest of the chapter is structured as follows. Section 4.2 reviews the existing literature on the economics of organised crime, seeking to highlight how the existing literature has focused more on its economic costs than on its microinteractions, while Section 4.3 focuses specifically on organised crime in Brazil, the empirical setting of this study. Section 4.4 introduces the data sources and Section 4.5 describes the methodology and results our empirical approach. Finally, Section 4.6 concludes.

4.2 The Economics of Organised Crime: A Literature Review

Organised crime groups, from mafias and cartels to street gangs, govern millions worldwide (Blattman et al., 2024). While organised crime may emerge to fill enforcement or protection voids when the state is absent or weak, state and gang rule could also be strategic complements. Evidence from Medellín suggests that gangs could minimize seizures and arrests by keeping neighbourhoods orderly and loyal, and that increasing state presence could increase incentives for gang rule (Blattman et al., 2024). While similar duopolies of coercion might emerge also in the presence of strong states, organised crime often imposes a high cost to society. While providing extralegal protection for businesses, mafias and gangs typically act as "primitive states", exercising territorial control and corrupting institutions (Gambetta, 1993; Skaperdas, 2001). While maximizing profits from illegal activities (e.g., drug trafficking, extortion), organised crime often imposes negative externalities like violence and corruption on society.

Our chapter builds on the existing theoretical models that attempt to generate an understanding of how these networks function. For example, Baccara and Bar-Isaac (2008) model the internal organization of criminal groups, highlighting a trade-off between efficiency and exposure: widely diffusing information inside a gang improves coordination but increases vulnerability to law enforcement. Their model rationalizes why some organizations adopt centralized hierarchies while others operate in decentralized "cell" structures as a means of limiting risk. Kugler, Verdier, and Zenou (2005) treat multiple criminal groups as competing firms that invest in corruption; when law enforcement is weak, tougher policing can perversely increase crime by driving gangs to bribe officials more, thereby reducing actual punishment.

These theoretical contributions underscore that organised crime has organizational and strategic dimensions distinct from ordinary crime. Empirically, however, studying gangs and mafias has proven challenging due to data scarcity. Levitt and Venkatesh (2000) constitutes an exception, as it analyses detailed financial records of a U.S. drug-selling gang. It shows that the gang's labour market resembles a tournament, with foot-soldiers earning only slightly above legal wages and gang leaders earning far more. The gang's pay structure appears highly skewed, and the tiny wage premium for low-level members does not appear to compensate for the enormous risks of death or imprisonment faced by gang members, suggesting that aspirants do not join for current pay but for the small chance of 'promotion' to a lucrative leadership role.

While very few systematic datasets exist on the internal hierarchy or network of a criminal organization, Mastrobuoni and Patacchini (2012) and Mastrobuoni (2015) use declassified FBI files on the 1960s Italian-American mafia network to map connections among hundreds of *mafiosi*. Their findings show an extremely hierarchical network, where a few bosses and captains are highly central, connecting to many others, while most members have few links (largely to their direct superior). Mastrobuoni (2015) also shows that a mobster's position in this network predicts his economic success, thus providing rare quantitative evidence of the payoff to climbing a gang's hierarchy, and suggesting that traditional economic incentives (higher earnings) are also present inside criminal organizations.

Our chapter contributes to these studies by moving beyond static hierarchies and trying to develop an understanding of how one can elicit important members of criminal organisations based on criminal cooperation dynamics. While some ethnographic accounts of mafias indicate that kinship and ethnic ties often underlie hierarchies (for example, the Sicilian mafia's reputed reliance on family lineage and patron-client relationships), these issues have been largely speculative in economics.[2]

Crucially, identifying "key players" within criminal networks remains challenging. The concept, introduced by Ballester, Calvó-Armengol, and Zenou (2006), seeks to identify the individual who leads to the largest reduction in crime when removed from the network. Traditional law enforcement relies on known hierarchies, but the existing network analysis reveals hidden important actors. For instance, Mastrobuoni (2015)'s network study shows that some mobsters with no famous status have high centrality and they appear deeply embedded in cooperation structures. The analysis in this chapter follows a similar logic: by mapping networks of communication and co-offending in Brazilian gangs, it aims to pinpoint individuals who serve as critical connectors or influencers, even if they are not labelled bosses by official intelligence. Our findings suggest that some "key players" have indeed been missed by official hierarchical classifications, despite our inter-centrality measures flagging them as potentially important.

Study of criminal networks is inherently linked to labour markets and local communities. Starting from Becker (1968), a large literature has studied the interaction between legitimate labour market opportunities and criminal participation. Organised crime can both respond to labour market conditions and actively shape them. On one hand, poor legal employment prospects, especially for youth, can push individuals into gang activity. Gangs often recruit from populations facing unemployment or low education, effectively acting as employers of last resort in deprived communities. In U.S. cities, *"alienation from legitimate labour markets"* drives some youth to seek income in gangs (Levitt and Venkatesh, 2000). In Peru, children exposed to higher-paying illegal coca farming (during price booms) face an increased risk of acquiring "criminal human capital" early and engaging in criminal careers as adults (Sviatschi, 2022).

On the other hand, organised crime itself may affect local communities by discouraging economic activity and electoral outcomes through business extortion and violence. In Italy, the presence of the mafia in certain southern provinces led to 16% lower GDP per capita over the post-war decades (Pinotti, 2015) and public re-

[2]In a working paper, Calderoni et al. (2020) use wiretap data on an Italian mafia to analyse promotion dynamics, finding that those who act as bridges between different communities of the network often rise to leadership, with brokers essentially becoming bosses.

sources being siphoned into patronage and away from useful infrastructure in Sicily (Dimico, Isopi, and Olsson, 2017) while also reducing the growth of new firm registrations (Detotto and Otranto, 2010). In municipalities with mafia presence, public funds (e.g., business subsidies) are misallocated to firms with mafia ties (Barone and Narciso, 2015), and mafia-tainted areas exhibit abnormal voting patterns and worse governance outcomes, consistent with mob interference in democracy (Daniele and Geys, 2015). In El Salvador, an abrupt expansion of gang control due to deportation of Los Angeles gang leaders to the state led residents inside gang-controlled neighbourhoods to earn significantly less income and attain lower education (Melnikov, Schmidt-Padilla, and Sviatschi, 2022). The mechanism at play appears to be that gangs restrict labour mobility, as they enforce territorial borders that prevent locals from commuting to jobs elsewhere in the city, in turn reducing employment options and depressing wages.

Starting from Schelling (1971)'s insight that organised crime tends to use *just enough* violence to conduct its business, evidence of the use of intimidation or violence by criminal organisations to influence local communities has also emerged in several contexts. In the Mexican Drug War, when Mexico's federal government shifted to a hardline stance (under the PAN party), municipalities that narrowly elected PAN mayors experienced sharp increases in drug-related homicides, suggesting that crackdowns weakened incumbent cartels in those towns, inviting violent turf wars as rival trafficking groups fought to fill the resulting power vacuums and to "usurp" the territory (Dell, 2015). In Italy, the presence of organised crime is associated with abnormal spikes in violence against politicians before elections, particularly when the electoral outcome is more uncertain, which in turn reduces voting for parties opposed by criminal organizations. Violence by the Sicilian Mafia also reduces anti-Mafia efforts by members of parliament appointed in Sicily, particularly from the parties that traditionally oppose the Mafia (Alesina, Piccolo, and Pinotti, 2019).[3]

One key difficulty in fighting organised crime groups and their illicit activities originates from the fact that prohibition may create monopoly power and rents that organised crime can exploit. For example, higher enforcement may raise drug prices by restricting supply, which could in turn increase profits for remaining traffickers, implying that moderate levels of enforcement might enrich cartels by eliminating competition, whereas extremely harsh enforcement could eliminate the market entirely (Becker, Murphy, and Grossman, 2006). In fact, there is evidence that a major interdiction effort in Colombia, which caused a shortage in cocaine supply, led Mexican cartels to fight more violently over the shrinking supply (Castillo, Mejía, and Restrepo, 2020), again highlighting the link between market forces and gang behaviour.

In sum, organised crime imposes multifaceted costs on society by distorting labour

[3]This complements Dal Bó, Dal Bó, and Di Tella (2006), who find that an increase in police presence in a context of high corruption had little effect on mafia activity in a region of South America, implying that when criminal groups have infiltrated the state apparatus, they can nullify enforcement by subverting the enforcers.

markets, fuelling violence, capturing political institutions, exploiting illicit trades, and hindering economic development. The existing literature has made progress in quantifying these effects and in theorizing about criminal organizations' behaviour under various incentives. Yet, a recurring theme is that we mostly observe what organised crime causes, rather than its microinteractions. Largely due to the scarcity of available data, the literature lacks detailed insight into who holds power within gangs, how networks are structured, and how that internal structure influences the gang's functioning and its vulnerabilities. Addressing this gap is not only of academic interest but also of practical importance for designing effective disruption strategies.

4.3 Organised Crime in Brazil

Brazil's contemporary gang landscape is rooted in the rise of prison-based criminal factions that now dominate urban organised crime. In the 1980s, overcrowded and chaotic prisons became incubators for gang formation (Misse, 2019). In Rio de Janeiro, groups like the *Comando Vermelho* (CV) emerged in prison as self-protection networks and later expanded into the drug trade of the favelas (Misse, 2019). São Paulo's *Primeiro Comando da Capital* (PCC), founded by inmates in the early 1990s, consolidated power throughout the prison system of São Paulo and eliminated rival factions (Misse, 2019). These prison gangs provided governance behind bars, enforcing codes of conduct and mediating conflicts, paradoxically reducing inmate-inmate violence (Dias and Salla, 2013). The PCC established itself as a centralised authority within São Paulo prisons, virtually monopolising dispute resolution and sharply curbing internal violence among prisoners (Dias and Salla, 2013). This newfound order came against a backdrop of mass incarceration: Brazil's inmate population exploded from around 60,000 in 1980 to more than 800,000 by 2019 (Misse, 2019), overwhelming the state capacity and enabling gangs to fill the void.

Outside prison walls, these factions quickly evolved into powerful urban criminal enterprises. The PCC and CV "spilled out" of the penitentiary system to dominate drug retail, extortion, and other illicit markets in cities (Stahlberg, 2021). They replicated their hierarchical yet federated structures on the streets, maintaining discipline through prison-based incentives and sanctions and leveraging the threat of prison retaliation to control street-level criminals (Lessing, 2017). By now, what began as prison self-protection rings have matured into sophisticated mafias that blur the line between carceral and street organised crime. While Brazil's gangs emerged in city slums and penitentiaries, organised crime has also proliferated in rural areas and borderlands, where criminal networks engage in illicit trades such as drug trafficking corridors, illegal mining, land grabbing, and arms smuggling. Major urban factions like the PCC have increasingly "colonized" these lucrative frontiers (Lessing, 2017).

Today, Brazil's urban gangs exhibit a mix of violent competition and negotiated order. On one hand, turf wars between rival factions drive extreme violence – notably the periodic CV–PCC confrontations and their offshoot conflicts. On the other hand, when a single gang achieves monopoly in a locale, a fragile peace can prevail. Recent research finds that gang fragmentation versus domination is, in fact, a key determinant of violence levels. In particular, the presence of multiple competing gangs in a state is associated with significantly higher homicide rates, whereas dominance of a single-gang correlates with lower violence (Stahlberg, 2021).

In response to this, a sizeable amount of evidence suggests that, by sweeping more young offenders into prison, crackdowns and harsher incarceration policies may unintentionally strengthen gang influence and expand the gangs' recruiting grounds and their coercive reach over communities (Lessing, 2017). Historical surges in imprisonment (often via anti-gang laws) in fact preceded leaps in PCC and CV power, as these groups gained the capacity to "govern and tax criminal markets" and even orchestrate violence as leverage over the state (Lessing, 2017). Conversely, strategic withdrawal of state presence could trigger violent instability. In fact, episodes of police strikes in Brazil have shown that within days of police withdrawal homicides surged by approximately 45%, overwhelmingly in gang-controlled areas, with the vast majority of victims being individuals tied to gang activity (Mancha, 2025). This in turn underscores the precarious balance between the state and gangs of Brazil: gangs may usually avoid high-profile violence to keep a low profile, but when the state "steps down", violent anarchy can erupt in contested gang turfs generating a highly non-linear relationship between police action and gang behaviour. Too little state presence invites gang warfare, yet indiscriminate repression could strengthen gangs in the long run.[4]

The entrenchment of gangs and organised crime in Brazil has wide-ranging implications for economic and social development. Monteiro and Rocha (2017) exploit variation in the timing and location of armed turf battles between drug factions to show that eruptions of gang violence significantly depress students' learning outcomes. Pupils exposed to nearby shoot-outs and prolonged conflict see their test scores drop and attendance falter, with disruptions extending beyond the immediate conflict period. Moreover, analysis of the "economics of the favela drug trade" using survey data from Rio de Janeiro gang members reveals that youths drawn into gangs tend to come from significantly poorer backgrounds and often have histories of school failure and early drug use, suggesting that lack of lawful opportunities and early-life disadvantages funnel individuals into criminal careers (Carvalho and Soares, 2016). Within the gang hierarchy, Carvalho and Soares (2016) also document an earnings structure similar to a tournament: wages are not tied to formal education, but they do rise with experience and especially with engagement in violent roles. This "promising" career trajectory is, of course, shadowed by extreme risk: the two-year fatality rate among the gang recruits in their sample is around 20% – a mortality hazard orders of magnitude higher than in legal jobs.

[4] This has led some to advocate focused deterrence and negotiated community policing as more effective means of reducing gang violence, citing temporary successes with Rio's Pacifying Police Units (UPPs) in the early 2010s (Shea, 2015).

These findings highlight a tragic equilibrium in Brazil's slums: gangs stunt human capital accumulation while offering immediate economic rewards and a semblance of status for disenfranchised youth, yet the prospects for longevity are grim. From a theoretical standpoint, this resembles a high-risk/high-reward segment of the labour market, sustained by a steady supply of desperate entrants and the gang's ability to enforce loyalty (e.g. through "protection" of members' families or the threat of retribution for desertion).

Rio Grande do Sul, the empirical setting of this study, exemplifies these dynamics on a localized scale. In Rio Grande do Sul, drug markets have intensified for the control of cocaine since the 2000s (Cipriani, Lien, and Santos, 2023) and today they are contested by numerous home-grown factions (with *Os Manos* among the oldest), which are locked in conflicts both inside and outside prisons (Misse, 2019; Cipriani, Lien, and Santos, 2023).[5] This profusion of factions has sometimes turned the state's prisons and border cities into battlegrounds. State authorities have grown alarmed at factions like *Os Manos* forging alliances with the PCC and expanding operations across the Uruguayan border.[6] Thus, even far from Rio and São Paulo, the prison-gang phenomenon has penetrated Brazil's peripheries, underscoring the nationwide diffusion of the *facção* model (Misse, 2019). This cross-border dimension has made Rio Grande do Sul a strategic node for transnational criminal logistics (Silva, 2021).

The factional landscape in Rio Grande do Sul remains fragmented, with over a dozen identified organizations engaged in violent turf competition, arms smuggling, and retail drug markets. In recent years, Rio Grande do Sul has seen substantial increases in organised crime–linked violence, particularly in municipalities with active drug factions and porous borders (Fórum Brasileiro de Segurança Pública, 2023). While PCC expansion may have brought some localized order through monopolization, especially in the south, this has often come at the cost of deeper criminal penetration in local governance and public contracting (UK Home Office, 2025). Breaking this cycle is a core concern of current policy debates and research: by studying the internal structure of gangs in Rio Grande do Sul and identify their key actors, this study makes an effort to rectify this omission.

[5] Some studies argue that drug markets result from the expansion of crime as a consequence of the crisis of the State, identifying criminal groups as an "organised counter-government" (De Simone, 2014), a "control over the entire State" (Smulders et al., 2017) or an "illegal governance" (Campana and Varese, 2018). However, other studies criticize the notion that crime is a reflection of the ineffectiveness of the criminal justice system and the administration of justice (Silva, 2004; Garzón, 2012; Sain, 2015; Feltran, 2018; Misse, 2019).

[6] See, for instance, https://insightcrime.org/brazil-organized-crime-news/rio-grande-do-sul-brazil/

4.4 Data

Our network analysis combines extremely rich administrative and intelligence data from police forces and Rio Grande do Sul institutions covering the 2014-2024 period at an unprecedented level of detail. These data include and merge five key datasets at the individual level: police occurrence data, police intelligence data, prison stay data, prison intelligence data, and the population registry[7].

Police Data

Data on occurrences were obtained from the police computer records of *Policia Civil* (i.e., the state-level investigative police) and cover all the incidents registered in Rio Grande do Sul, resulting in more than 12 million observations since 2014. Each occurrence is identified with a unique id linked with detailed information about the occurrence: type of crime/incident, date and hour of the occurrence, date and hour of registration by police officers, location, participants, role of participants (victim, offender, witness, etc.), physical state of participants (injured, dead, etc.). Since the participants of each occurrence and the nature of their participation are observable in the data, we are able to identify co-offenders to form our network links. Our analysis includes all the occurrences since the beginning of 2014 that involve individuals who belong to our network of interest.

Sensitive, restricted-access police intelligence data were also made available for this study. These data include individual-level information about organised crime membership, area of action, and hierarchical position within the organised crime group. For each individual in the state, the data contain information on whether the individual is known to be affiliated to a criminal organisation, which group the individual belongs to, the position of the individual in the hierarchy of the criminal organization, and where and how the individual is known to operate.[8] These data are collected and curated by the intelligence officers of police forces and they inform policing strategies designed by the security services in the state to tackle organised crime. The sample of 4,072 individuals present in the data is skewed towards highly ranked members of organised crime group, due to the primary use of this source by the police operators.[9] Although the data include individuals acting as money launderers, drug logistic suppliers and other relevant roles in criminal organisations, we focus mainly on the state, region and municipality leadership to guide the construction of our network.

The second police intelligence data set used within this analysis is a list of gang-

[7]Table D.3 summarises each of these data sources.

[8]Due to the extreme sensitivity of the data, we are not permitted to name the exact data titles and sources.

[9]Within the police's intelligence, around 60% of named individuals are assigned a gang hierarchy. Within this hierarchy, around 13% are of high rank, 57% are middle-ranked, and 30% are lower ranked. Given the pyramidal structure of most gang hierarchies, the highest ranked individuals are over-sampled.

related offences that were compiled based on a survey administered to several police officers who specialise in organised crime. The resulting list summarises the set of offences that are disproportionately associated with gang activity and includes crimes such as drug trafficking, robbery, extortion, and kidnapping. A more detailed summary can be found in table D.1. These gang-related crimes represent 15% of all recorded crimes within our data. However, within our network of co-offenders, gang-related crimes account for 48% of their total recorded offences.

Prison Data

Prison data for the universe of inmates are collected administratively through prison police and management services (i.e., *Superintendência dos Serviços Penitenciários* - SUSEPE) with a high level of detail. These microdata cover all inmates in the prison system of the state of Rio Grande do Sul and record all their movements. In particular, we are able to observe all the movements of inmates within and between prisons (e.g., transfers, isolation periods, medical visits, temporary releases, exits for sentencing hearings) and all the entries and releases that correspond to the start and end of prison sentences. We have individual-level information on a wide array of characteristics of each inmate both at the time of entry into prison and during the prison stay: for example, we have records both at the time of imprisonment and while in prison of their tattoos, which in Brazil often indicate the organised group one belongs to, the lawyer that assists them, which may also be indicative of the organised group one belongs to, and their stated compatibility in prison, i.e., which organised crime group they deem themselves to be compatible with at the time of entry. We can observe the visits inmates receive, including the date of the visit, the relationship between the visitor and the inmate (e.g., partner, mother, father, sibling, daughter, son, friend, etc), the items brought to the inmate, whether permitted (e.g., food and medicines) or not (e.g., drugs and weapons). We also have information on fights and other criminal occurrences within the prison, time spent in solitary confinement, and details of medical treatment. Importantly, on any given day, we can track the location of the inmate inside the prison system at the level of the prison gallery cell. Our sample includes over 3.1 million movements since 2014, and details the stays of 168,748 inmates within 5,855 gallery cells across 113 prisons.

We also have privileged access to two sources of prison intelligence data: individual risk profiles and prison-gallery gang controlled mappings.[10] Risk profiles are attributed to specific inmates who require special arrangements for their own security and safety, as well as the security and safety of the prison officers who interact with them. The risk profiles capture several dimensions including safety risk due to the nature of the crimes committed (e.g., sex offenders cannot be allocated with other inmates as their lives would be in danger), former occupation (e.g., former police officers require special arrangements to ensure their security in prison) or criminal profile

[10]Due to the extreme sensitivity of the data, we are not permitted to name the exact data titles and sources.

(ex: leaders of organised crime groups pose security and safety concerns with terms of escort needed while being moved). From a total of 974 inmates with risk profiles, we identify 662 individuals with a risk profile compatible with leadership roles in organised crime to inform our network construction.[11] The offence list of these 662 individuals includes 8,902 crimes, 4,663 of which are flagged as gang-related.

The second set of sources from the prison intelligence are maps of prison galleries that are controlled by organised crime gangs and information on which gangs control each gallery. From 2017 to 2024, we have a yearly mapping of which galleries are controlled by which gangs.[12] For galleries with no gang affiliation, we also observe whether they are signposted for a particular use (e.g. a medical ward) or to house certain types of prisoner (e.g. religious or political prisoners). For all galleries, we observe their designed capacity. Considering the important role of prisons in the origins of gangs in Brazil, for which Rio Grande do Sul is no exception, data on the control exerted by gangs inside prison is of primary relevance in our study of the gang affiliation of individuals.

Population Registry

Finally, all the data described so far are linked at the individual level to the main registry of residents and people born in Rio Grande do Sul collected by the General Institute of Forensics (i.e., *Instituto-Geral de Perícias*, IGP). This dataset includes the near universe of people born or residing in the state, approximately 24 million individuals across our time period. In the data, we have individual information regarding year of birth, year of death, sex, race, marital status, and municipality of birth and residence. We also have access to detailed information regarding their education and assets. Importantly, this data source is used as "masterkey" to link and identify every individual in the other datasets as it contains full names, parents' names, federal fiscal identification number and federal person identification numbers for each individual. These data allow us to link every individual, including inmates, to their family members and their respective rich array of information contained in the datasets described above.[13] Additionally, even individuals not born or residing in Rio Grande do Sul enter the registry as soon as they have a police occurrence and/or a prison stay.

[11] We consider individuals labelled by the prison intelligence as: organised crime group leaders, narco leaders, prison gallery leaders, gang leaders, and those with substantial negative influence.

[12] In some cases, we have bi-yearly gallery mappings.

[13] The identification of individuals across data sources is made by the state's data processing expert team. Our research team works with anonymized identifiers resulting from the data processing.

4.5 Methods and Results

4.5.1 Network Construction

The combination of the merged data described above allows us to study organised crime structures and interactions in the streets and inside prisons with an unprecedented level of detail. We start our analysis by constructing a network of criminal individuals involved in organised crime in order to explore their co-offending patterns. More formally, social network analysis (SNA) is a process through which social structures are analysed using graph theory and its visualization. This approach fully characterizes the existing networks in terms of their agents (called nodes) and links. The first step towards an SNA is to build an adjacency matrix. In the simplest case, this matrix G is a square matrix of dimension n (the total number of nodes), and the element g_{ij} (known as a dyad) equals 1 when there is a link from node i to node j, and 0 otherwise. In this case, we create an adjacency matrix in which the g_{ij} element is equal to 1 when individuals i and j are co-offenders. Such a task can be challenging due to its dimensions.[14] Although this approach is merely descriptive of the dataset and potential networks, it is also extremely enlightening and provides a first approach to the existing criminal links.

Within the network, each node is a suspected gang member and each link between nodes represents a criminal occurrence in which the two linked individuals co-offended. We identify these links using criminal occurrences from January 2014 to September 2024. We restrict our analysis to the set of gang-related offences outlined by the police intelligence.[15] To identify the nodes of the network, we begin with a group of 662 high-profile gang members who have been identified by the police and prison intelligences. These high-ranking individuals are identified as gang leaders within the gang hierarchy and are denoted as 'zero order' nodes within the network. In particular, we consider identified leaderships at state, region, municipality and prison levels to form the 'zero order' group.

The remaining nodes within the network are identified by 'growing out' the network from our initial list of zero order nodes using their co-offences (links). In other words, we identify additional gang members that the prison intelligence did not specify by observing who the specified members co-offend with. We define first order nodes as any individual that has co-offended with a zero order individual in a gang-related occurrence (as defined using the police intelligence data) and is not already included in the network. The network is then expanded by iterating the process to identify second-order nodes (those that co-offend with first-order individuals).

[14] Every time the network is grown by including the co-offenders of each existing node, the number of nodes expands exponentially. As the network matrix is nxn, its dimensions also expand rapidly.

[15] To prevent the possibility of two members of opposing gangs being connected as co-offenders, we exclude violent events where co-offenders could have been fighting against each other. These are defined as violent events that do not occur at the same time as other crimes (such as robbery), and violent events where an individual is simultaneously defined as an offender and a victim.

Grown from an initial set of 662 individuals, our network contains 28,159 individuals. This includes 7,455 first order members and 20,042 second order members. In table 4.1, we show that individuals within this network are typically involved in multiple serious crimes and 70% have been to prison. The average individual has 10.2 recorded crimes, 4.8 of which are gang related, including 1.6 drug and trafficking offences and 0.7 murders. The average individual has also stayed in 8.5 different prison galleries. There is notable heterogeneity between the different orders of individuals; order zero individuals are older, have stayed in more prison galleries (20.3), and have committed more murders (2.3). All crime statistics decrease as order increases.[16] This pattern is consistent with order zero representing more senior gang members with a larger number of severe gang-related occurrences, and higher order individuals representing lower ranked and less experienced gang members.

4.5.2 Assigning Individual Gang Labels

Considering the contextual importance of in-prison interactions for organised crime in Brazil, each individual within the network is assigned a gang affiliation based on the history of their prison stays. Within Brazilian prisons, galleries are often associated with one or more gangs that show compatibility and offer some guarantee of peaceful cohabitation. Individuals may express a preference to stay in a certain gallery if they are already affiliated with that gang or if they live within an area controlled by that gang. Not all gangs have isolated galleries in every prison of the state; however, prison police services triage aims to ensure compatibility between gang affiliation and the leadership of the gallery where the inmate is placed - with life preservation being the first priority in gallery assignment. Some prisons have non-gang aligned galleries but, due to capacity constraints, it is not guaranteed that an inmate without gang affiliation can be always placed in these neutral galleries. For non-aligned individuals placed into a gang controlled gallery, developing an affiliation with that gang in order to gain protection and other privileges whilst in prison is almost inevitable (Lessing, 2017). Therefore, we begin by assuming that an individual's prison stay history is a good proxy of their gang affiliation.

We use data for all prison stays within the state between January 2017 and September 2024 to measure individual gang affiliation. First, for each day an individual spends in a gang controlled gallery, they are assigned the corresponding gang label from the intelligence mapping that is closest in time. The 11 mappings of prison gallery affiliations are described in Table D.2. They collectively provide labels for the stays of 115,098 inmates in 4,219 different gallery cells across 76 prisons, naming 39 different gangs. To improve labelling, very small regional gang labels are replaced with their larger state-level counterparts.[17] Our gang labelling exercise uses

[16] All differences between groups are statistically significant at the 1% level, excluding the share that are female

[17] A detailed discussion of the methodology used to map smaller local gangs to larger gangs can

115,639 unique gang-affiliated stays in 3,565 gallery cells for 14,125 members of our network. For individuals within the network with no prison stay history, we do not infer their individual gang affiliation.

The individual's gang affiliation is then determined by comparing the share of their total prison time spent in galleries associated with each gang. For an individual to receive a 'unique gang label', for example gang A, they must have spent at least 60% of their time in gang-affiliated galleries in gang A galleries. Individuals may receive a 'multiple gang label', for example gang A and B, if multiple gangs meet this condition. This is possible when two gangs coexist within a gallery. An individual will have no label if no gang meets these criteria, or if they have spent less than 30 days in gang affiliated prison galleries.

As the next section assigns gang labels to entire communities within the network, it is important to emphasise that individual gang label assignment is not influenced by any features of the network structure. The information used to infer an individual's gang affiliation stems from individual prison stay data and prison intelligence of galleries, whereas the information used to construct the network comes from police occurrence data and prison intelligence of individual profiles. This separation helps ensure that the labels assigned to communities within the network are not simply reinforcing a single source of existing intelligence.

4.5.3 Community Detection

Community Detection - Markov Stability Model

Since our aim is to study the criminal interactions within and between organised crime groups, one of the main challenges we face is that of analysing large networks without a prior labelling of groups that compose them. Detecting communities/clusters of individuals at a granularity scale that best approximates the "ground" interactions between agents and groups is a complex task (Schaub et al., 2019). In our context, this is particularly relevant since we know that organised crime members operate mainly through gangs with coordinated actions at different spatial levels (cities, regions, prisons, etc.).

As researchers, the challenge is even greater since we rely on observed linkages between agents of the network (co-occurrences) to back up the true communities. In fact, the literature on community detection has established that an optimal algorithm for all possible community detection tasks does not exist (Kleinberg, 2002; Peel, Larremore, and Clauset, 2017). Community detection modelling therefore needs to be goal-specific considering the properties of the network and subsequent partitions that matter the most to the application at hand.

With the objective of identifying communities of individuals that reflect group affiliation and cooperation dynamics, we take a non-deterministic approach with re-

be found in Appendix A.

spect to the best scale of aggregation (Peel, Larremore, and Clauset, 2017). A priori, it is not straightforward to assume which scale of pre-determined aggregation (i.e., neighbourhoods, prison galleries) may be best to study organised crime dynamics in this setting. On the one hand, it could be argued that considering each gang at the statewide level would be a reasonable scale, as it would group individuals that identify themselves under the same "umbrella" gang. However, we also know that these organizations sometimes operate at a local level (city and/or neighbourhoods) with a certain degree of independence. This results in regionalised crime dynamics which make statewide ties between criminal structures weaker. Therefore, we require a community detection modelling approach that allows for different scales of grouping of individuals and can be tested against the ground truth in order to select an optimal scale.

Leveraging on the concept of generalised Markov Stability for community detection, we apply an unsupervised community detection algorithm that identifies optimal partitions at different scale levels. In particular, a scale is Markov stable if the group of communities resulting from the network partition remains confined to the same regions of the network for a considerable amount of Markov time (walk length) and is robust to changes occurring locally and elsewhere in the network. Formally, optimal Markov stable communities are such that the distance between the Markov dynamics and the reference process, which is usually selected to be the stationary distribution of the Markov chain, is maximized for a given length of Markov time. The intuition behind the model is that resorting to dynamics of *random walks*, communities correspond to regions of the network where a "walker" ends up spending most of the time due to the denser links within communities and the sparse connections across communities. Since the model allows the "walker" to move between nodes based on the network transition matrix over different lengths of iteration, one can have multiple partitions of the network corresponding to local minima that satisfy the Markov Stability objective function for a sustained interval of Markov time. We use the python package *PyGenStability* that detects optimized communities of a network graph at different levels of scale (time) by maximizing generalized Markov Stability (Arnaudon et al., 2024).

Figure 4.1 shows the scales of the different robust optimized partitions estimated by the Markov stability algorithm.[18] A partition is considered robust for a given scale t if the average Normalized Variation of Information $NVI(t)$ across all pairs of partitions is low and the distance between the optimal partitions at scales t and t' remains low across scales $NVI(t, t')$.[19] One can observe that for our network there are

[18]The Markov Stability model requires the network to be fully connected in order to identify stable partitions, therefore our analysis is performed in the largest component (subgraph) of the entire network which corresponds to 27760 out of 28159 individuals of the original network.

[19]In information theory, Variation of Information between two partitions X and Y is given by: $VI(X, Y) = E(X) + E(Y) - 2MI(X, Y)$, where $E(.)$ stands for entropy and $MI(.)$ mutual information. Normalized Variation of Information, NVI, rescales VI by the join entropy of both partitions, $NVI(X, Y) = VI(X, Y)/E(X, Y)$

7 Markov scales with stable and robust partitions and correspondent communities. These scales imply significantly different levels of aggregation in the final communities: while the largest community of scale 1 has 80 individuals, the largest community of scale 7 has 26,917 individuals.

Community Detection - Labelling

After retrieving the sets of stable communities at different optimal scales, we use the information about the individuals that belong to each community to infer the dominant gang association for that community. In order to do so, we explore the individual prison gang affiliations defined in Section 4.5.2 to calculate the share of individuals with prison records associated with a certain gang through their prison gallery history within each community, $S_c^g = \frac{1}{N_c}\sum_{i=1}^{N_c} 1[I_i^g = g]$. We then create threshold majority rules [20] to assert which gang label to attribute to each community:

- If $N_c > 20$:
  1. If the highest gang share within community c is larger than 60%, then the community will be labelled according to the gang that corresponds to the highest share (Gang A)
  2. If there is no gang whose share is higher than 60%, then we consider the two gangs with the largest shares. If the sum of their shares is at least 80%, then the community will have a double label corresponding to each of the gangs (Gang A/B)
  3. If none the previous conditions are met, we conclude that there is not enough evidence to label the community and therefore it is considered "Mixed Uncertain"

- If $N_c \leq 20$:
  1. If the highest gang share within community c is larger than 60%, then the community will be labelled according to the gang that corresponds to the highest share (Gang A)[21]
  2. If none of the previous conditions are met, we conclude that there is not enough evidence to label the community and therefore it is considered "Mixed Uncertain"

[20] A potential concern about the proposed labelling method relates to its reliance on a certain predefined value of the thresholds. To ensure that labelling results are robust to the choices of threshold levels, we perform a sensitivity analysis with different values of thresholds while preserving the logic of majority rule and find robust results in terms of labelling.

[21] Due to the smaller size of the clusters and the potential for measurement error, we calculate the share of individuals in the whole cluster with a given gang label instead of conditioning the share to only individuals with prison stay labels.

We consider the possibility of dual labelling in larger communities for three context-specific reasons. Firstly, it is not uncommon for gangs to cooperate among themselves through alliances that can have both variability in terms of locality and duration. Two different gangs may cooperate in certain regions where both share common interests over territory and market control; this is often the case between the two biggest gangs with statewide expression and smaller localized gangs whose operations are more geographically concentrated. Secondly, as previously explained, not all prisons have the engineering capacity to have separate galleries to accommodate each gang; therefore, members of some gangs may be allocated to a gallery with inmates predominantly affiliated to another gang but that show "compatibility" in terms of alignment. These "compatibility" arrangements can be agreed upon between leadership outside or inside the prison and become a co-habiting norm. Thirdly, among lower ranks of the gang structure, there can be alignment changes resulting in members switching between gangs responding to different incentives (monetary rewards, power and prestige or survival and protection). As a consequence of the institutional features described above, there can be groups of individuals with prison gallery affiliations that differ but have occurrence links that place them within the same community in the network.[22]

We perform this labelling exercise separately for each of the 7 optimal scales resulting from the Markov Stability model. Tables 4.2 and 4.3 summarise the percentages of individuals and communities in the network by community label type at each optimal scale. One can observe that the percentage of communities with a dual label tends to increase with the scale of aggregation as communities become bigger and affiliation within community becomes more heterogeneous. Of particular relevance is the sharp decline in the share of single gang labelled communities at the highest scale of aggregation, in favour of a large increase in communities for which there is no clear majority of affiliation (even when considering two gangs) which are therefore labelled as uninformative ("Mixed Uncertain"). These patterns seem to indicate that there is an optimum scale that best approximates the true affiliation of individuals in the communities that lies within the middle range of network scale resolution.

Community Detection - Optimal Scale Validation

Having successfully identified and labelled criminal communities at different scales, we proceed to assess which scale best approximates the ground truth of gang affiliation of each individual. When addressing community detection problems, researchers rarely know the true affiliation of individuals in the network (unobserved labels), or the available labels describe individual characteristics that are imperfect proxies of affiliation (measurement error in the labels).

In our setting, we have the true affiliation for a subsample of individuals that come from the police intelligence databases. Importantly, our community detection model

[22]Table 4.2 shows that share of individuals with dual label fluctuates between 1% and 27% depending on the scale of the partitions

and labelling do not make use of the intelligence gang affiliation data as input at any stage of the process, and the police agencies compiling the intelligence list used for validation are not the same as those that produce the intelligence mapping of prison galleries used in our labelling methodology. Despite the fact that both police intelligences share a certain degree of information on these individuals, we know that these databases are kept separately and are not accessible to the members of the other police intelligence body without prior institutional request - strengthening our reasoning for using one of these data sources as validation.

We use this subsample of data from the police intelligence databases to perform validation of our clustering and labelling by calculating the accuracy of our model community labels with respect to individual intelligence labels.[23] The intelligence list of individuals eligible for validation is comprised of 4,072 observations, 2,165 of which are found in the constructed network. The individuals present in both samples form the final subsample to assess the performance of the labelling exercise. We consider a label to be accurate when the intelligence label and the label of the community in which an individual is placed match each other.[24]

In Figure 4.2, we plot the labelling accuracy rate at each optimal scale of the Markov Stability in our validation subsample. We can see that the accuracy curve approximates an inverse U-shape with accuracy being lower at both low and high levels of scale aggregation. We interpret this result in the following way. While very low levels of aggregation imply small sized communities which may not provide enough information to confidently label communities, large aggregation levels will merge communities with weaker connection linkages and bias the labelling exercise towards the gangs with larger membership. Indeed, when we look at the accuracy of labelling by different gangs in Figure 4.3 we can see that accuracy remains relatively high for the two biggest gangs in the state (Gang A and B), while medium-sized gangs accuracy falls sharply at higher scales as they begin to merge into communities of the two dominant gangs.

The scale with the highest accuracy rate of 71% is scale 112, which includes 243 communities of different sizes ranging from 5 to 449 members per community. The largest gang, gang A, shows a large accuracy rate of 85% which is not totally surprising considering its dominance in the organised crime landscape in the state. However, smaller gangs have accuracy levels not too far away from those of gang A: gang B shows an accuracy rate of 74% and gang C displays an even larger rate of 87%. Gang D, however, displays a lower accuracy rate of 59%. Overall, these metrics show that communities estimated in scale 112 are a good approximation to the true affiliation

[23]Although we are confident that intelligence affiliation has minimal measurement error, we acknowledge that the individuals represented in the subsample are skewed toward higher ranked individuals and so will the classification performance metric (accuracy) used in the validation. Table D.4 outlines the match rate by rank order.

[24]For the cases of dual labelled communities, we consider a label to be accurate if any of the two gang labels correspond to the individual gang affiliation in the intelligence data. We deflate these matches by a factor of 0.5 when counting accuracy.

and relevant reference group of organised crime in the state. Therefore, they are a useful dimension to consider in the peer effects estimation and key player analysis that follows.

4.5.4 Peer effects

With the aim of identifying the main actors, or key players, within our network and communities we proceed to identify and estimate a peer effects model which enables us to estimate the degree of feedback effects that each individual represents within the network structure. In short, peer effects indicate how much of an individual's behaviour stems from the behaviours of their peers. These effects have been shown to be relevant across several economic settings, and crime is no exception (see, e.g., Bayer, Ross, and Topa 2008; Liu et al. 2012; Lindquist and Zenou 2014).

When it comes to peer effects in networks, two different types of models have been studied in the literature (Liu, Patacchini, and Zenou, 2014). In the first model, called the "local average", individuals conform to the social norm and are therefore influenced by the *average* behaviour of their peers. One can think of the individual's decision on whether or not to engage in crime depending on the social norm across his peers as an example where the local average model would be a good approximation of the underlying determinants of criminal participation. The same rationale can be adapted to reason about how the intensity or type of crime of an individual can be influenced by the criminal *norm* of his peers: for example, if perpetuating crime of a violent nature is a common way to demonstrate loyalty and commitment to the criminal group (Carvalho and Soares, 2016), then each individual would face a penalty attached to deviating from committing this type of crime.

In the second model, individuals are influenced by the *total* or *aggregate* behaviour of their peers, therefore being known as the "local aggregate" model. This model results from cases in which the decision of how much crime an individual should engage in depends on the total crime levels of their peers. This model rationale is compatible with a scenario in which the total crime effort of a criminal's peers is complementary to his own criminal effort. In this case, the individual faces incentives to increase his level of effort as a function of the total level of effort of his peers due to strategic complementarities. For the present analysis, we estimate a local aggregate model. We do so as we believe that the *total* crime that surrounds an individual is more influential than the *average* crime committed by each of its peers.

Liu, Patacchini, and Zenou (2014) show that, in the case of the local aggregate model, the best-reply strategy function of an individual with quadratic utility on effort and strategic complementarities with total peers' effort is a linear function of the total effort of his peers and his ex-ante individual heterogeneity in the return to effort. Formally, let us write the individual's utility function as:

$$u(y_i) = \pi_i y_i - \frac{1}{2} y_i^2 + \alpha \sum_{j=1}^{N} g_{ij} y_j y_i \tag{4.1}$$

where y is criminal effort, g_{ij} are the co-offending adjacency links as described in 4.5.1 between individuals i and j, π is ex-ante heterogeneity in returns to criminal effort and $\alpha \geq 0$ is the social multiplier (peer effect). Then, the best-reply response of individual i is given by:

$$y_i = \pi_i + \alpha \sum_{j=1}^{N} g_{ij} y_j \tag{4.2}$$

Under certain stability assumptions, namely that $0 \leq \alpha < 1/\lambda_r^*$ [25], Liu, Patacchini, and Zenou (2014) show that the local-aggregate network game has a unique interior Nash equilibrium in pure strategies consistent with the best-reply response described in equation (4.2).

Assuming that ex ante individual heterogeneity can be linearly decomposed into individual and peer observable characteristics and an unobserved community effect, the econometric model for the local aggregate model of peer effects can be written as follows:

$$Y = \alpha GY + \beta X + \gamma G^* X + \eta + \varepsilon \tag{4.3}$$

Where Y is the number of offences, G is the adjacency matrix, G^* is the row-normalized adjacency matrix, GY is the total number of offences by peers, X is a set of individual characteristics, and G^*X is a set of average peer characteristics. η are community fixed effects. In this model, α identifies peer effects and γ identifies contextual effects.

The usual identification challenge in peer effects models arises from the reflection problem (Manski, 1993), which does not enable the researcher to disentangle peer effects from contextual effects. Using the terminology of social networks, the reflection problem arises when networks are complete. That is, when all agents are connected to (and influenced by) all other agents in the network and therefore G is not full rank. However, many networks (such as the criminal networks studied in this chapter) are not complete; not every individual is connected to all other individuals in the network. Bramoullé, Djebbari, and Fortin (2009), Lee, Liu, and Lin (2010), Liu and Lee (2010), Liu et al. (2012), and others have shown that in cases when the network is incomplete, G is full rank and invertible and therefore the architecture of social networks can be used to identify endogenous peer effects. The identifying variation relies on the fact that a person is not always friends with all of his friends' friends, i.e. the incomplete architecture of social networks provides an exclusion restriction that enables us to identify peer effects net of contextual effects, thereby solving the reflection problem.

[25] λ_r^* stands for highest eigenvector of the adjacency matrix G_r.

Formally, if the network is incomplete and so G is invertible, we can rewrite the model in equation (4.3) as the following reduced form for the r^{th} community:

$$Y_r = (I_{n_r} - \alpha G_r)^{-1} (\beta X_r + \gamma G_r^* X_r + \eta_r I_{n_r} + \varepsilon_r) \tag{4.4}$$

From equation (4.4), Liu et al. (2012) show that the econometric model can be identified using $G^{*2}X$ (average characteristics of peers of peers) and GL (total number of peers) as instruments for GY conditional on X, GX and η. Specifically, let $Z_r = \{G_r^{*2}X, G_r L\}$, we assume that the characteristics of peers of peers and the total number of peers are as good as random conditional on community fixed effects and other individual and peer covariates, $E[Z_r'\varepsilon_r|\eta_r, X_r, GX_r] = 0$. Conditioning the variation within communities to control for similarity in unobservable individual characteristics and contextual environment effects (correlated effects within communities) is important for the plausibility of the exogenousity of the instruments and shows the relevance of the community detection modelling of Section 4.5.3.

Table 4.4 presents the estimates of the peer effect model described in equation (4.3) using different estimators. Column (1) starts by presenting the OLS estimates consistent with a positive peer effect of 1.6% increase in crime relative to the counterfactual scenario where a criminal does not have any co-offenders. Columns (2) to (5) present estimates for different instrumental variable estimators. We start by estimating the IV model via 2SLS and then consider estimators to deal with the documented bias of 2SLS in the presence of many instruments (Bekker, 1994) and potential efficient losses of under exploring the over-identification of IV models (LIML, GMM, GMM-CUE). The resulting estimates of the peer effects parameter are remarkably consistent across estimators at 0.015.[26] Reassuringly, the F-statistic assessing the strength of the first stage indicates our instruments are strong and relevant,[27] which may explain the little variability when using different estimators, considering that these estimators tend to exhibit larger differences in the presence of many weak instruments (Chao and Swanson 2005; Mikusheva and Sun 2021; Keane and Neal 2024). Equally importantly, the test for overidentifying restrictions suggests that the model is correctly specified and the instruments proposed are not invalid.

The value of the peer effects coefficient then indicates that having one friend only increases an individual's crime by 1.5% compared to a scenario where the individual does not have any peers. Given that in this setting the average number of peers in is 10.66, the implied average social multiplier ($\frac{1}{1-E(GL)\alpha}$) is 1.19. This value is somewhat closer to Dominguez (2021), but smaller than Lindquist and Zenou (2014). In both cases, the networks are smaller than in our current context.

[26] Note that this estimate satisfy the condition for the existence of unique interior Nash equilibrium, $0 \leq \alpha < 1/\lambda_r^*$, as $1/\lambda_r^* = 1/51.04 = 0.0196$ in the communities used in our analysis.

[27] The first stage estimates can be found in Table 4.5

4.5.5 Key Players

The next step in our analysis is to identify so-called "key players" in our setting. According to Ballester, Calvó-Armengol, and Zenou (2006), the key player in a criminal group is "the individual who, when removed, leads to the most significant reduction in the group aggregate crime". The identification of such player is crucial in terms of policy making as theory would indicate that it would be a strategy with large effects, at a relatively small cost (as only one individual would be removed). We identify the most influential individuals in this setting as those with the highest contextual intercentrality measure. We also compare it to other more standard centrality measures in the network literature. We also compare it to the order categorisation and underlying hierarchy stemming from police intelligence and network growing method. These exercises are of relevance to potentially detect important players that are "missed" by more readily available metrics.

The seminal paper by Ballester, Calvó-Armengol, and Zenou (2006) first identifies the key player as the individual in the network with the largest intercentrality measure. This is a measure the authors built specifically for the purpose of identifying key players in networks, and has been widely used in the empirical networks literature since[28]. However, in this paper, we follow a more refined version intercentrality proposed in Ballester and Zenou (2014).

In Ballester, Calvó-Armengol, and Zenou (2006), intercentrality measures an agent's overall importance for network activity by calculating the drop in aggregate equilibrium effort when that agent is removed from the network. In this specific case, it measures how much total gang crime across the whole network would drop if a given individual were removed. It is derived from the Bonacich centrality framework, so it reflects both direct productivity and the indirect spillover effects an agent generates through their position. Ballester and Zenou (2014) retain the same strategic interaction framework but introduce contextual intercentrality. The contextual measure differs from its global counterpart in that it incorporates two distinct channels: the direct contextual effect, reflecting the individual's own contribution to crime within the overall context, and the indirect contextual effect, reflecting their influence on other gang members whose activities are context-relevant.

It is worth noting that identifying key players via the contextual intercentrality measure requires two assumptions. The first assumption needed is that the gang is

[28]Ballester, Calvó-Armengol, and Zenou (2006) build the intercentrality measure as follows:

$$c_i(g,a) = \frac{b_i(g,a)^2}{m_{ii}(g,a)} \tag{4.5}$$

where $c_i(g,\phi)$, is the centrality of individual i, considering network g and a scalar ϕ. The network G is the one considered above, and we will consider our peer effect estimate stemming from the previous section as the scalar ϕ. $b_i(g,\phi)$ is the Bonacich centrality of node i (which counts the number of paths in g that stem from i), $b(g,\phi) = \sum_{j=1}^{n} m_{ij}(g,\phi)$, $M(g,\phi) = [I - \phi G]^1$, and $m_{ii}(g,\phi)$ is the ii element of matrix M.

fixed, meaning that the network g does not vary after the removal of an individual, and no new links are created. The second assumption requires that the criminal ability of an individual α does not depend on the gang but depends on their characteristics and the characteristics of its peers [29].

For individual i, peer effects ϕ, ability α, and gang g, Ballester and Zenou (2014) define the contextual intercentrality measure as follows:

$$\delta_i(g, \phi, \alpha) = \frac{b_{\alpha_{\langle i \rangle},i}(g, \phi) \cdot \sum_{j=1}^{n} m_{ji}}{m_{ii}} + B_\alpha(g, \phi) - B_{\alpha_{\langle i \rangle}}(g, \phi) \tag{4.6}$$

where $\alpha_{\langle i \rangle} = (X_i, \alpha^{[-i]})'$ describes the context where individual i has not yet been removed, so they have their attributes X, and the vector α is computed from the network when individual i is removed from the network, namely $\alpha^{[-i]}$. $B_\alpha(g, \phi)$ and $B_{\alpha_{\langle i \rangle}}(g, \phi)$ measure total Bonacich network centrality under contextual effects α and $\alpha_{\langle i \rangle}$ respectively.

$\delta_i(g, \phi, \alpha)$ considers two effects. The first one, corresponding to the first term in equation (4.6), is a network effect derived from the centrality measure by Ballester, Calvó-Armengol, and Zenou (2006) with contextual effects α. It captures the direct effect on crime from removing individual i and the indirect effect on others' criminal activity from the removal of that individual from the network while keeping the vector $\alpha_{\langle i \rangle}$ unchanged. The second and novel effect with respect to Ballester, Calvó-Armengol, and Zenou (2006) is the contextual one, which is captured by the last two terms in equation (4.6). It stems from the change in context from α to $\alpha_{\langle i \rangle}$.

Table 4.6 presents average characteristics for all gang members and those identified in the top 10% of the contextual intercentrality distribution. From the first panel, it is possible to see that the most central individuals are very similar to the average across most observable characteristics. However, they seem to be, on average, a larger proportion of men and a larger share is alive. In terms of their criminal characteristics, some differences arise. Central individuals according to the contextual intercentrality measure have committed a larger number of offences. Moreover, they are more likely to be classified as order 0 or 1 in the intelligence labelling. This last feature is not trivial. First, because the contextual intercentrality measure does not use the information on orders at all, their empirical correlation was not guaranteed. Second, because it confirms to a certain extent that the orders in the police intelligence data are identifying many of the central individuals, according to network theory. Nonetheless, it also indicates that some individuals that are classified as order 2 are central individuals in the network. Therefore, it seems that key players are mostly in line with the hierarchical classification from the police intelligence, but there may be some "missed" individuals that we are able to detect with network analysis. Finally, the third panel of Table 4.6 indicates that the most central individuals

[29] Operationally $\alpha_i = \beta_0 + X_i\beta_1 + \bar{X}_i\beta_2 + u_i$, where X is a vector of individual observable exogenous characteristics and $\bar{X}$ is a matrix with the average characteristics of individual i's peers.

are not concentrated in one gang only. Rather, they seem to follow the distribution of the general gang member population.

Figure 4.4 displays the network subgraph of one distinct cluster to demonstrate how intercentrality can reveal more heterogeneity within the network than simple network hierarchy.[30] The measures are in general agreement with each other; The average intercentrality rank of zero order nodes is greater than the average rank of first order nodes, which is in turn greater than that of second order nodes. Zero order gang members in this subgraph are also all concentrated within the top two deciles of the intercentrality measure. However, the measure also displays considerable variation within the nodes of each order. There also is considerable overlap; there are many second order nodes that have higher intercentrality than first order nodes.

To consider how the intercentrality measure may vary in its targeting compared to police intelligence, we directly compare the 662 zero order gang members to the same number of individuals with the highest intercentrality values. Only 65 individuals are found in both lists, corresponding to an intersection rate of 9.8%. Furthermore, figure 4.5 shows that the top key players are also concentrated within fewer clusters compared to order zero gang members. The key players are concentrated in 44 clusters, with the 5 most frequent clusters containing 43% of these key players. In comparison, order zero gang members are concentrated within 179 clusters, with the 5 most frequent clusters containing only 10.5% of these order zero gang members. Table 4.8 compares the distribution of these same two groups across assigned gang labels. Compared to the order zero nodes, the key players are more likely to belong to the most common gang (gang 3). The key players are also less likely to be assigned a double label (e.g. gang 1/2) or to be labelled as "Mixed Uncertain".

Lastly, table 4.7 presents a rank-rank correlation of the intercentrality measure with other standard network centrality measures. The results indicate a positive correlation among these measures that varies in strength. While the contextual intercentrality measure seems to have a high correlation with degree centrality, which measures the number of direct peers, its correlation with other centrality measures is much lower. This could indicate that relying only on readily available network centrality measures provides a partial picture to detect and identify the key players in the network.

4.6 Conclusion

Organised crime imposes large costs on society, as it distorts labour markets, it fuels violence, it captures political institutions, it exploits illicit trades, and it has the potential to ultimately hinder economic development. This chapter contributes to the existing literature that has documented these effects by investigating the network

[30]The chosen cluster is the largest cluster identified when using Markov stability analysis at the optimal time value, as discussed in section .

structure and microinteractions of gangs in Rio Grande do Sul, Brazil's southernmost state.

Largely due to the scarcity of available data, the literature lacks detailed insight into who holds power within gangs, how networks are structured, and how that internal structure influences the gang's functioning and its vulnerabilities. Using a rich integration of intelligence data, prison stay data, occurrence data, and detailed data on individual characteristics, we apply a network community-detection model (Markov Stability) to identify communities (network subgraphs) within Brazilian gangs operating in our setting. Our analysis of these communities using various network statistics to discern different levels of centrality indicates the potential hierarchy and roles of individuals within these criminal structures. Comparing these network-derived insights with official designations provided by police intelligence, such as known leadership roles, suggests that the conventional wisdom about these gangs' hierarchy is incomplete. Only 9.8% of the top bosses identified by the intelligence data are identified as key players. Although these individuals hold relatively low rank within the network, their extensive connections across subgroups could position them to be key liaisons. Removing them from the criminal organisation could do particular harm to these organisations.

The contribution of our study is twofold: firstly, it incorporates methods from network analysis to investigate hierarchy formation; secondly, it uncovers actionable intelligence on how law enforcement might remove key individuals from existing gangs more effectively by targeting those who hold the network together. By shedding light on the internal workings and network structures of Brazilian gangs, this study advances our understanding of these illicit organizations and how they can be disrupted. In doing so, it builds on on the existing literature that has documented why they are so harmful to modern economies and societies.

4.7 Tables

Table 4.1: Summary statistics of network

	Order 0	Order 1	Order 2	Total
Share female	0.06	0.21	0.19	0.19
	(0.24)	(0.40)	(0.39)	(0.39)
Share in prison at end of period	0.36	0.27	0.22	0.24
	(0.48)	(0.45)	(0.41)	(0.42)
Share ever in prison	0.96	0.76	0.67	0.70
	(0.20)	(0.43)	(0.47)	(0.46)
Average age	44.15	34.57	33.29	33.86
	(8.82)	(9.51)	(9.84)	(9.87)
Average number of gallery stays	20.34	9.09	7.95	8.54
	(15.93)	(12.87)	(13.23)	(13.34)
Average number of occurrences	13.45	10.52	9.90	10.15
	(14.17)	(8.86)	(9.00)	(9.14)
Average number of gang-related occurrences	7.04	5.18	4.54	4.77
	(9.07)	(4.85)	(4.45)	(4.74)
Average number of drugs and trafficking occurrences	2.26	1.94	1.51	1.64
	(4.15)	(2.23)	(2.01)	(2.15)
Average number of murder occurrences	2.31	0.98	0.57	0.72
	(5.37)	(2.04)	(1.21)	(1.70)
Count	662	7,455	20,042	28,159

Standard errors are shown in parentheses. Differences between groups for all variables excluding share female are statistically significant at the 1% level.

Table 4.2: Share of Individuals Assigned to Different Cluster Types Using Stable Markov Partitions

Markov Time Value	Mixed Gang Label	Mixed Uncertain Label	No Cluster Label Assigned	Single Gang Label
3	0.11	0.23	0.01	0.64
65	0.14	0.14	0.00	0.71
112	0.17	0.12	0.00	0.71
154	0.15	0.10	0.00	0.74
203	0.22	0.11	0.00	0.67
255	0.22	0.27	0.00	0.51
296	0.97	0.01	0.00	0.02

Table reports the share of individuals assigned to each type of cluster for each of the six optimal partitions, as determined using the Markov Stability algorithm. The total number of individuals within the network in each row is 27,760 - this represents the largest fully connected sub-graph in our network.

Table 4.3: Share of Cluster Types Using Stable Markov Partitions

Markov Time Value	Total Clusters	Dual Gang Label	Mixed Uncertain Label	No Cluster Label Assigned	Single Gang Label
3	1170	0.07	0.37	0.07	0.49
65	552	0.11	0.25	0.02	0.61
112	243	0.16	0.19	0.00	0.65
154	125	0.14	0.16	0.01	0.70
203	52	0.17	0.17	0.00	0.65
255	21	0.14	0.14	0.00	0.71
296	8	0.12	0.50	0.00	0.38

Table reports the share of clusters assigned each type of label for each of the six optimal partitions, as determined using the Markov Stability algorithm. The total number of individuals within the network in each row is 27,760 - this represents the largest fully connected sub-graph in our network.

Table 4.4: Peer Effect Estimates

	(1) OLS	(2) 2SLS	(3) LIML	(4) GMM	(5) GMM-CUE
GY	0.016***	0.015***	0.015***	0.015***	0.015***
	(0.0010)	(0.0009)	(0.0010)	(0.0007)	(0.0007)
Observations	27,760	27,760	27,760	27,760	27,760
First stage F		86.7	86.7	86.7	86.7
J Over-Identification (p-value)		0.276	0.276	0.276	0.287
Controls	X, $G^{*}X$, η				

Notes: This table reports peer effects estimates following Eq.(4.3). Column 1 to 5 presents results for different estimators: OLS, 2SLS, LIML, GMM and GMM-CUE. Individual and average peer characteristics are included: sex, cohort of birth, race, schooling level, single status, in prison as of end date and region of birth. All models include community fixed effects are included as controls. The observational unit is the individual. Standard errors clustered at the community level are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.5: First-Stage Estimates

GL	
Number of Peers	11.809***
	(0.299)
$G^{*2}X$	
Birth Cohort	0.123
	(0.106)
White	-9.910
	(6.624)
Basic Education	21.492**
	(7.571)
> Basic Education	-9.281
	(9.053)
Single	32.909***
	(9.909)
In Prison at End Period	46.863***
	(5.878)
Female	-18.195**
	(8.970)
Observations	27,760
Controls	X, $G^{*}X$, η

Notes: The coefficients of peers-of-peers' region-of-birth shares are not displayed but are included in the first-stage estimation. Standard errors clustered at the community level are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.6: Average characteristics, full sample and key players

	Full Sample	Top 10 intercentrality
Female	19.0	15.4
YOB	1990	1989
White	73.2	73.2
Single	90.4	89.8
Died	5.0	2.73
No Education	19.4	18.4
Basic Education	60.9	62.0
> Basic Education	19.6	19.6
Number of offences	10.22	19.8
In prison	23.7	40.4
Order 0	1.43	5.37
Order 1	26.55	56.92
Order 2	72.02	37.72
Gang 1	13.69	14.01
Gang 1/2	0.18	0.18
Gang 1/3	13.91	14.70
Gang 1/4	0.27	0.68
Gang 3	50.35	49.86
Gang 3/5	0.16	0.04
Gang 3/6	2.74	2.49
Gang 4	2.85	2.59
Gang 6	3.98	3.31
Mixed Uncertain	11.84	11.85

Notes: This table presents mean characteristics of all gang members in our sample, and mean characteristics of individuals identified in the top 10% of the intercentrality measure of Ballester and Zenou (2014).

Table 4.7: Rank-rank correlation among centrality measures

	Degree	Betweeness	Closeness	Eigencentrality
Intercentrality	0.642	0.494	0.395	0.293

Notes: This table presents rank-rank correlations of the intercentrality measure of Ballester and Zenou (2014) with standard centrality measures: degree, betweenness, closeness, and eigencentrality.

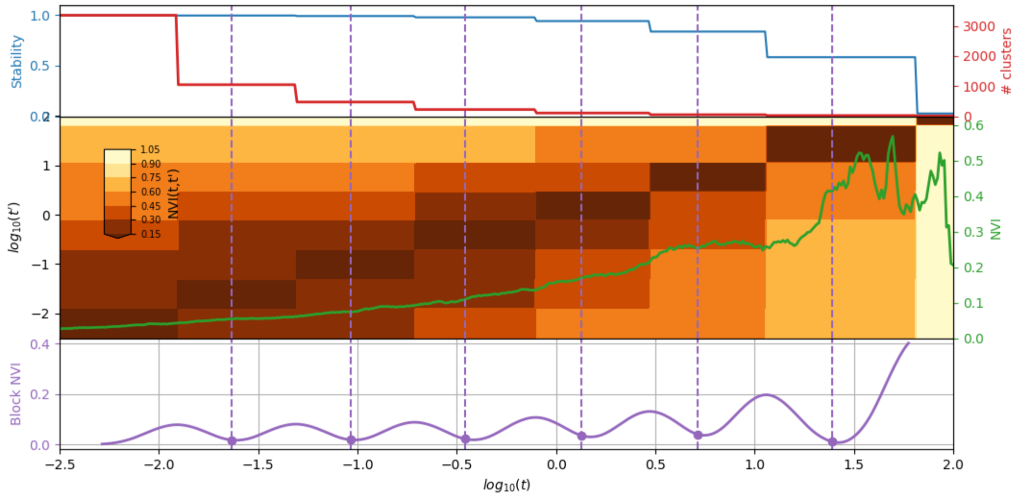
Table 4.8: Comparing key players to order zero individuals: Gang label distribution

	Order Zero	Key Players
Gang 1	0.13	0.20
Gang 1/2	0.01	0.00
Gang 1/3	0.16	0.04
Gang 1/4	0.01	0.01
Gang 3	0.48	0.60
Gang 3/5	0.02	0.00
Gang 4	0.02	0.05
Gang 5	0.02	0.02
Mixed Uncertain	0.16	0.09
Total Observations	396	396

Notes: This table demonstrates the share of individuals assigned to each gang label (using the method described in section 4.5.3) within two groups. The first group is the 396 order zero individuals that were assigned gang labels. The second group is a sample of key players of the same size, in other words the 396 labelled individuals with the highest calculated intercentrality values.

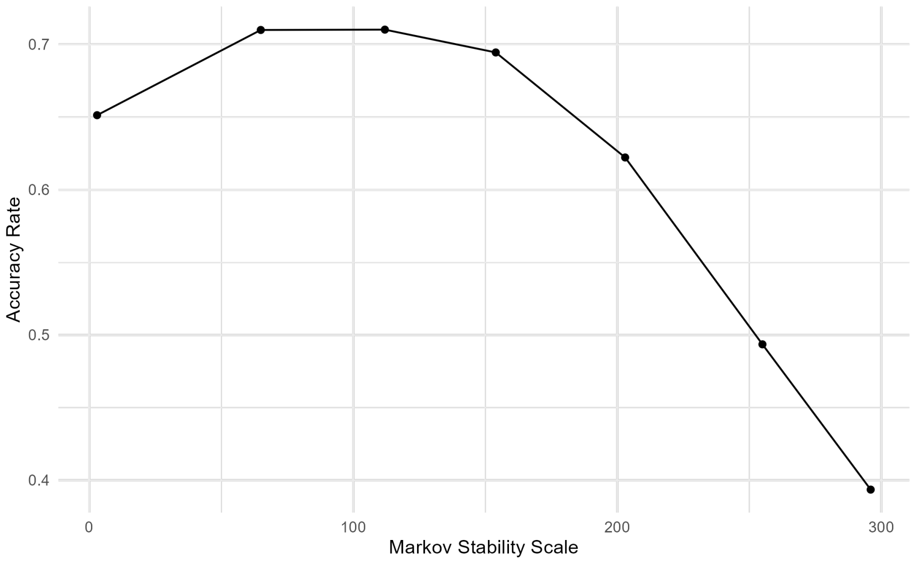
4.8 Figures

Figure 4.1: Stable Partitions of the Markov Stability Algorithm



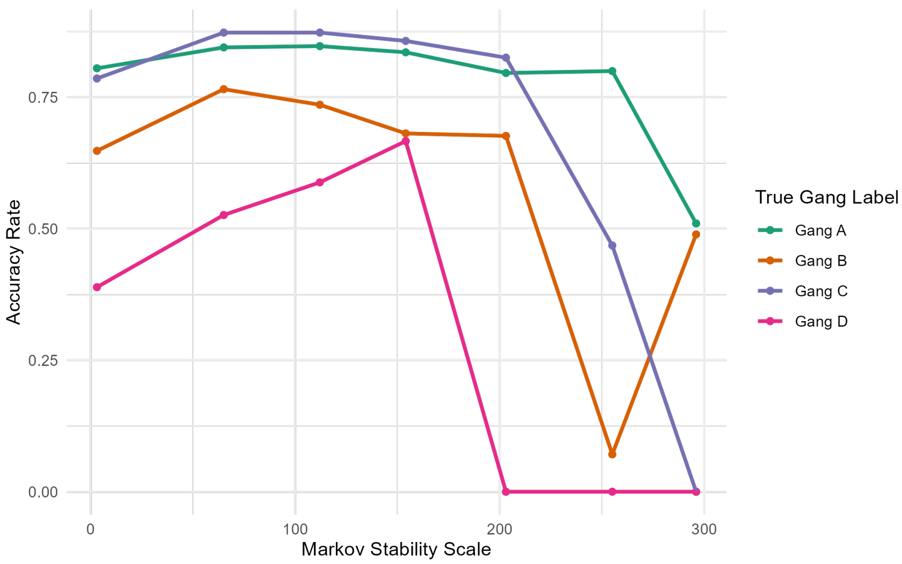
Denotes the community outcomes for different time scales. A longer time allows for a longer random walk within the network, and therefore segments the network into fewer, larger communities. The Normalized Variation of Information (NVI) is a metric to determine the (dis)similarity between partitions. The purple vertical lines pinpoint the six time scales that result in a stable partition set.

Figure 4.2: Community Label Accuracy Rate by Markov Stability Scale



Each data point captures the share of individuals in the validation set whose community label fully or partially matches their true gang label. The true gang label refers to the gang label of the individual within the validation set.

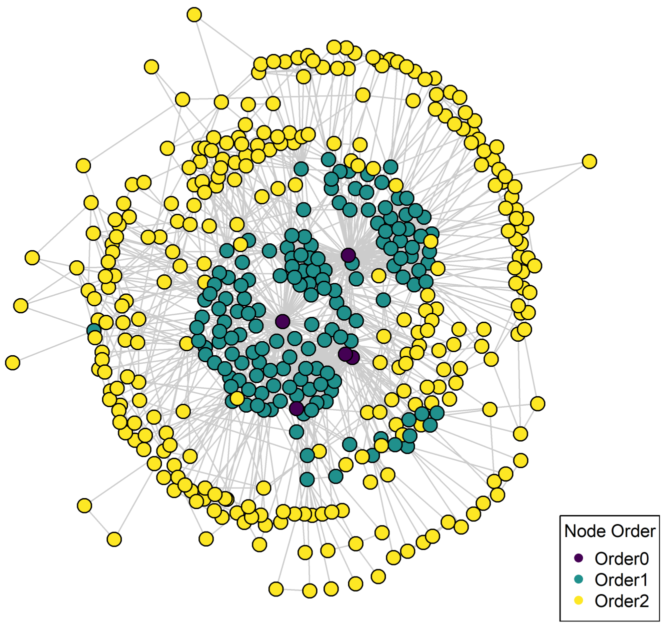
Figure 4.3: Community Label Accuracy Rate by Markov Stability Scale - By True Gang Affiliation



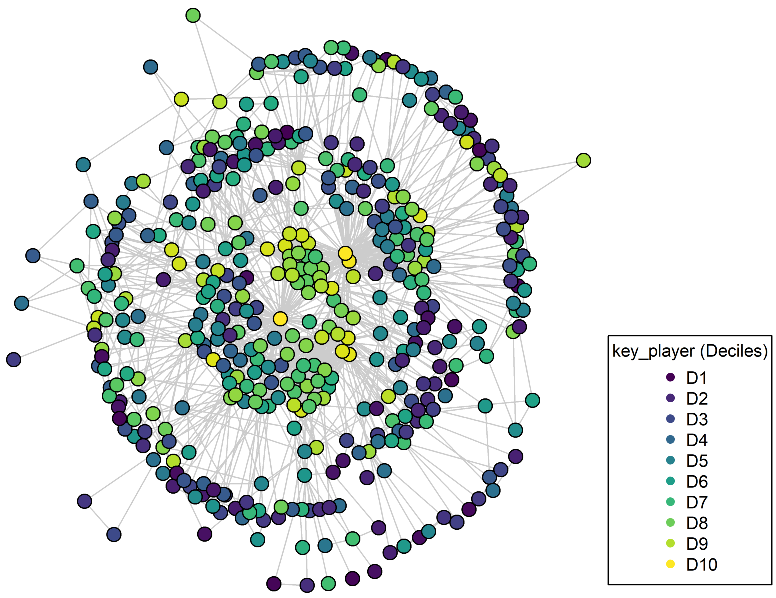
Each data point captures the share of individuals in the validation set whose community label fully or partially matches their true gang label. The true gang label refers to the gang label of the individual within the validation set.

Figure 4.4: Co-offending Network Subgraphs of Largest Detected Markov Stability Cluster

4.4a: By Node Order of Connection

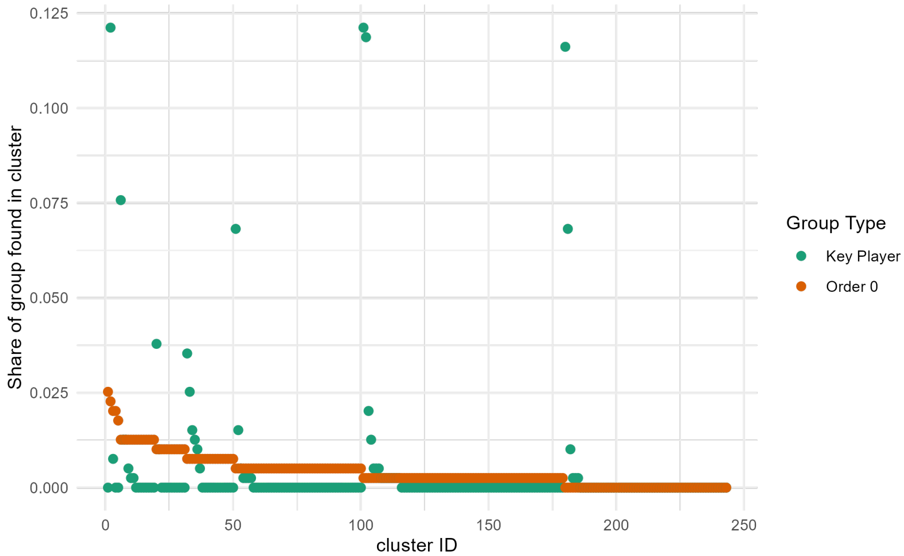


4.4b: By Intercentrality



Co-offending Network Subgraphs of the Largest Detected Markov Stability Cluster. The network displays subgraphs of the set of individuals in the largest cluster at the chosen Markov stability scale. Each vertex connects a pair of co-offenders. In (a) the node colour denotes the inferred hierarchy of the individual, and in (b) the node colour denotes the decile of the individual's calculated intercentrality.

Figure 4.5: Distribution of Key Players and Order Zero Individuals across Markov Stability Clusters



The figure demonstrates the share of individuals assigned to each Markov Stability Cluster within two groups. The first group is the 396 order zero individuals that were included in the cluster detection exercise. The second group is a sample of key players of the same size, in other words the 396 individuals in the network with the highest calculated intercentrality values. For ease of comparison, the clusters are ordered by the share of order zeros included in the cluster.

Bibliography

Abebe, Girum, A Stefano Caria, and Esteban Ortiz-Ospina. 2021. "The Selection of Talent: Experimental and Structural Evidence from Ethiopia." *American Economic Review*, 111(6): 1757–1806.

Adams, Lorna, Sarah Coburn-Crane, Alfie Sanders-Earley, Rachel Keeble, Harry Harris, James Taylor, and Becky Taylor. 2023. "Working lives of teachers and leaders – wave 1." Department for Education.

Akee, Randall and Maggie R Jones. 2019. "Immigrants' Earnings Growth and Return Migration from the U.S.: Examining their Determinants using Linked Survey and Administrative Data." *NBER Working Paper Series.*

Alesina, Alberto, Salvatore Piccolo, and Paolo Pinotti. 2019. "Organized crime, violence, and politics." *The Review of Economic Studies*, 86(2): 457–499.

Allegretto, Sylvia. 2022. "The Teacher Pay Penalty Has Hit a New High: Trends in Teacher Wages and Compensation through 2021." Economic Policy Institute.

Allen, Rebecca, Chris Belfield, Ellen Greaves, Caroline Sharp, and Matt Walker. 2014. "The Costs and Benefits of Different Initial Teacher Training Routes." Institute for Fiscal Studies.

Amior, Michael and Jan Stuhler. 2024. "Immigration, Monopsony and the Distribution of Firm Pay." *CEP Discussion Paper, CEPDP1971.*

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

Arnaudon, Alexis, Dominik J Schindler, Robert L Peach, Adam Gosztolai, Maxwell Hodges, Michael T Schaub, and Mauricio Barahona. 2024. "Algorithm 1044: PyGenStability, a Multiscale Community Detection Framework with Generalized Markov Stability." 50(2).

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S. Lee. 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.

Baccara, Mariagiovanna and Heski Bar-Isaac. 2008. "How to Organize Crime." *Review of Economic Studies*, 75(4): 1039–1067.

Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean Marc Robin. 2014. "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics." *American Economic Review*, 104(6): 1551–96.

Ballester, Coralio and Yves Zenou. 2014. "Key Player Policies When Contextual Effects Matter." *Journal of Mathematical Sociology*, 38(4): 233–248.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou. 2006. "Who's who in networks. Wanted: The key player." *Econometrica*, 74(5).

Barone, Guglielmo and Gaia Narciso. 2015. "Organized crime and business subsidies: Where does the money go?" *Journal of Urban Economics*, 86: 98–110.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy*, 116(6): 1150–1196.

Becker, Gary S. 1968. "Crime and Punishment: an Economic Approach." In *The Economic Dimensions of Crime*. 13–68. Palgrave Macmillan UK.

Becker, Gary S, Kevin M Murphy, and Michael Grossman. 2006. "The market for illegal goods: the case of drugs." *Journal of Political Economy*, 114(1): 38–60.

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

Bekker, Paul A. 1994. "Alternative Approximations to the Distributions of Instrumental Variable Estimators." *Econometrica*, 62(3): 657–681.

Bell, Brian, Nicholas Bloom, and Jack Blundell. 2022. "Income dynamics in the United Kingdom and the impact of the Covid-19 recession." *Quantitative Economics*, 13(4).

Biasi, Barbara. 2024. "Salaries, Pensions, and the Retention of Public-Sector Employees: Evidence from Wisconsin Teachers."

Biasi, Barbara, Chao Fu, and John Stromme. 2021. "Equilibrium in the Market for Public School Teachers: District Wage Strategies and Teacher Comparative Advantage."

Bisca, Paul M, Vu Chau, Paolo Dudine, Raphael A Espinoza, Jean-Marc Fournier, Pierre Gu'erin, Niels-Jakob H Hansen, and Jorge Salas. 2024. "Violent Crime and Insecurity in Latin America and the Caribbean: A Macroeconomic Perspective." International Monetary Fund 2024/009, Washington, D.C.

Blattman, Christopher, Gustavo Duncan, Benjamin Lessing, and Santiago Tobón. 2024. "Gang rule: Understanding and countering criminal governance." *Review of Economic Studies*.

Blau, Francine D and Lawrence M Kahn. 2017. "The gender wage gap: Extent, trends, & explanations." *Journal of Economic Literature*, 55(3).

Boeri, Tito, Giulia Giupponi, Alan B Krueger, and Stephen Machin. 2020. "Solo self-employment and alternative work arrangements: A cross-country perspective on the changing composition of jobs."

Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *Journal of Labor Economics*, 3(4).

Borjas, George J. 2015. "The slowdown in the economic assimilation of immigrants: Aging and Cohort effects revisited again." *Journal of Human Capital*, 9(4).

Borjas, G J. 1995. "Assimilation and changes in cohort quality revisited: what happened to immigrant earnings in the 1980s?" *Journal of labor economics*, 13(2).

Bowyer, Georgina and Morag Henderson. 2020. "Race Inequality in the Workforce." Centre for Longitudinal Studies.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009. "Identification of peer effects through social networks." *Journal of econometrics*, 150(1): 41–55.

Britton, Jack and Carol Propper. 2016. "Teacher pay and school productivity: Exploiting wage regulation." *Journal of Public Economics*, 133: 75–89.

Britton, Jack, Ian Walker, and Ben Waltmann. 2022. "How much does it pay to get good grades at university?" Institute for Fiscal Studies, London.

Brynin, Malcolm and Ayse Güveli. 2012. "Understanding the ethnic pay gap in Britain." *Work, Employment and Society*, 26(4): 574–587.

Bueno, Carycruz and Tim R. Sass. 2018. "The Effects of Differential Pay on Teacher Recruitment and Retention."

Burdett, Ken and Melvyn Coles. 2003. "Equilibrium wage-tenure contracts." *Econometrica*, 71(5): 1377–1404.

Calderoni, Francesco, Gian Maria Campedelli, Tommaso Comunale, Martina E Marchesi, and Ernesto U Savona. 2020. "Recruitment into organised criminal groups: A systematic review." *Trends and Issues in Crime and Criminal Justice*, , (583): 1–28.

Calvó-Armengol, Antoni and Matthew O Jackson. 2004. "The effects of social networks on employment and inequality." *American Economic Review*, 94(3).

Campana, Paolo and Federico Varese. 2018. "Organized crime in the United Kingdom: Illegal governance of markets and communities." *The British Journal of Criminology*, 58(6): 1381–1400.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *The Quarterly Journal of Economics*, 128(3).

Carvalho, Leandro and Rodrigo R Soares. 2016. "Living on the Edge: Youth Entry, Career and Exit in Drug-Selling Gangs." *Journal of Economic Behavior & Organization*, 121: 77–98.

Castillo, Juan Camilo, Daniel Mejía, and Pascual Restrepo. 2020. "Scarcity without leviathan: The violent effects of cocaine supply shortages in the mexican drug war." *Review of Economics and Statistics*, 102(2): 269–286.

Chao, John C and Norman R Swanson. 2005. "Consistent Estimation with a Large Number of Weak Instruments." *Econometrica*, 73(5): 1673–1692.

Chevalier, Arnaud, Peter Dolton, and Steven McIntosh. 2007. "Recruiting and Retaining Teachers in the UK: An Analysis of Graduate Occupation Choice from the 1960s to the 1990s." *Economica*, 74(293): 69–96.

Chiswick, Barry R. 1978. "The Effect of Americanization on the Earnings of Foreign-born Men." *Journal of Political Economy*, 86(5).

Christian, Alvin, Matthew Ronfeldt, and Basit Zafar. 2024. "College Students and Career Aspirations: Nudging Student Interest in Teaching."

Cipriani, Marcelli, Nicolas Santiago Lien, and Alana Barros Santos. 2023. "Between Factions and Clans: The Structuring of the Drug Market in Porto Alegre and Córdoba from the Dynamics of Their Criminal Collectivities." *Novos estudos CEBRAP*, 42: 81–99.

Clark, Ken and Steve Nolan. 2021. "The Changing Distribution of the Male Ethnic Wage Gap in Great Britain." IZA – Institute of Labor Economics.

Clark, Ken and William Shankley. 2020. "Ethnic minorities in the labour market in Britain." *Ethnicity and Race in the UK*, 127–148.

Clark, Kenneth and Stephen Drinkwater. 2000. "Pushed out or pulled in? Self-employment among ethnic minorities in England and Wales." *Labour Economics*, 7(5).

Clark, Kenneth and Stephen Drinkwater. 2002. "Enclaves, neighbourhood effects and employment outcomes: Ethnic minorities in England and Wales." *Journal of Population Economics*, 15(1).

Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor. 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, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. "Teacher-student matching and the assessment of teacher effectiveness." *Journal of Human Resources*, 41(4).

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

Coffman, Lucas C., John J. Conlon, Clayton R. Featherstone, Judd B. Kessler, and Jessica Mixon. 2023. "Liquidity for teachers: Evidence from Teach For America and LinkedIn." *Economics of Education Review*, 97.

Commission on Race and Ethnic Disparities. 2021. "The report of the Commission on Race and Ethnic Disparities."

Crawford, Rowena and Richard Disney. 2018. "Wage Regulation and the Quality of Police Applicants." *Economica*, 85(340): 701–734.

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

Dal Bó, Ernesto, Pedro Dal Bó, and Rafael Di Tella. 2006. ""Plata o Plomo?": bribe and punishment in a theory of political influence." *American Political science review*, 100(1): 41–53.

Daniele, Gianmarco and Benny Geys. 2015. "Organised crime, institutions and political quality: Empirical evidence from italian municipalities." *The Economic Journal*, 125(586): F233–F255.

De Falco, Adriano, Benjamin Hattemer, and Sofía Sierra Vásquez. 2024. "Recruiting Better Teachers? Evidence from a Higher Education Reform in Chile."

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

Dell, Melissa. 2015. "Trafficking Networks and the Mexican Drug War." *American Economic Review*, 105(6): 1738–1779.

Department for Education. 2011. "Training our next generation of outstanding teachers Implementation plan."

Department for Education. 2014. "Record Number of Teachers as 9,000 Join Schools in England."

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

De Simone, Martin. 2014. "Crimen organizado en Argentina. Una mirada con perspectiva democrática y desde los derechos humanos." *Buenos Aires, Asociación Civil por la Igualdad y la Justicia/La otra trama.*

Detotto, Claudio and Edoardo Otranto. 2010. "Does crime affect economic growth?" *Kyklos*, 63(3): 330–345.

Dias, Camila Nunes and Fernando Salla. 2013. "Organized Crime in Brazilian Prisons: The Example of the PCC." *International Journal of Criminology and Sociology*, 2: 397–408.

Dimico, Arcangelo, Alessia Isopi, and Ola Olsson. 2017. "Origins of the Sicilian mafia: The market for lemons." *The Journal of Economic History*, 77(4): 1083–1115.

Disney, Richard, Carl Emmerson, and Gemma Tetlow. 2010. "The Value of Teachers' Pensions in England and Wales." *Fiscal Studies*, 31(1): 121–150.

Dolton, Peter and Wilbert Van der Klaauw. 1999. "The Turnover of Teachers: A Competing Risks Explanation." *Review of Economics and Statistics*, 81(3): 543–550.

Dolton, Peter, Lea Samek, and Powen She. 2019. "Total Reward and Pensions in the UK by Occupation in the Public and Private Sectors." Office of Manpower Economics.

Dolton, P. J. 1990. "The economics of UK teacher supply: the graduate's decision." *Economic Journal*, 100(400): 91–104.

Dominguez, Magdalena. 2021. "Sweeping up gangs: The effects of tough-on-crime policies from a network approach." *IEB Working Paper 2021/03*.

Dustmann, Christian and Joseph Simon Görlach. 2015. "Selective out-migration and the estimation of immigrants' earnings profiles." Vol. 1.

Elsby, Michael W L, Donggyun Shin, and Gary Solon. 2016. "Wage adjustment in the great recession and other downturns: Evidence from the united states and Great Britain." *Journal of Labor Economics*, 34(S1).

Engstrom, David M. 1997. "The Economic Determinants of Ethnic Segregation in Post-War Britain."

Falch, Torberg. 2011. "Teacher Mobility Responses to Wage Changes: Evidence from a Quasi-natural Experiment." *American Economic Review*, 101(3): 460–65.

Falch, Torberg. 2017. "Wages and recruitment: Evidence from external wage changes." *Industrial and Labor Relations Review*, 70(2): 483–518.

Feltran, Gabriel. 2018. *Irmãos: uma história do PCC.* Editora Companhia das Letras.

Feng, Li and Tim R. Sass. 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, Jeremy D. and Charles M. Achilles. 1999. "Tennessee's Class Size Study: Findings, Implications, Misconceptions." *Educational evaluation and policy analysis*, 21(2): 97–109.

Forth, John, Van Phan, Felix Ritchie, Damian Whittard, Lucy Stokes, Alex Bryson, and Carl Singleton. 2022. "ASHE – 2011 Census Data Linkage: User Guide Drop 1 of the ASHE – 2011 Census Dataset."

Fórum Brasileiro de Segurança Pública. 2023. "Anuário Brasileiro de Segurança Pública 2023."

Fullard, Josh. 2021*a*. "Bad economy, good teachers? The countercyclicality of enrolment Into Initial Teacher Training Programmes in the UK." Institute for Social and Economic Research (ISER).

Fullard, Joshua. 2021*b*. "Relative Wages and Pupil Performance, evidence from TIMSS." Institute for Social and Economic Research (ISER).

Fullard, Joshua and James Zuccollo. 2021. "Local pay and teacher retention in England." Education Policy Institute.

Gaddis, S Michael. 2015. "Discrimination in the credential society: An audit study of race and college selectivity in the labor market." *Social Forces*, 93(4).

Gambetta, Diego. 1993. *The Sicilian Mafia: The Business of Private Protection.* Cambridge, MA:Harvard University Press.

Garzón, Juan Carlos. 2012. "La rebelión de las redes criminales: El crimen organizado en América Latina y las fuerzas que lo modifican." *Woodrow Wilson International Center for Scholars. Recuperado el*, 3.

Giupponi, Giulia and Xiaowei Xu. 2020. "What does the rise of self-employment tell us about the UK labour market?" *Institute of Fiscal Studies.*

Goldhaber, Dan, John M. Krieg, Stephanie Liddle, and Roddy Theobald. 2024. "Out of the Gate, But Not Necessarily Teaching: A Descriptive Portrait of Early Career Earnings for Those Who Are Credentialed to Teach." *Education Finance and Policy*, 19(1): 81–105.

Green, Francis. 2021. "British teachers' declining job quality: evidence from the Skills and Employment Survey." *Oxford Review of Education*, 47(3): 386–403.

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

Hanushek, Eric A., Marc Piopiunik, and Simon Wiederhold. 2019. "The Value of Smarter Teachers." *Journal of Human Resources*, 54(4): 857–899.

Heath, Anthony F and Valentina Di Stasio. 2019. "Racial discrimination in Britain, 1969–2017: a meta-analysis of field experiments on racial discrimination in the British labour market." *British Journal of Sociology*, 70(5).

Hendricks, Matthew D. 2014. "Does it pay to pay teachers more? Evidence from Texas." *Journal of Public Economics*, 109: 50–63.

Inter-American Development Bank. 2017. *The Costs of Crime and Violence: New Evidence and Insights in Latin America and the Caribbean.* Washington, DC.

Ioannides, Yannis M. and Linda Datcher Loury. 2004. "Job Information Networks, Neighborhood Effects, and Inequality." *Journal of Economic Literature*, 42(4): 1056–1093.

Keane, Michael P. and Kenneth I. Wolpin. 1997. "The career decisions of young men." *Journal of Political Economy*, 105(3): 473–522.

Keane, Michael P and Timothy Neal. 2024. "A Practical Guide to Weak Instruments." *Annual Review of Economics*, 16: 185–212.

Kleinberg, Jon. 2002. "An Impossibility Theorem for Clustering." *Advances in Neural Information Processing Systems*, 15.

Kugler, Maurice, Thierry Verdier, and Yves Zenou. 2005. "Organized Crime, Corruption and Punishment." *Journal of Public Economics*, 89(9-10): 1639–1663.

Leaver, Clare, Owen Ozier, Pieter Serneels, and Andrew Zeitlin. 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.

Lee, Lung-fei, Xiaodong Liu, and Xu Lin. 2010. "Specification and estimation of social interaction models with network structures." *The Econometrics Journal*, 13(2): 145–176.

Lessing, Benjamin. 2017. "Counterproductive punishment: How prison gangs undermine state authority." *Rationality and Society*, 29(3): 257–297.

Levitt, Steven D and Sudhir A Venkatesh. 2000. "An Economic Analysis of a Drug-Selling Gang's Finances." *Quarterly Journal of Economics*, 115(3): 755–789.

Lindquist, Matthew and Yves Zenou. 2014. "Key Players in Co-Offending Networks." IZA – Institute of Labor Economics 8012.

Liu, Xian. 2012. "Appendix A: The Delta Method." In *Survival analysis: models and applications*. Chapter Appendix A, 405–406. Wiley.

Liu, Xiaodong and Lung-fei Lee. 2010. "GMM estimation of social interaction models with centrality." *Journal of Econometrics*, 159(1): 99–115.

Liu, Xiaodong, Eleonora Patacchini, and Yves Zenou. 2014. "Endogenous peer effects: Local aggregate or local average?" *Journal of Economic Behavior and Organization*, 103.

Liu, Xiaodong, Eleonora Patacchini, Yves Zenou, and Lung-Fei Lee. 2012. "Criminal networks: Who is the key player?"

Lubotsky, Darren. 2007. "Chutes or ladders? A longitudinal analysis of immigrant earnings." *Journal of Political Economy*, 115(5).

Mancha, André. 2025. "When the State steps down: Reduced police surveillance and gang-related deaths in Brazil." *World Development*, 189: 106925.

Manning, Alan and Barbara Petrongolo. 2017. "How local are labor markets? Evidence from a spatial job search model." *American Economic Review*, 107(10): 2877–2907.

Manning, Alan and Joanna Swaffield. 2008. "The gender gap in early-career wage growth." *Economic Journal*, 118(530).

Manning, Alan and Rebecca Rose. 2021. "Ethnic minorities and the UK labour market." *CentrePiece - The magazine for economic performance*.

Manski, Charles F. 1993. "Identification of endogenous social effects: The reflection problem." *The review of economic studies*, 60(3): 531–542.

Mastrobuoni, Giovanni. 2015. "The Value of Connections: Evidence from the Italian-American Mafia." *Economic Journal*, 125(586): F256–F288.

Mastrobuoni, Giovanni and Eleonora Patacchini. 2012. "Organized crime networks: An application of network analysis techniques to the American Mafia." *Review of Network Economics*, 11(3).

McGregor-Smith, Ruby. 2017. "Race in the workplace: The McGregor-Smith Review."

McLean, Dawson, Sarah Tang, and Jack Worth. 2023. "The impact of training bursaries on teacher recruitment and retention: An evaluation of impact and value for money." National Foundation for Education Research.

Melnikov, Nikita, Carlos Schmidt-Padilla, and Maria Micaela Sviatschi. 2022. "Gangs, Labor Mobility and Development."

Mikusheva, Anna and Liyang Sun. 2021. "Inference with Many Weak Instruments." *The Review of Economic Studies*, 89(5): 2663–2686.

Misse, Michel. 2019. "The Puzzle of Social Accumulation of Violence in Brazil: Some Remarks." *Journal of Illicit Economies and Development*, 1(2): 177–182.

Monteiro, Joana and Rudi Rocha. 2017. "Drug Battles and School Achievement: Evidence from Rio de Janeiro's Favelas." *Review of Economics and Statistics*, 99(2): 213–228.

Murphy, Kevin M and Finis Welch. 1990. "Empirical Age-Earnings Profiles." *Journal of Labor Economics*, 8(2).

Nagler, Markus, Marc Piopiunik, and Martin R. West. 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." Department for Education.

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

OECD. 2015. "Education at a Glance 2015." OECD Indicators, Paris:OECD.

Peel, Leto, Daniel B Larremore, and Aaron Clauset. 2017. "The ground truth about metadata and community detection in networks." *Science Advances*, 3(5): e1602548.

Phan, Van, Carl Singleton, Alex Bryson, John Forth, Felix Ritchie, Lucy Stokes, and Damian Whittard. 2022. "Accounting for Firms in Ethnicity Wage Gaps Throughout the Earnings Distribution." IZA – Institute of Labor Economics.

Pinotti, Paolo. 2015. "The Economic Costs of Organized Crime: Evidence from Southern Italy." *Economic Journal*, 125(586): F203–F232.

Podgursky, Michael and Ruttaya Tongrut. 2006. "(Mis-)Measuring the Relative Pay of Public School Teachers." *Education Finance and Policy*, 1(4): 425–440.

Podgursky, Michael, Ryan Monroe, and Donald Watson. 2004. "The academic quality of public school teachers: an analysis of entry and exit behavior." *Economics of Education Review*, 23(5): 507–518.

Propper, Carol and John Van Reenen. 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.

Rho, Deborah and Seth Sanders. 2021. "Immigrant earnings assimilation in the united states: A panel analysis." *Journal of Labor Economics*, 39(1).

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica*, 73(2): 417–458.

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

Sain, Marcelo Fabian. 2015. *La regulación del narcotráfico en la provincia de Buenos Aires.* UMET, Universidad Metropolitana para la Educación y el Trabajo.

Scafidi, Benjamin, David L. Sjoquist, and Todd R. Stinebrickner. 2006. "Do Teachers Really Leave for Higher Paying Jobs in Alternative Occupations?" *Advances in Economic Analysis & Policy*, 6(1).

Schaefer, Daniel and Carl Singleton. 2020. "Nominal Wage Adjustments and the Composition of Pay: New Evidence from Payroll Data."

Schaub, Michael T, Jean-Charles Delvenne, Renaud Lambiotte, and Mauricio Barahona. 2019. "Multiscale dynamical embeddings of complex networks." *Physical Review E*, 99(6): 62308.

Schelling, Thomas C. 1971. "What is the business of organized crime?" *The American Scholar*, 643–652.

Shea, Susan C. 2015. "Crime and Security in Brazil: Brazil's Pacification Efforts in the Favelas of Rio de Janeiro."

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

Sibieta, Luke. 2023. "What has happened to teacher pay in England?" Institute for Fiscal Studies.

Silva, Leonardo. 2021. "Criminal Dynamics at the Brazil–Uruguay Border: The Rise of Organized Crime in Rio Grande do Sul." *Revista Brasileira de Segurança Pública*, 15(1): 115–138.

Silva, Luiz Antonio Machado da. 2004. "Sociabilidade violenta: por uma interpretação da criminalidade contemporânea no Brasil urbano." *Sociedade e estado*, 19: 53–84.

Sims, Sam and Asma Benhenda. 2022. "The effect of financial incentives on the retention of shortage-subject teachers: evidence from England."

Skaperdas, Stergios. 2001. "The Political Economy of Organized Crime: Providing Protection When the State Does Not." *Economics of Governance*, 2(3): 173–202.

Smulders, Anna, Stephanie Corte, Sarah Gohary, and Moravia De La O. 2017. *" Control... Over the Entire State of Coahuila": An Analysis of Testimonies in Trials Against Zeta Members in San Antonio, Austin, and Del Rio, Texas.* Human Rights Clinic, University of Texas School of Law.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming up inequality." *Quarterly Journal of Economics*, 134(1).

Stahlberg, Stephanie Gimenez. 2021. "Organized Crime and Violence in Brazil." PhD diss. Johns Hopkins University.

Stinebrickner, Todd R. 1998. "An empirical investigation of teacher attrition." *Economics of Education Review*, 17(2): 127–136.

Stinebrickner, Todd R. 2001*a*. "A Dynamic Model of Teacher Labor Supply." *Journal of Labor Economics*, 19(1): 196–230.

Stinebrickner, Todd R. 2001*b*. "Compensation Policies and Teacher Decisions." *International Economic Review*, 42(3): 751–780.

Stinebrickner, Todd R. 2002. "An analysis of occupational change and departure from the labor force: Evidence of the reasons that teachers leave." *Journal of Human Resources*, 37(1): 192–216.

Sviatschi, Maria Micaela. 2022. "Making a Narco: Childhood Exposure to Illegal Labor Markets and Criminal Life Paths." *Econometrica*, 90(4): 1835–1878.

The National Association of Schoolmasters Union of Women Teachers. 2024. "Teachers Wellbeing Survey Report 2024."

Tomaskovic-Devey, Donald, Melvin Thomas, and Kecia Johnson. 2005. "Race and the accumulation of human capital across the career: A theoretical model and fixed-effects application."

Tsao, Carolyn. 2025. "It's Not (Just) About the Money: The Value of Working Conditions in Teaching."

United Nations Office on Drugs and Crime. 2010. "Organized Crime Has Globalized and Turned into a Security Threat." *Press Release, UNODC (17 June 2010).*

United Nations Office on Drugs and Crime. 2011. "Estimating Illicit Financial Flows Resulting from Drug Trafficking and Other Transnational Organized Crimes." UNODC, Vienna.

United Nations Office on Drugs and Crime. 2019. *Global Study on Homicide 2019.* Vienna:UNODC.

Wood, Martin, Jon Hales, Susan Purdon, Tanja Sejersen, and Oliver Hayllar. 2009. "A test for racial discrimination in recruitment practice in British cities." Department for Work and Pensions 607.

World Bank. 2025. “Organized Crime and Violence Hold Back Latin America and the Caribbean.” *World Bank Press Release (28 April 2025).*

Worth, Jack. 2021. “Do bursaries change who applies to teacher training?” National Foundation for Education Research.

Worth, Jack and Dawson McLean. 2022. “What teachers do after leaving & implications for pay-setting.” National Foundation for Educational Research.

Worth, Jack, Sarah Tang, and Maria Galvis. 2022. “The impact of pay and financial incentives on teacher supply.” National Foundation for Educational Research.

Xu, Xiaowei. 2023. “Increasing concentration of high-skilled jobs in London means graduates elsewhere cannot fully capitalise on their education.” Institute for Fiscal Studies 3-4, London:Routledge.

Zwysen, Wouter and Neli Demireva. 2020. “Ethnic and migrant penalties in job quality in the UK: the role of residential concentration and occupational clustering.” *Journal of Ethnic and Migration Studies*, 46(1): 200–221.

Appendix A

Appendix to Chapter 1

A.1 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 wage gap measure is outlined in detail in chapter 2.11, but I describe minor differences in the process below. Whilst I use only one measure of wages in my main analysis, I construct multiple alternative measures for robustness. In all analysis I restrict my sample to full-time working age UK-born individuals working in England.

In the same process as described in chapter two, 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. These shares are more granular than those calculated in chapter 2.11 because the ITT trainee dataset contains information on an individual's undergraduate degree, something missing from the SWC. 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 with undergraduate classification d and subject j. 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, otherwise known as the continuous weighted measure in chapter 2.11. 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[1]. 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. This differs to the method of estimation in chapter 2.11, where I estimate teacher wages using ASHE and age, because it is important to generate wages using a precise measure of tenure (which is not included in ASHE). As can be seen in figure 2.3, ASHE and SWC wages closely match each other for early career teachers.

I predict both outside wages and teacher wages for trainees for each year since their training 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 A.1 below. I estimate the wage gap as a raw difference, rather than the difference in logs, so that the wage gap coefficient is directly comparable to the bursary coefficient in the main regressions.

$$\widehat{\text{Wage Gap}}_{i,y,x,t} = \widehat{\text{Teacher Wage}}_{i,y,t} - \widehat{\text{Outside Wage}}_{i,y,x} \tag{A.1}$$

I also construct alternative measures of outside wages, which mirror the methods described in chapter 2.11. These include the 'binary weights' outside wage (which includes only jobs with over 80% graduates), the 'leavers' wage (which includes only jobs held by ex-teachers), and the 'LFS Graduate' wage[2]. I also establish a 'basic' wage gap which calculates the continuous job weights for just two groups: STEM and

[1]coefficients on sex and ethnicity are not included as they are not consistently present in the SWC data, but were small and not significant when included.

[2]For this exercise, the controls included in the LFS regression are year, age, age squared, sex, region, a full set of their interactions, undergraduate classification, undergraduate subject, and ethnicity

non-STEM graduates.

The continuous weights measure of outside wage is the median measure of the five methods. Graduate (binary weights) wages predict the highest average, followed by the LFS measure. The basic average wage is second lowest, and the leaver measure predicts the lowest wages. When controlling for outside wages in tables 1.4, 1.5, 1.6, and 1.7, 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 measures can produce more extreme estimates. Most importantly, the coefficient of the bursary level on key outcomes is reliably consistent across these measures.

A.2 Model Simulations

In this section, I discuss how I simulate the effects of the model outlined in section 1.7. 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 a 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 with a standard deviation of £2k that follow an AR(1) error procedure with autocorrelation of 0.7.

At t=0, the agent observes 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[3]. An individual then selects their job for t=1 in order to 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.

A.3 Cost-Benefit Analysis

I evaluate the cost-effectiveness three retention and recruitment policies: Bursaries, increasing early-career salaries[4], 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 teaching 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

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

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

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[5].

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 1.1 and 1.2. 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 A.2. 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 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, Tang, and Galvis (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

[5] 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]

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)[6].

[6] 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).

A.4 Tables

Table A.1: 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 A.2: 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
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 A.3: 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 A.4: Cohort Size Pooled Regression Results: Subject-Year Level

	All Subjects		Stem Subject		Non-Stem Subject	
	Trainees	Teachers	Trainees	Teachers	Trainees	Teachers
Bursary	0.196**	0.067	0.223	0.026	0.150	0.077
	[0.034]	[0.334]	[0.360]	[0.920]	[0.174]	[0.474]
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 A.5: Cohort Size Pooled Regression Results: Subject-Year Level, controlling for competing subjects

	All Subjects		Stem Subject		Non-Stem Subject	
	Trainees	Teachers	Trainees	Teachers	Trainees	Teachers
Bursary	0.234**	0.146*	0.263	0.161*	0.174	-0.040
	[0.034]	[0.098]	[0.220]	[0.054]	[0.116]	[0.372]
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 and the average bursary offered in competing subjects in that year, with wild-bootstrapped standard errors clustered at the subject level.

Table A.6: 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.

Appendix B

Appendix to Chapter 2

B.1 Delta Method Standard Errors

I employ the delta method to estimate the standard errors of the predicted wage gaps, as outlined in Liu (2012). Let the mean of the estimated wage gap be W. Given that W is a linear combination of two predicted wages, the wage gap can be written as:

$$\bar{W} = \bar{X}_T\hat{\beta}_T - \bar{X}_G\hat{\beta}_G \tag{B.1}$$

Where $\bar{X}_T$ and $\bar{X}_G$ are vectors of the weighted average values of the regressor variables in the teaching and graduate regressions, and $\hat{\beta}_T$ and $\hat{\beta}_G$ are the corresponding coefficient estimates. The wage gap is a linear combination of the $\bar{X}$ vectors and the random variables $\hat{\beta}$. The approximate standard error using the delta method is therefore:

$$\mathrm{SE}(\bar{W}) = \sqrt{\bar{X}_T\,\widehat{\mathrm{Var}}(\hat{\beta}_T)\,\bar{X}_T^\top + \bar{X}_G\,\widehat{\mathrm{Var}}(\hat{\beta}_G)\,\bar{X}_G^\top} \tag{B.2}$$

In other words, the standard error is the square root of the estimated sum of variances for the average teaching wage and the average graduate wage. Each variance is a product of the covariance matrix of the estimated coefficients, $\hat{\beta}_T$ and $\hat{\beta}_G$, and the weighted average of the regressors, $\bar{X}_T$ and $\bar{X}_G$.

B.2 Tables

Table B.1: Average predicted wage gaps by source - consistent sample size

Data source	Method	Mean	Std. Error (delta)	Total Observations
LFS	Grads	0.057	0.107	1,371,583
LFS	Grads, stem	0.018	0.136	1,371,583
ASHE	Grads, Continuous weights	0.137**	0.068	1,371,583
ASHE	Grads, Binary weights	-0.024	0.077	1,371,583
ASHE	Leaver weights	0.321***	0.083	1,371,583

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors were calculated using the delta method. Table reports the average predicted wage gap,
$ln(\text{Weekly Teacher Wage}) - ln(\text{Weekly Counterfactual Wage})$ for the sample of teachers in the SWC using different alternative wage counterfactuals. The teacher wage is estimated using full-time state school teachers.

B.3 Figures

Figure B.1: Distribution of full-time teacher hourly wages by data source

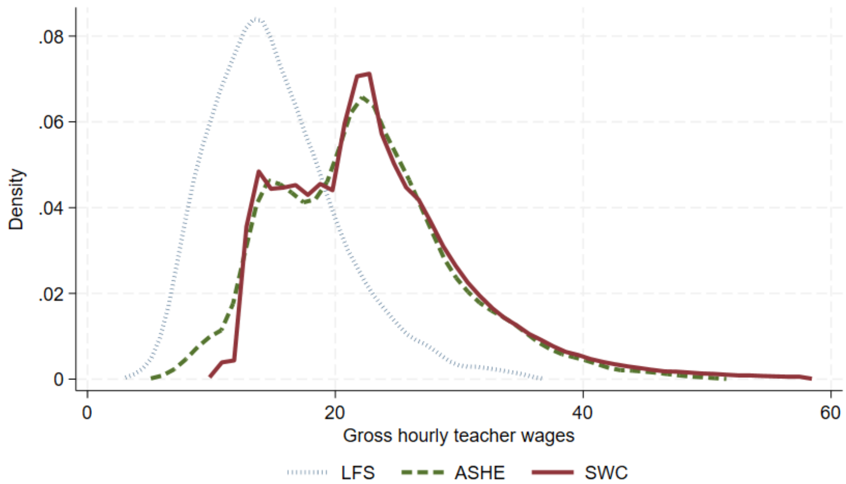


Figure outlines the kernel density of hourly full-time pay for state school teachers in England.

Figure B.2: Estimated average teacher wage gap by age

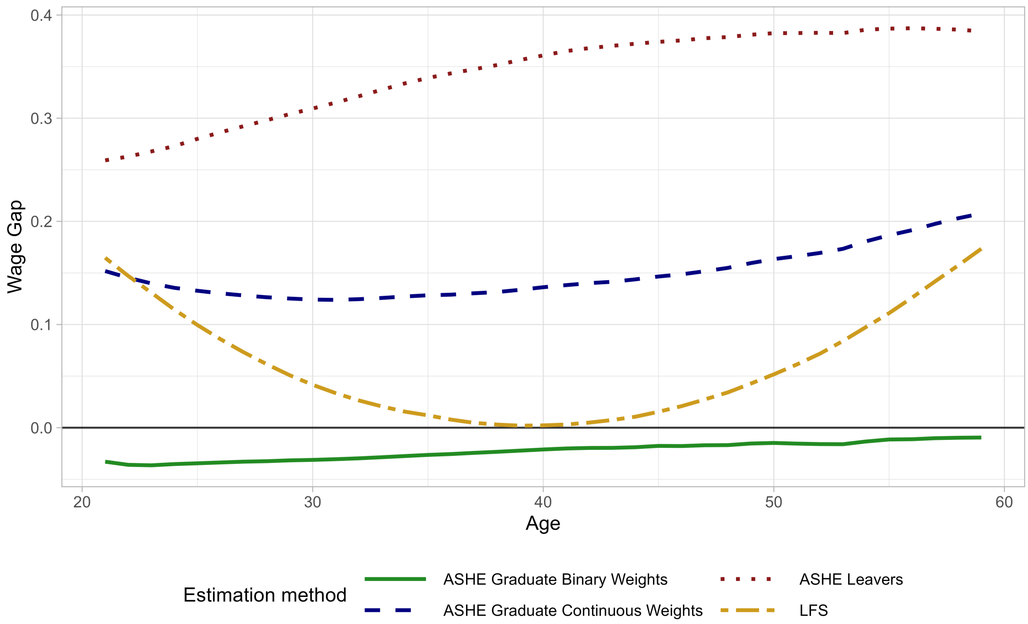


Figure estimates the average predicted wage gap, ln(Weekly Teacher Wage) − ln(Weekly Counterfactual Wage) for different age groups using the four key estimation methods.

Figure B.3: Estimated average teacher wage gap by region: Leavers method

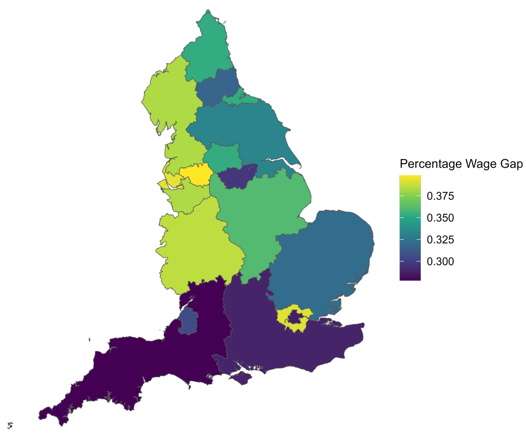


Figure estimates the average predicted wage gap, ln(Weekly Teacher Wage) − ln(Weekly Counterfactual Wage) for different regions in England. Major subregions, such as the metropolitan areas of Bristol, Manchester, and Tyne and Wear are considered separately. Note that the legend is not the same as the corresponding maps for the other wage gaps. This allows for sufficient variation in colour within maps.

Appendix C

Appendix to Chapter 3

C.1 Addressing potential selection in the ASHE-Census 2011 dataset

In Table C.1, we compare the entire population of employees aged 20 – 50 in the 2011 Census to the sub-set of those in the 2011 Census which have been matched to ASHE records. Migrants and ethnic minority workers are under-represented in the ASHE – 2011 Census sub-sample. This is because linkage quality is lower in London.

Non-random linkage between ASHE and the 2011 Census might bias our estimated pay gaps, particularly if linkage quality is also correlated with our outcome variable, basic hourly pay. To explore this, we estimate a probit model for employees aged 20 – 50 in ASHE 2011 in which a dummy variable identifying Census linkage is regressed on basic log hourly pay. Linkage rates tend to rise with basic hourly pay, but this correlation is only modest with a marginal effect of log hourly pay on the probability of Census linkage of 0.114 and overall pseudo-R-squared of 0.0119. If we include the controls we use in our main regressions (year, region, age fixed effects) in this probit model, the marginal effect falls to 0.0616.

Furthermore, our results are robust to applying sample weights which have been constructed by the ONS to adjust the earnings profile of the linked ASHE-Census to match the profile of the full ASHE sample.

As a further indication that the sample selection problem is not severe, we use the Quarterly Labour Force Survey (QLFS) to check the representativeness of our ASHE – 2011 Census sample. We replicate our results using the Labour Force Survey (LFS) from 1999 to 2018 using the same population sub-samples. Black Caribbean and Black African groups have been merged order to maintain consistent ethnicity definitions across the sample years. In Table C.5 we find that compared to the ASHE-Census sample, the LFS sample is on average younger with lower pay. Some migrant groups also have markedly different shares of people holding degrees or living in London.

The analysis in Table 3.3 is replicated in Table C.6, with the exclusion of firm-level controls. Most wage gaps in column 1 are within three percent of their previous estimate excluding those of white migrants whose wage gaps are over ten percent larger in this sample. Male UK-born Indians have an estimated gap seven percent smaller than estimated in the ASHE-census sample. Controlling for education in column 2 reduces the disparity for all groups, excluding the male UK born Pakistani-Bangladeshi and Indian groups where the difference is a similar magnitude. Column 3 has larger disparities due to the absence of firm controls.

Figure C.1 demonstrates similar patterns in wage progression compared to figure 3.3, particularly the UK-born groups. Within the male migrant group, the combined Black group has more positive wage growth relative to the Pakistani and Bangladeshi group. The Indian migrant group shows more wage growth at earlier ages, but a pronounced wage cut after age 30 for both men and women. The Bangladeshi and Pakistani male migrant group also experience wage stagnation in later years.

In summary, the ASHE-Census findings are generally consistent with those produced using the LFS. The largest differences can be found when comparing migrant results. This suggests that the ASHE-Census could contain a more representative and reliable resource for analysis of migrant and groups.

C.2 Supplementary Tables

Table C.1: Descriptive statistics and balance. All employees aged 20 – 50.

	2011 Census (%)	2011 Census linked with ASHE (%)
Male	47.88	50.17
UK born	82.39	93.73
Bachelor's degree (or higher)	35.55	36.42
White	85.19	93.12
Black Caribbean	1.24	1.21
Black African	2.04	1.26
Indian	3.10	2.92
Pakistani / Bangladeshi	2.40	1.49

Source: 2011 Census in England and Wales and ASHE-Census linked sample in England and Wales

Table C.2: Probit regression for census linkage on log hourly pay.

	Census link (1)	Census link (2)
Log hourly pay (marginal effect at mean)	0.114***	0.0616***
	(0.00056)	(0.00703)
N	2577835	2031094
Pseudo R-squared	0.0119	0.0380
FE		
Region - year		X
Age		X

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: ASHE - 2011 Census linked sample in England and Wales.

Table C.3: Pooled cross-sectional wage penalty regression - Men

	(1)	(2)	(3)	(4)
Men	log wage	log wage	log wage	log wage
White Migrant	0.0993***	0.0415***	0.0276***	0.00755
	(0.00543)	(0.00480)	(0.00514)	(0.00428)
Black Caribbean Migrant	-0.319***	-0.216***	-0.133***	-0.0868***
	(0.0142)	(0.0137)	(0.0133)	(0.00990)
Black Caribbean UK	-0.267***	-0.221***	-0.197***	-0.123***
	(0.00795)	(0.00732)	(0.00744)	(0.00614)
Black African Migrant	-0.358***	-0.414***	-0.308***	-0.165***
	(0.00837)	(0.00761)	(0.00738)	(0.00546)
Black African UK	-0.157***	-0.279***	-0.195***	-0.121***
	(0.0196)	(0.0188)	(0.0176)	(0.0145)
Indian Migrant	-0.0720***	-0.124***	-0.0851***	-0.0833***
	(0.00738)	(0.00635)	(0.00665)	(0.00490)
Indian UK	-0.0230**	-0.0978***	-0.129***	-0.0836***
	(0.00717)	(0.00653)	(0.00658)	(0.00523)
Pakistani/Bangladeshi Migrant	-0.231***	-0.239***	-0.149***	-0.109***
	(0.00996)	(0.00838)	(0.00869)	(0.00665)
Pakistani/Bangladeshi UK	-0.114***	-0.151***	-0.155***	-0.110***
	(0.00855)	(0.00774)	(0.00743)	(0.00606)
Cons	2.626***	2.630***	2.633***	2.630***
	(0.000767)	(0.000670)	(0.000528)	(0.000431)
N	437587	437522	427293	427292
R-squared	0.224	0.415	0.697	0.801
FE				
Region - year	X	X	X	X
Age	X	X	X	X
Educ - year		X	X	X
Firm			X	X
Occupation				X

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Regression uses various controls and using sample weights. Source: ASHE - 2011 Census linked sample in England and Wales.

Table C.4: Pooled cross-sectional wage penalty regression - Women

	(1)	(2)	(3)	(4)
Women	log wage	log wage	log wage	log wage
White Migrant	0.0928***	0.0316***	0.0120**	0.00791*
	(0.00453)	(0.00391)	(0.00424)	(0.00334)
Black Caribbean Migrant	-0.174***	-0.160***	-0.127***	-0.0762***
	(0.00883)	(0.00785)	(0.00797)	(0.00620)
Black Caribbean UK	-0.0569***	-0.0493***	-0.0499***	-0.0235***
	(0.00663)	(0.00602)	(0.00623)	(0.00470)
Black African Migrant	-0.171***	-0.252***	-0.196***	-0.106***
	(0.00675)	(0.00575)	(0.00615)	(0.00470)
Black African UK	-0.0364**	-0.202***	-0.188***	-0.101***
	(0.0129)	(0.0129)	(0.0131)	(0.0100)
Indian Migrant	-0.104***	-0.133***	-0.102***	-0.0667***
	(0.00625)	(0.00531)	(0.00541)	(0.00384)
Indian UK	0.0667***	-0.0384***	-0.0765***	-0.0351***
	(0.00672)	(0.00606)	(0.00668)	(0.00499)
Pakistani/Bangladeshi Migrant	-0.143***	-0.131***	-0.112***	-0.0510***
	(0.0125)	(0.0103)	(0.0116)	(0.00759)
Pakistani/Bangladeshi UK	-0.0279***	-0.0779***	-0.0834***	-0.0461***
	(0.00837)	(0.00742)	(0.00800)	(0.00580)
Cons	2.407***	2.412***	2.416***	2.413***
	(0.000732)	(0.000614)	(0.000518)	(0.000399)
N	440365	440334	430321	430317
R-squared	0.165	0.416	0.649	0.792
FE				
Region - year	X	X	X	X
Age	X	X	X	X
Educ - year		X	X	X
Firm			X	X
Occupation				X

Standard errors in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001. Regression uses various controls and using sample weights. Source: ASHE - 2011 Census linked sample in England and Wales.

Table C.5: Sample characteristics by sex, ethnicity, and migrant status.

UK Born	**Men**					**Women**				
	Hourly wage	Age	Degree (%)	In London (%)	N	Hourly wage	Age	Degree (%)	In London (%)	N
White	13.10	36.53	36%	8%	262,503	10.29	36.69	37%	7%	287,396
	(8.58)	(8.41)				(6.77)	(8.45)			
Black	11.90	36.02	36%	49%	2.320	11.50	36.23	46%	55%	3,456
	(7.01)	(7.88)				(6.48)	(7.89)			
Indian	14.38	32.45	58%	31%	2,511	12.17	32.11	60%	33%	2,762
	(9.89)	(7.26)				(7.33)	(7.32)			
Pakistani/ Bangladeshi	11.65	30.30	50%	22%	1,573	10.23	30.19	50%	24%	1,588
	(8.58)	(7.02)				(6.31)	(7.19)			
Total	13.10	36.45	36%	8%	268,907	10.32	36.60	38%	8%	295,202
	(8.59)	(8.41)				(6.77)	(8.45)			

Migrant	**Men**					**Women**				
	Hourly wage	Age	Degree (%)	In London (%)	N	Hourly wage	Age	Degree (%)	In London (%)	N
White	14.71	35.20	42%	26%	20,285	9.87	37.29	48%	27%	22,918
	(10.80)	(7.68)				(5.38)	(7.71)			
Black	11.00	37.34	47%	44%	3,721	9.87	37.29	47%	51%	4,515
	(6.49)	(7.70)				(5.38)	(7.71)			
Indian	14.79	37.44	53%	37%	4,853	11.24	37.77	47%	39%	4,298
	(10.25)	(7.29)				(7.71)	(7.38)			
Pakistani/ Bangladeshi	9.52	35.40	34%	33%	3,258	9.67	35.23	37%	22%	1,278
	(7.14)	(7.54)				(6.94)	(7.88)			
Total	13.77	35.81	43%	30%	32,117	11.38	35.76	47%	32%	33,009
	(10.15)	(7.67)				(7.68)	(7.95)			

Means reported and standard deviations in parentheses. Source: LFS Individual – 1999-2018.

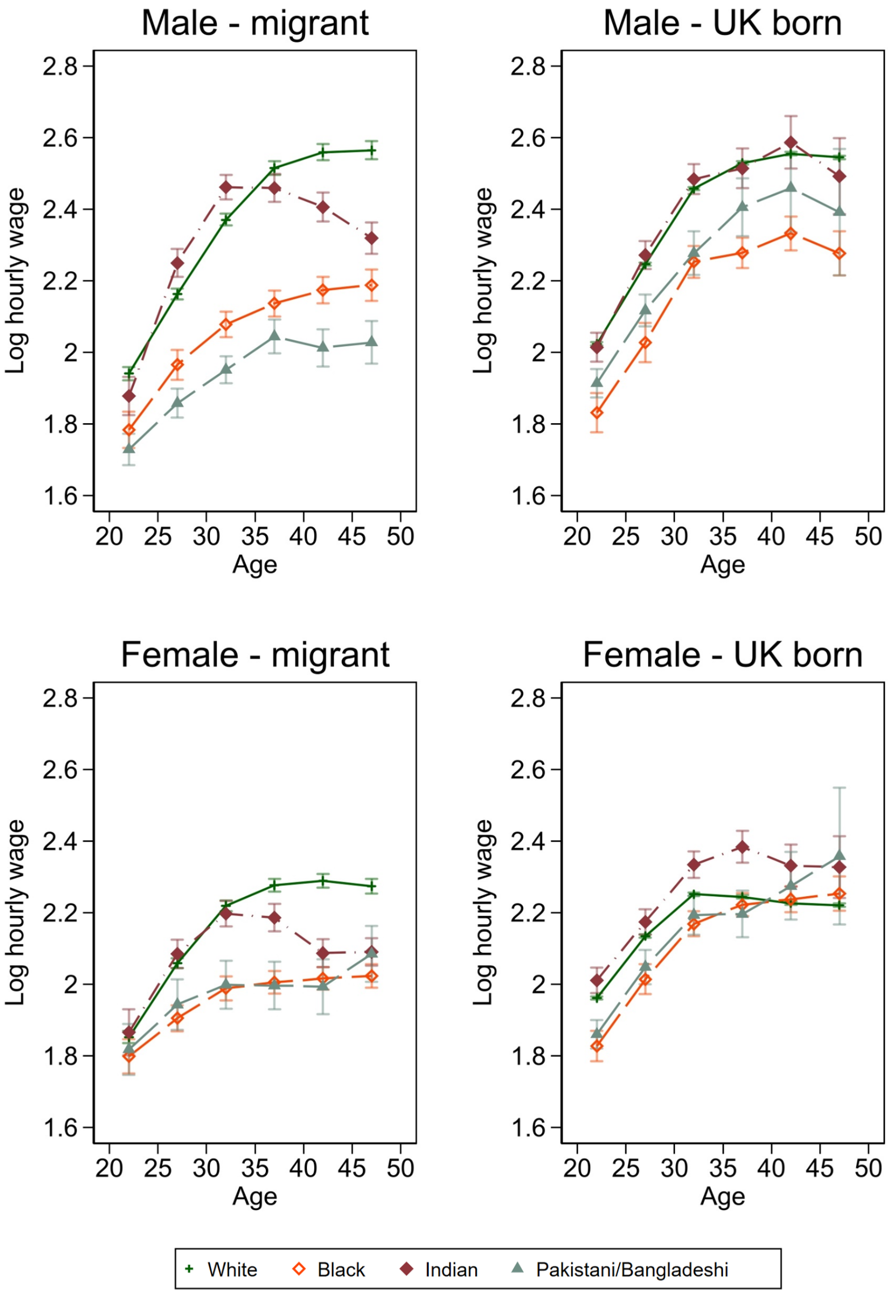
Table C.6: Cross-sectional regression table of log hourly wage gap with white UK born

VARIABLES	Men (1)	Men (2)	Men (3)	Women (1)	Women (2)	Women (3)
White migrant	-0.0393***	0.0135***	-0.114***	-0.0161***	0.00490	-0.0287***
	(0.00419)	(0.00402)	(0.00848)	(0.00376)	(0.00361)	(0.00665)
Black migrant	-0.342***	-0.314***	-0.0663***	-0.235***	-0.207***	-0.0373***
	(0.00862)	(0.00810)	(0.00860)	(0.00751)	(0.00687)	(0.00742)
Black UK born	-0.219***	-0.184***	-0.102***	-0.0556***	-0.0626***	-0.0802***
	(0.0105)	(0.00960)	(0.0108)	(0.00836)	(0.00768)	(0.0101)
Indian migrant	-0.0810***	-0.0815***	0.0145***	-0.0941***	-0.0746***	0.00236
	(0.00884)	(0.00792)	(0.00331)	(0.00865)	(0.00788)	(0.00298)
Indian UK born	0.0199*	-0.0545***	-0.155***	0.0719***	-0.00865	-0.105***
	(0.0107)	(0.00965)	(0.00669)	(0.00925)	(0.00843)	(0.00591)
Pakistani/Bangladeshi migrant	-0.451***	-0.342***	-0.0857***	-0.222***	-0.158***	-0.0651***
	(0.00997)	(0.00885)	(0.00621)	(0.0152)	(0.0135)	(0.00625)
Pakistani/Bangladeshi UK born	-0.115***	-0.153***	-0.214***	-0.0768***	-0.104***	-0.0853***
	(0.0133)	(0.0119)	(0.00752)	(0.0125)	(0.0114)	(0.0119)
Constant	2.421***	2.418***	2.413***	2.190***	2.189***	2.187***
	(0.000999)	(0.000908)	(0.000794)	(0.000946)	(0.000841)	(0.000716)
Observations	301,024	298,353	298,256	328,211	326,000	325,944
R-squared	0.213	0.356	0.520	0.178	0.353	0.535
Controls						
Region-Year	X	X	X	X	X	X
Age	X	X	X	X	X	X
Education-Year		X	X		X	X
Occupation			X			X

Robust Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Source: LFS Individual – 1999-2018.

C.3 Supplementary Figures

Figure C.1: Log hourly wages by ethnic group and age group



Year-region controls. Source: LFS Individual – 1999-2018

Appendix D

Appendix to Chapter 4

D.1 Supplementary Details

Parenting Gangs Labelling

In Section 4.5.2, we described our method for assigning gang labels to individuals in the network based on their prison stay histories. This process relies on gallery-level mappings that associate specific prison galleries with particular gang affiliations. The dataset identifies 39 distinct gangs, encompassing both large, state-wide criminal factions and smaller, neighbourhood- or city-level groups (hereafter referred to as *local gangs*).

For analytical purposes, we consolidate many of these local gangs into corresponding *parent gangs*—larger criminal organizations that exert varying degrees of influence over their local counterparts. These parent-local relationships often involve the provision of protection, access to drug and arms supply chains, and support in dispute resolution. In exchange, local gangs typically adopt the parent gang's norms, symbols, and strategic objectives, such as coordinated territorial control or alignment in broader inter-gang conflicts. These affiliations range from loose alliances to more hierarchical forms of subordination, but they consistently entail some level of operational interdependence.

We apply this same consolidation to the police intelligence dataset used for validation, which records known gang affiliations of a sample of individuals. Harmonizing gang labels across both the inferred and validation datasets ensures consistency and reflects the collaborative dynamics between factions observed in the field.

We adopt this aggregation strategy for two primary reasons. First, merging local gangs with their parent factions allows us to assign gang labels to a larger portion of individuals. Local gangs often have few or mixed gallery associations - e.g., a gallery may be jointly labelled as affiliated with both Gang A and Gang B[1] - making it difficult to assign a clear label based solely on prison history. Individuals affiliated with such gangs are frequently housed in galleries linked to larger factions with which they collaborate or peacefully coexist. Without merging these affiliations, many local gang members would remain unlabelled due to insufficient exposure to any single gang's galleries. However, if local gangs are functionally integrated with a parent faction, then assigning the parent gang's label still carries meaningful interpretative value. These relationships are often reflected in the network structure, where members of the parent and local gangs are embedded within the same communities. Indeed, when we apply the optimal stable Markov partitioning algorithm without accounting for parent-local merges, we observe several dual-gang communities that align with known local-parent affiliations from intelligence sources[2]. Thus, aggregating labels improves label coverage while preserving interpretability in terms of underlying inter-gang cooperation.

Our second motivation is to improve the validity of comparisons with the police intelligence dataset. Not all gangs present in the gallery label dataset appear in the validation data, and vice versa. For example, if an individual is known (via police intelligence) to be a member of Gang X, which is a local affiliate of parent Gang A, but the gallery data lacks any reference to Gang X, the individual may still be labelled as belonging to Gang A based on prison history. If we did not absorb local gangs into parent factions, such a case would be mistakenly treated as a labelling error, despite reflecting the true underlying affiliation structure. Consolidating labels ensures that functionally equivalent affiliations are appropriately matched across datasets.

To determine which local gangs operate under or in collaboration with larger factions, we drew on intelligence provided by three law enforcement bodies: the civil police, the military police, and prison intelligence. Each agency independently supplied a list of local gangs and their associated parent factions. We then harmonised the information across the three sources and convened discussions among the agencies

[1] Often galleries associated with local gangs are mixed: for example, a gallery may be labelled as "Gang A / Gang B".

[2] This is observed within the analysis when not accounting for parent gangs: we find dual-gang communities that are consistent with known parent-local alignments provided by police intelligence.

to resolve discrepancies and establish a final, consensus-based crosswalk. Using this mapping, we replaced all local gang labels with their corresponding parent gang labels in both the prison gallery and validation datasets. As a result of this harmonisation, the 39 gangs originally identified in the gallery data were condensed into 15 parent gangs, and the 60 gangs identified in the validation dataset were condensed into 24.

D.2 Supplementary Tables

Table D.1: Summary list of gang related offences

Offence
Contraband/smuggling
Counterfeiting money
Crimes occurring during the investigation or obtaining of evidence related to organized crime.
Criminal/unlawful association
Disappearance of an individual
Kidnapping
Drug related offences (trafficking, possession, manufacture, prescription of narcotics)
Extortion
Favouring prostitution/ sexual exploitation
Firearm related offences (shooting, illegal possession, theft, trading, trafficking)
Fraud/false identity
Homicide/Murder
Human trafficking for the purpose of sexual exploitation
Illegal weapon offences (possession, manufacture, trading)
Illegal gambling
Incitement to crime
Money laundering
Manslaughter
Other crimes against life
Prison escape/riot
Providing child/adolescent weapons, ammunition, or products that cause dependence
Robbery (resulting in injuries, rape, death)
Suicide instigation/aid
Tampering with the identification number of a motor vehicle
Theft (over 65 varieties including theft of phones, vehicles, weapons and banks)
Torture

This table summarises the set of offences that are disproportionately associated with gang activity. The list was compiled based on a survey administered to several police officers who specialise in organized crime.

Table D.2: Summary of coverage: Snapshots of prison gallery gang affiliation

	Coverage		
Date of snapshot	Number of blocks	Number of prisons	Number of inmates
05/2017	1,597	20	26,118
04/2018	2,347	36	35,603
04/2019	2,112	34	32,108
10/2019	2,727	42	32,752
04/2020	2,692	49	36,670
05/2021	3,163	55	43,686
06/2022	3,485	60	44,954
06/2023	3,619	65	39,712
12/2023	3,331	69	29.993
01/2024	3,428	71	31,288
07/2024	3,506	73	33,076

The number of blocks denote the number of prison blocks, or sub-galleries. This unit of measurement is used as the size and definition of a gallery varies between prisons.

Table D.3: Summary List of Data Sources

Data	Data Source	Data Type	Unit(s) of observation
occurrence data	police	administrative	individual-occurrence
gang membership hierarchies	police	intelligence	individual
gang related offence list	police	intelligence	offence
prison stay data	prison	administrative	individual
risk profiles	prison	intelligence	individual
gallery affiliation mappings	prison	intelligence	gallery-block
population registry	population registry	administrative	individual

This table summarises the properties of the data used in the study. A description of each data set is outlined in section 4.4.

Table D.4: Share of Gang Network Matched with Intelligence Validation List

Order of Individual	Share Matched	Count Unmatched	Count Matched
0	0.52	319	343
1	0.15	6,349	1,106
2	0.04	19,326	716

This table outlines the share of the network, generated using prison intelligence and police occurrence data, that are matched with the validation data. The validation data is police intelligence data of gang membership and hierarchies. The order of the individual is their hierarchy within the network: where order zero indicates a potential gang leader and order two indicates that they are two links (co-offences) away from a gang leader.