I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17273

The Effect of Wages on Job Vacancy Duration: Evidence from a Spatial Discontinuity

Charles Carter
Judith M. Delaney
Kerry L. Papps

SEPTEMBER 2024

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17273

The Effect of Wages on Job Vacancy Duration: Evidence from a Spatial Discontinuity

Charles Carter
University of Bath

Judith M. Delaney
University of Bath, University College London and IZA

Kerry L. Papps
University of Bradford and IZA

SEPTEMBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effect of Wages on Job Vacancy Duration: Evidence from a Spatial Discontinuity*

We exploit a spatial discontinuity in the wages paid by the United Kingdom's National Health Service to examine how wages affect the length of time a vacancy is advertised. NHS workers in inner London are required by law to be paid 4.3% more than those who work in outer London. We use a regression discontinuity design and estimate an elasticity of vacancy duration with respect to wages of -6.3. This number is larger than reported by previous studies and suggests that firms can fill worker shortages faster by raising wages. This also highlights the importance this margin of worker recruitment when analysing firm search and job match. Our results are robust to various checks including a placebo test using fictitious borders and are robust to changes in the bandwidth and the duration measure. The estimates are similar across all occupational groups in the NHS and are not limited to jobs that require specific skills such as nurses and therapists. Our results provide evidence for policy makers which suggests that increasing the wages paid to NHS workers may lead to increased cost savings by reducing the need to hire expensive agency staff and may also lead to better health outcomes of the population through reduced staff shortages.

JEL Classification: J22, J23, J31, J38

Keywords: vacancy duration, wages, employer search

Corresponding author:
Judith M. Delaney
University of Bath
Claverton Down
Bath
BA2 7AY
Great Britain

E-mail: judithmdelaney@gmail.com

* We would like to thank Jonathan James, Cristina Lafuente, and Chris Martin for helpful comments as well as participants at the Royal Economic Society (RES) Annual Conference in Belfast 2024 and at the 6th IDSC of IZA Workshop on Matching Workers and Jobs.

1. Introduction

Economic models often predict that where firms have some degree of wage setting power, paying higher wages should enable them to hire more workers. In particular, the existence of search frictions in the labour market means that a firm faces a trade-off between the wage it offers and the search costs associated with an open vacancy (Manning, 2003). By offering a higher wage a firm can fill a vacancy faster, at the expense of lower profit per worker. This theoretical relationship between wages and vacancy duration has found limited support in the empirical literature (Faberman and Menzio, 2018; Mueller et al., 2024). These studies have relied mainly on observational data and so may be hampered by the endogeneity of wages and vacancy durations. Since firms may post higher wages or change their hiring efforts for those vacancies which are harder to fill, and because workers may perceive higher paying jobs as those with more competition, it is very difficult to parse out the direct causal effect of wages on vacancy duration.

In this paper, we examine whether the speed at which a vacant post is filled is related to the wage at which the post is advertised by exploiting a spatial discontinuity in the wages paid by the UK's National Health Service (NHS), one of the largest employers in the world. Due to the higher costs of living in London compared to the rest of the UK, UK Government regulations require that the NHS pays a 15% salary premium to workers in outer London and a 20% premium for workers in inner London compared to the nationally defined pay scale. We compare otherwise identical vacancies on either side of the border between inner and outer London and examine how long they are advertised for. By exploiting an exogenous source of variation in wages, we are able to circumvent the problem of endogeneity between posted wages and vacancy duration. In addition, unlike other studies that have been plagued by measurement error in wages, our estimates do not rely on wage data as the wage differential is determined by government policy. Using a sharp regression discontinuity design, we find that a given vacancy is advertised for approximately 6 fewer days on average in inner London, and that the elasticity of vacancy duration with respect to wages is -6.3. This is considerably larger than the estimates found by most previous empirical studies, suggesting that those estimates are significantly biased.

Our findings are particularly important for firm behaviour. Large firms face a constant need to replace departed workers and unfilled vacancies can impose significant costs, in the form of reduced production, higher overtime payments to remaining workers or the need to hire costly temporary workers. If a small increase in wages significantly reduces the time required to fill a post, then firms are likely to be more willing to use this mechanism to mitigate the costs associated with job turnover than has previously been assumed.

The previous evidence on the relationship between vacancy duration and wages has been mixed, both in terms of the direction and magnitude. Faberman and Menzio (2018) find a positive relationship between posted wages and vacancy duration. They use US survey data from 1980 and suggest that one of the reasons for this positive relationship could be due to a lack of data on detailed job titles. The underlying dynamics of this seemingly paradoxical relationship is further explained by Marinescu and Wolthoff (2020) who examine the significance of wage information displayed on job listing platforms on the number of applicants a posting receives. They find that controlling for a vacancy's Standard Occupational Classification (SOC) code produces a negative relationship between the posted wage and the number of applicants the vacancy attracts which would imply a positive relationship between posted wages and vacancy duration. After controlling for the more specific job title, however, they find that higher wages attract more applicants and correspondingly result in lower vacancy duration. Similarly, Mueller et al. (2024) using administrative Austrian data from 1997-2014 find the positive relationship between starting wages and vacancy duration they observe becomes small and negative with sufficient controls for job and firm heterogeneity.[1]

While controlling for job title prevents one source of bias, the estimated relationship between duration and wages may still be driven by unobserved heterogeneity. The paper that is closest in spirit to ours is the recent paper by Bassier et al. (2023). They use UK data on online job adverts and exploit within-firm, discrete wage changes as well as external pay settlements to analyse the elasticity between posted wages and vacancy duration. By focusing their analysis on wage changes that are arguably unrelated to the difficulties in finding a worker or the wage offers of other firms they can get closer to a causal estimate than other studies. They find relatively large elasticities ranging between -3 and -5. Our work is similar to theirs as we also use an exogenous source of variation in wages, but our work does not rely on an actual estimate of wages (since we know the government mandated pay premium) which limits the impact of measurement error. We find estimates that are similar but slightly larger than those found by Bassier et al. (2023), suggesting that the use of identification strategies that rely on exogenous wage variation indicates that wages indeed have a very large effect on the duration of job vacancies.

The rest of the paper proceeds as follows: Section 2 outlines the institutional setting. Section 3 discusses the data we use, Section 4 outlines the empirical strategy, and Section 5 discusses the results. Section 6 includes several robustness tests, Section 7 examines heterogeneity, Section 8 discusses policy implications, and Section 9 concludes.

[1] These findings are also consistent with the results of Belot et al. (2018), who undertake an experimental study and find that higher wage vacancies tend to attract a greater number of 'saves' from job seekers. This aligns with the concept of a trade-off between higher wages and the perceived competition for a job, which can shed light on the positive relationship between wages and the observed duration of job vacancies, as observed by Faberman and Menzio (2018) and Banfi and Villena-Roldan (2019). In addition to these empirical observations, Belot et al. (2018) formulate a directed search model that incorporates multiple applications and on-the-job search. Through calibration using UK data, this model effectively replicates the negative relationship between wages and job vacancy duration that their empirical investigations uncover.

2. Institutional Setting

The NHS is a publicly funded healthcare system with the aim of providing services that are free at the point of use for all UK citizens. However, evidence suggests that the NHS is critically understaffed (The Health Foundation, 2020; Morgan, 2022). As of September 2021, the NHS had a waiting list of almost 7 million people and had 99,460 vacant positions. This number of shortages corresponded to a job vacancy rate of 12% for nurses and 8% for doctors (Heath and Social Care, 2022).

The pay for most NHS workers is set out by a system called "Agenda for Change". Doctors, dentists and senior managers are covered by a separate system and are excluded from our analysis. As part of Agenda for Change, staff working at locations in and around London are entitled to a High Cost Area Supplement (HCAS) to account for the higher cost of living in London. Specifically, NHS staff working in fringe, outer or inner London are paid an additional 5%, 15% and 20% on top of their base salary, respectively. There is a minimum HCAS payment, which is paid to workers with salaries below a published level. Similarly, there is a maximum HCAS payment for workers earning above a certain amount.

In this paper, we compare vacancies in inner and outer London. An NHS worker in inner London is paid a 4.35% higher wage than a worker doing an identical job in outer London, due to the difference in HCAS rates.[2] There are two main reasons why we decide to exclude fringe London from the analysis. First, the vicinity of the inner-outer border is densely populated, well connected by public transport and served by many NHS practices. In comparison, the outer-fringe London border and fringe-non-London border are much longer and run through relatively sparsely populated areas with distinct and poorly linked population

[2] To show this, consider a base salary, X. Workers in inner London earn 1.2X while workers in outer London earn 1.15X. The percentage difference is therefore $(1.2X - 1.15X) / 1.15X \approx 4.35\%$.

centres. Second, the outer-fringe London border coincides with the Greater London Authority border and with the border for the Integrated Care Boards (ICB). The ICBs are responsible for planning health services for their local population, managing the NHS budget, and coordinating local providers of NHS services, such as hospitals and GP practices. Therefore, there is a greater potential that relevant factors other than pay vary on either side of these borders.

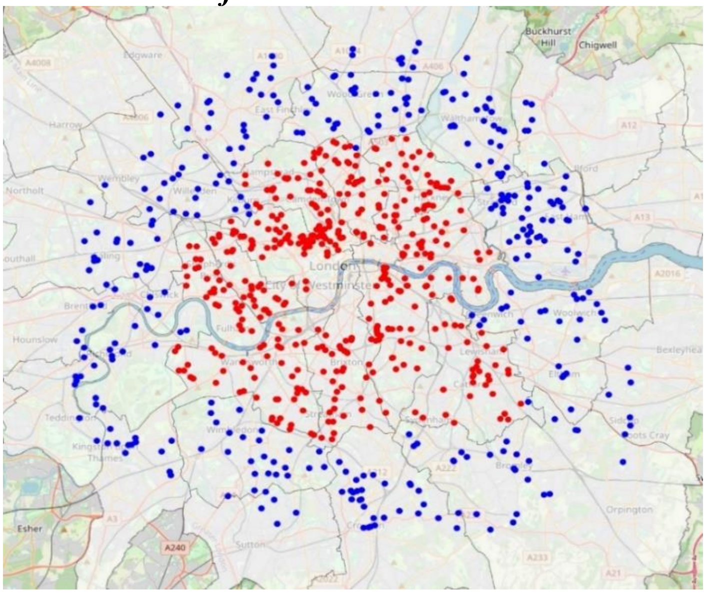
Within London, there are 5 ICBs, each covering approximately the same number of people and grouping together 5-8 boroughs. These are North Central London, North East London, North West London, South East London and South West London. Each ICB spans the inner-outer London border. We control for these in our analysis in order to hold constant differences in management or recruitment practices. In addition, we control for 5 geographical areas to ensure we are comparing NHS practices that are close together and well connected by transport links. These geographic segments cover all the area in our analysis and also span the inner-outer London border. The segments are West London, North London, East London, South East London, and South West London.[3]

The geographical boundary of inner London is defined in the NHS Terms and Conditions (NHS, 2023).[4] The locations of all NHS hospitals, general practitioner surgeries (equivalent to primary care physicians in the US) and clinics in inner and outer London are shown in Figure 1.

[3] These are defined according to postcode area, which groups parts of London that lie on the same mail delivery routes.

[4] The boroughs in inner London are Hammersmith, Fulham, Kensington & Chelsea, Westminster, Camden, Islington, City & Hackney, Tower Hamlets, Lambeth, Lewisham and Southwark. The boroughs in outer London are Brent, Ealing, Harrow, Hillingdon, Hounslow, Barnet, Enfield, Haringey, Barking & Dagenham, Havering, Newham, Redbridge, Waltham Forest, Bexley, Bromley, Greenwich, Croydon, Kingston, Richmond & Twickenham, Sutton & Merton.

Figure 1: The location of NHS Practices in Inner and Outer London



Note: Each dot represents an NHS practice, i.e., hospital, GP surgery or clinic. The red dots represent practices located in Inner London and blue are for those in Outer London. The map shows practices spanning 5 km on either side of the border.

To ensure fairness and non-discriminatory practice, the Agenda for Change system specifies 11 pay bands, ranging from Band 2 (the most junior) to Band 9 (the most senior).[5] The band assigned to a job is determined by the level of responsibility, required skills, and experience. Workers within each band have a predetermined period before they are eligible for a pay rise, based on their experience. The pay bands are shifted upwards by varying amounts each financial year, which runs from April to March, to accommodate inflation. The pay ranges of each band are set out in the NHS Handbook (2023), which provides clear guidelines for remuneration and progression within the band system.

3. Data

We use data from the UK Government's "Find a job" job posting website (https://www.gov.uk/find-a-job). This is a free service that is managed by the Department for Work and Pensions and is intended to facilitate efficient matching between employers and job

[5] Band 1 was closed for new applicants as of 1 December 2018. There are a total of 11 bands between 2 and 9 because Band 8 is split in to 4 sub-bands.

seekers, reducing the costs for both. Any UK company may advertise its vacant positions on "Find a job". The NHS jobs comprise about 40% of all adverts on the website.[6] Web scraping software was used to collect data on all open ads every week between July 2020 and February 2024. We collected information on the job title, pay, pay frequency, the dates the vacancy was posted and closed, the name and postcode of the employer, and the job description. An example job advert from "Find a job" is provided in Figure A1 in the Appendix.

A full list of NHS practices was obtained from the NHS website. The latitude and longitude of each practice was found using Google Maps and this was used to calculate the shortest linear distance from the practice to the inner-outer London boundary. The distance was then merged with the job vacancy data using the postcode.[7]

While our main RDD regression does not rely on the wage, we want to verify that the government mandated pay premium holds in our data and so we need to calculate the wage. An hourly wage is calculated for each vacancy using the information on pay and pay frequency. Almost all NHS job vacancies on "Find a job" report a pay range and in the analysis we use the mid-point of the pay range as our measure of the posted wage. The pay band of each vacancy was determined by comparing the posted annual salary with the exogenous minimum and maximum salaries associated with each band in a given financial year, after subtracting 4.35% from salaries for inner London vacancies (to account for the HCAS).

The duration of a posted vacancy is defined as the difference in days between the opening date of the vacancy and the closing date listed in the job ad. However, some ads are taken down before the published closing date, presumably because the position has been filled early. Conversely, some ads stay up past their published closing date, presumably because the

[6] A comparison of the number of NHS job ads at any point in time with the number of job ads listed on the NHS's own jobs website shows that around of 90% of NHS publicly listed vacancies are cross listed on "Find a job".
[7] Each postcode normally contains around 15 addresses and so represents a very narrowly defined area.

position remained unfilled at the time of the original closing date. Therefore, we adjust our duration calculation to account for ads that closed earlier or later than originally announced. In doing so we make use of the fact that we have a weekly snapshot of all open ads. We define early-closing ads as those where the last date the ad was observed is more than 7 days before the listed closing date. Similarly, we define late-closing ads as those where the last date the ad was observed is after the listed closing date. Since the ads were only scraped weekly we do not know the exact date these ads were taken down, only to within a range of a week. Therefore, for early- and late-closing ads, we adjust the vacancy duration to be equal to the difference between the opening date and the last observed date plus 3.5.[8]

In practice we find that only 2% of jobs close early while 4.5% close late. Therefore, most of the variation we observe in posting lengths is driven by practices anticipating the effects of their choice of duration at the time they post an ad, rather than by adjustments in duration after an ad is posted. In Section 6 we show that if we use duration in weeks rather than days (and therefore have an exact measure of duration for all ads) we get very similar estimates.

We drop ads which we observe as having been open for less than one week as we think these may be subject to measurement error and because we are unable to capture all similarly short-lived vacancies since the scraping takes place only once per week. We also exclude ads which are posted for more than 90 days as these are likely to be "phantom" vacancies which have been left up on the website.[9]

Each ad includes a job title. Unfortunately, these are not recorded in a sufficiently consistent fashion to allow us to compare vacancies on this basis. Therefore, we use the information in the job title to assign each vacancy to one of 14 occupations. These correspond

[8] We find that whether we added 3.5 or 7 or zero days to this duration measure has no real effect on our findings.
[9] The phenomenon of "phantom" vacancies occurs whereby vacancies that have already been filled have been left up as an advert erroneously (Albrecht et al. (2020)).

to the job categories on the NHS Health Careers website and are: Mental Well-being Professionals, Domestic Services, Estate Services, Nurses, Midwives, Administration, Support Services, Allied Health Professionals, Clinical Support Staff, Healthcare Support Staff, Pharmacists, Research and Science, Corporate Services, and Management. Similar to Adams-Prassl et al. (2022) we take a machine learning approach to assign job categories from the job title. Details are given in the appendix.

We restrict our sample to adverts that were posted between October 2021 and September 2023. We exclude data from before October 2021 to mitigate the distortionary effects of Covid-19 lockdowns and the resulting social distancing measures. We also drop ads posted after October 2023 so that we can determine whether the ad stayed up no longer than 90 days after the posting date. There are a total of 100,300 NHS vacancies in our sample which are within 5 km of the inner-outer London border and have no missing wages (wages are missing for about 12% of the sample). Restricting the sample to have a 2 km bandwidth on either side of the border (which we use in our baseline regressions) results in 38,784 observations. Finally, dropping observations with vacancy durations less than 7 or more than 90 days leaves us with a total of 36,435 observations.

Table 1 displays the summary statistics for the dataset showing the mean salary and duration for each job category overall and by inner and outer London. The average posted salary is £43,000 and the average duration of a vacancy is 20.84 days. The average salary is almost £500 higher in inner London than in outer London, possibly due to the HCAS premium but also due to differences in the distribution of jobs advertised. The average duration in inner London is 20.66 days which is lower than the average duration of 21.76 days in outer London. The most prevalent occupation is nurses, which makes up more than one third of our sample. The lowest paid occupation is administration while the highest paid is management.

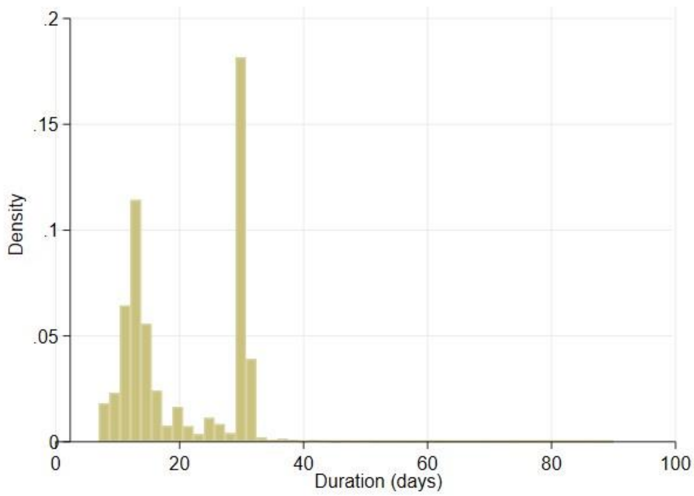
Table 1: Mean Salary and Duration by Occupation

	Inner			Outer			Total		
	N	Salary (£)	Duration (Days)	N	Salary (£)	Duration (Days)	N	Salary (£)	Duration (Days)
Administration	3586	31498	19.76	440	31295	20.89	4026	31476.11	19.89
Allied Health	2018	45235	20.51	491	43662	21.17	2509	44926.88	20.64
Clinical Support	958	44803	21.24	67	42034	20.40	1025	44622.21	21.18
Corporate	721	49968	20.73	77	49510	18.80	798	49924.21	20.54
Estates	763	37616	20.30	83	39644	20.89	846	37815.18	20.36
Healthcare Support	1683	33786	19.74	413	34065	20.88	2096	33840.64	19.97
Management	2542	53026	20.12	404	50426	20.82	2946	52669.63	20.21
Midwife	874	48217	20.77	18	44958	20.61	892	48151.20	20.77
Nurse	10953	42073	21.21	1505	40424	21.74	12458	41874.14	21.27
Pharmacists	1470	46216	20.49	232	43582	21.00	1702	45856.95	20.56
Public Health	151	43513	21.54	32	41767	14.27	183	43207.46	20.27
Research	679	46185	21.01	26	42622	25.19	705	46053.71	21.17
Therapy	4179	48464	20.66	2070	46062	22.82	6249	47668.33	21.37
Total	**30577**	**43003.27**	**20.66**	**5858**	**42526.22**	**21.76**	**36435**	**42926.57**	**20.84**

Note: This table shows the mean outcomes for variables in the vacancy data using jobs.gov.uk from 2021 to 2023.

Figure 2 shows a histogram of the duration of job vacancies in days. As noted earlier, the sample is truncated below 7 days. We see that the majority of vacancies are posted for less than one month with peaks at 13 and 30 days.[10]

Figure 2: Distribution of Vacancy Duration



Note: This histogram includes all ads in our estimation sample. The duration is calculated as the difference between the opening date and the published closing date of an ad (or the difference between the opening date and the last date the ad was observed plus 3.5 if the last observed date is more than a week earlier than or is later than the closing date).

[10] Our discussions with NHS administrators revealed that there are no "preset" posting lengths and no national guidance regarding advertising durations. Rather, the length of time an ad is posted is at the discretion of the practice that is hiring. The spikes seen in Figure 2 therefore seem to reflect a tendency to post ads for a standard period, either two weeks or one month.

4. Methodology

We use a local linear regression discontinuity design to estimate the causal effect of wages on vacancy duration. We include all vacancies within 2 km of the inner-outer London border. By selecting a relatively small bandwidth, we should ensure that there are no significant cross-border differences in any unobserved factors that influence vacancy duration. Consider two NHS practices situated immediately on either side of the inner-outer London boundary. Given their proximity to each other, it is reasonable to assume that these practices operate under very similar economic conditions and that factors such as the cost of living, labour market conditions, geographical mobility, and demand for NHS services do not vary significantly on either side of the border. In addition, London has excellent transport links, which mitigates geographical immobility and reduces the likelihood that transport costs play a significant part in restricting job seekers' search decisions. Therefore, if we find a difference in the average vacancy duration on either side of the border, we can interpret this as evidence that an exogenous increase in wages affects vacancy duration. Our regression specification is as follows:[11]

$$\ln D_i = \alpha Inner_i + \beta Dist_i + \gamma Inner_i \times Dist_i + \mathbf{X}_i \boldsymbol{\delta} + e_i \tag{1}$$

Here $\ln D$ is the log of the number of days vacancy i is posted, *Inner* is an indicator variable that takes the value of 1 if vacancy i is for a job in inner London and 0 if the vacancy is in outer London, and *Dist* is a continuous variable denoting the distance between the location of vacancy i and the border in kilometres (with distances on the inner London side of the border expressed as negative numbers). We also include a vector of controls $\mathbf{X}$, which includes month-by-year dummies, ICB dummies, geographic segment dummies, occupation dummies, pay

[11] This strategy is similar to the technique employed by Greaves and Sibieta (2019) who use HCAS zones in London to estimate the impact of teacher salaries on students' educational standards.

band dummies, dummies for the type of practice a job is with (specifically, whether the job is at a GP practice, hospital or clinic), and a set of dummies for job type (specifically, a dummy for whether the job is full-time or part-time and dummies for whether the job is permanent, temporary or an apprenticeship). We cluster the standard errors at the postcode level.

5. Baseline Results

The HCAS implies that an NHS position in inner London should pay 4.35% more than an identical position in outer London (since the former attracts a 20% premium and the latter a 15% premium compared to the national pay scales). To test whether this difference in pay across the border does indeed translate into an observed difference in posted wages in our sample, we regress the log of the posted wage on the same set of controls as in equation 1. The results of this regression are reported in column 1 of Table 2. This shows that NHS vacancies on the inner London side of the border have posted wages that are 4% higher on average than comparable vacancies on the outer London side of the border. This difference is not statistically different from the 4.35% premium that is mandated by law.

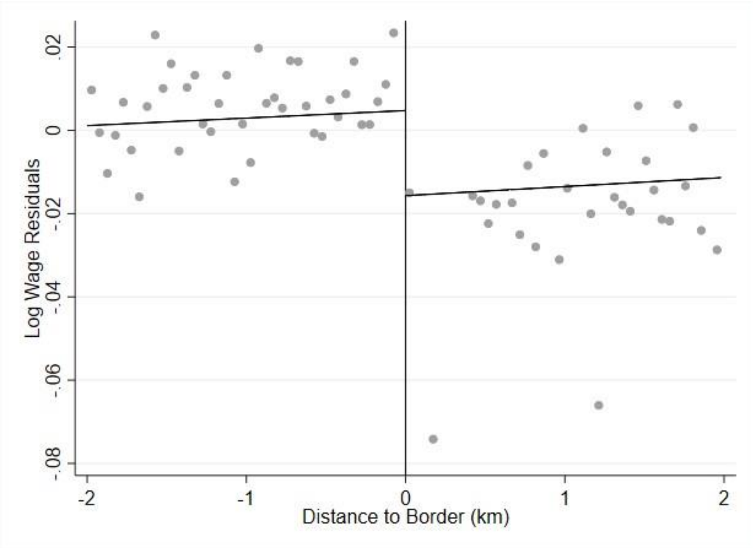
Table 2: The Effect of Inner London Premium on Log Wage and Log Duration

Variables	(1) Log (Wage)	(2) Log (Duration)	(3) Log (Duration)
Inner London	0.040***	-0.273***	-0.273***
	(0.002)	(0.100)	(0.054)
Distance	0.008***	-0.043	-0.114***
	(0.002)	(0.047)	(0.037)
Inner London × Distance	-0.007***	-0.094	0.095**
	(0.002)	(0.073)	(0.038)
Observations	36,435	36,435	36,435
R-squared	0.988	0.029	0.191
Month-by-Year FE	Yes	No	Yes
Integrated Care Board FE	Yes	No	Yes
Geographic Segment FE	Yes	No	Yes
Occupation FE	Yes	No	Yes
Pay Band FE	Yes	No	Yes
Practice Type FE	Yes	No	Yes
Job Type FE	Yes	No	Yes

Note: Standard errors are clustered by postcode and are reported in parentheses. *** $p<0.01$; ** $p<0.05$. A bandwidth of 2 km is used in all regressions. The Practice Type fixed effects include controls for whether the practice is a hospital, GP, or clinic. The Job Type fixed effects include controls for whether the job is permanent, temporary or an apprenticeship, and whether it is full-time or part-time.

This regression is shown graphically in Figure 3 which further confirms that there is a discontinuity in pay at the inner-outer London border. In this figure, we regress log wages on all the controls in **X**, take the residuals and average them into 40 equally spaced bins on either side of the border (i.e. 50-metre bins). We see that the posted wages barely change as one moves away from central London. However, we do see a sharp decrease in wages at the border going from inner to outer London (that is, from left to right).

Figure 3: The Effect of the London Pay Premium on Log Wages



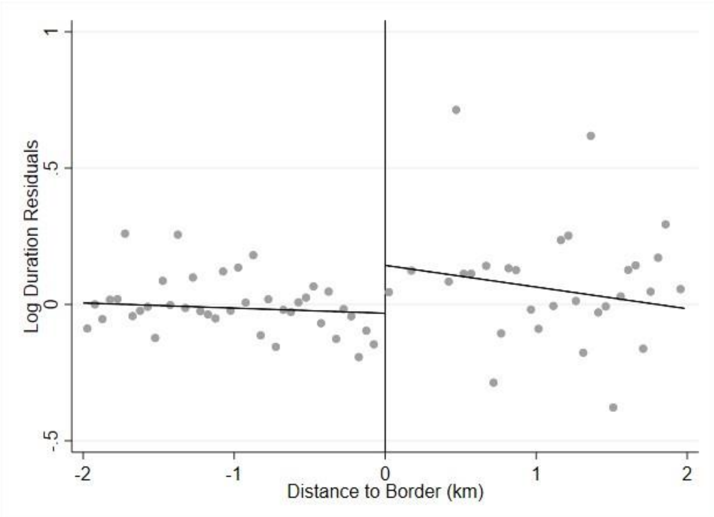
Note: Each marker represents the average residual from a regression of log wages on our set of controls reported in equation 1, within 40 equally spaced bins on either side of the border. The lines depict the lines of best fit, separately on each side of the border.

Next, we examine how the London pay premium affects the length of time a vacancy is posted. In column 2 of Table 2 we estimate equation 1, leaving out the vector of control variables, **X**. This shows that the duration of vacancies in inner London is 27.3% lower than the duration of vacancies in outer London. When we add our full set of controls in column 3, this coefficient does not change at all but the standard error almost halves. Given that the government mandated pay premium is 4.35%, this implies a duration-wage elasticity of -6.3. Since the average duration of a job vacancy in outer London is 21.76 days, this means that the inner London pay premium reduces vacancy durations on average by about 6 days.

Our estimated elasticity is much higher than those of Faberman and Menzio (2018) and Mueller et al. (2024). However, it is similar to the estimates found by Bassier et al. (2023), who use variation in discrete wage changes brought about by firm-wide revisions in pay levels and find an elasticity between -3 and -5. Bassier et al. (2023) rely on patterns in the wage data and information from a secondary wage settlement dataset in order to identify exogenous wage changes. In contrast, we rely on a discontinuity arising from a well-documented organizational pay policy, which should minimize the potential for measurement error. Therefore, this could explain why our estimates are even larger than those in Bassier et al. (2023).

To depict our estimates graphically, we follow the same approach as in Figure 3 and take the residuals from a regression of log duration on the controls in **X**, average these within 40 equally spaced bins on each side of the border, and plot these against distance to the border. Figure 4 shows that the duration of a vacancy does not change much as one moves from inner London towards the border (that is, from left to right) but there exists a jump in duration at the border such that job vacancies throughout outer London have higher durations that those in inner London.

Figure 4: The Effect of the London Pay Premium on Log Duration



Note: Each marker represents the average residual from a regression of log wages on month, area, occupation, band and job type fixed effects, within 40 equally spaced bins on either side of the border. The lines depict the lines of best fit, separately on each side of the border.

6. Robustness Checks

Although our estimates so far control for the most likely determinants of vacancy duration that vary across the inner-outer London border, it is possible that they may still be driven by various modelling choices, including the length of the bandwidth or our definition of vacancy duration. In addition, the effects we find may be driven by underlying geographic patterns of vacancy duration within London, so that dividing any part of London into an inner and outer part would lead to the same estimates. To address these concerns, in this section we show what happens to our estimates when we use various bandwidths and use a different measure of vacancy duration. We also show what happens to the estimates when we replace the true inner-outer London border with fictitious borders that mimic the inner-outer London region border.

Adjusting the Bandwidth

There is a trade off in any bandwidth adjustment as choosing a small bandwidth means that there will be fewer observations in our sample and therefore more noise in our estimates. Conversely, a higher bandwidth may serve to reduce variance but will also introduce bias as it is more likely that there will be unobserved determinants of vacancy length that vary from one side of the border to the other.

Table 3 shows the estimates using different bandwidths, from 1 km to 3 km. We see there is a clear pattern with smaller bandwidths resulting in larger coefficients. With a bandwidth of 1 km the estimate is -0.40, while with a bandwidth of 3km the estimate falls to -0.18. The standard errors are much larger for the 1 km bandwidth than for the other bandwidths. However, the coefficient estimates are significant in all cases. The number of observations drops off significantly once the bandwidth is reduced below 2 km. As we are interested in examining heterogeneity in the coefficient by occupation, we think that 2 km is a justifiable choice of bandwidth.[12]

[12] Greaves and Sibieta (2019) also use a bandwidth of 2 km when they study variation in teacher salary across the outer and fringe London borders. We attempted to use the optimal bandwidth adjustment method of Calonico et al. (2019) but the suggested bandwidth was 200 metres which was not feasible for our analysis.

Table 3: The Effect of Varying the Bandwidth

Variables	(1) 1 km	(2) 1.5 km	(3) 2 km	(4) 2.5 km	(5) 3 km
Inner	-0.404***	-0.368***	-0.273***	-0.240***	-0.183***
	(0.140)	(0.057)	(0.054)	(0.053)	(0.057)
Distance	-0.172	-0.213***	-0.114***	-0.075***	-0.036
	(0.172)	(0.055)	(0.037)	(0.029)	(0.026)
Inner × Distance	0.000	0.161**	0.095**	0.053*	0.019
	(0.171)	(0.075)	(0.038)	(0.032)	(0.026)
Observations	11,706	21,589	36,435	40,637	48,031
R-squared	0.214	0.197	0.191	0.186	0.191

Note: Standard errors are clustered by postcode and are reported in parentheses. *** p<0.01; ** p<0.05; * p<0.10. In all columns, we also include fixed effects for Month-by-Year, Integrated Care Board, Geographic Segment, Occupation, Pay Band, Practice Type, and Job Type. The Practice Type fixed effects include controls for whether the practice is a hospital, GP, or clinic. The Job Type fixed effects include controls for whether the job is permanent, temporary or an apprenticeship, and whether it is full-time or part-time.

Alternative Measure of Vacancy Duration

We collected data on vacancies once per week. Therefore, we had to approximate the exact day an ad was taken down if a vacancy was closed earlier or later than the pre-specified closing date. To examine whether this affected our results, we repeated our baseline regression using vacancy duration measured in weeks rather than days. Duration in weeks is measured exactly, even for ads that closed early or late. Table 4 shows that when we use log of vacancy duration in weeks the coefficient is -0.26 (with or without controls). This implies that vacancies in inner London take on average 26% fewer weeks to fill. Given that the average number of weeks a vacancy is posted is 3.17, this corresponds to an effect of 5.7 days per week, which is very similar to our baseline estimate of 5.9 days. Using weeks also allows us to directly compare our estimates with those of Bassier et al. (2023), who also measure duration in weeks. The estimate in Table 4 implies an elasticity of -5.9, compared to a range of -3 to -5 in Bassier et al. (2023).

Table 4: The Effect of Inner London Premium on Log Wage and Log Duration in Weeks

Variables	(1) Log (Duration in Weeks)	(2) Log (Duration in Weeks)
Inner	-0.257**	-0.258***
	(0.099)	(0.052)
Distance	-0.036	-0.108***
	(0.051)	(0.037)
Inner × Distance	-0.091	0.093**
	(0.074)	(0.037)
Observations	36,435	36,435
R-squared	0.021	0.132
Month-by-Year FE	No	Yes
Integrated Care Board FE	No	Yes
Geographic Segment FE	No	Yes
Occupation FE	No	Yes
Pay Band FE	No	Yes
Practice Type FE	No	Yes
Job Type FE	No	Yes

Note: Standard errors are clustered by postcode and are reported in parentheses. *** p<0.01; ** p<0.05. A bandwidth of 2 km is used in all regressions. The Practice Type fixed effects include controls for whether the practice is a hospital, GP, or clinic. The Job Type fixed effects include controls for whether the job is permanent, temporary or an apprenticeship, and whether it is full-time or part-time.

Placebo Test using the Border Location

Next, we present the results of a placebo test to check whether the duration elasticities are truly driven by comparing vacancies on either side of the inner-outer London border or whether they reflect a general geographical pattern in London, unrelated to the inner London pay premium. We create artificial borders by moving the HCAS border towards or away from the centre of London by varying distances, maintaining a 2 km bandwidth each time. To avoid comparing observations that straddle the true inner-outer London border, we move the border by at least 2.5 km in either direction. Given a bandwidth of 2 km, this means that when we shift the border inwards toward London we only use inner London observations and vice versa. At these newly-defined borders there are no government mandated HCAS premiums and so if we find a discontinuity in vacancy duration using these fictitious borders, it is evidence that the changes in vacancy duration reported earlier are likely driven by some factor other than wages.

Table 5 below shows the estimated coefficients for the artificial borders. In columns 1-3, we shift the border inwards (i.e. towards central London) by 2.5 km, 3 km and 4 km, respectively. In columns 4-6, we shift the border outwards by the same distances. Reassuringly, using a border that does not correspond to the true government defined border between inner and outer London gives non-statistically significant estimates, as expected.

Table 5: Placebo Test by Shifting the Border Inwards and Outwards

Variables	(1) -4 km	(2) -3 km	(3) -2.5 km	(4) +2.5 km	(5) +3 km	(6) +4 km
Inner	-0.040	-0.053	-0.051	-0.045	-0.157	0.094
	(0.067)	(0.057)	(0.049)	(0.052)	(0.101)	(0.062)
Distance	0.007	0.059	0.077**	-0.038	-0.097**	0.022
	(0.112)	(0.037)	(0.032)	(0.025)	(0.047)	(0.040)
Inner x Distance	0.066	-0.075	-0.171***	0.076	0.115*	0.183
	(0.115)	(0.051)	(0.047)	(0.069)	(0.068)	(0.131)
Observations	24,074	28,012	26,567	47,279	51,809	31,609
R-squared	0.172	0.167	0.171	0.144	0.164	0.196

Robust standard errors clustered by postcode in parentheses. *** $p<0.01$; ** $p<0.05$; * $p<0.10$. Positive distances mean the border is shifted away from central London; negative distances mean the border is shifted towards central London. Bandwidth of 2 km is used in all regressions. In all columns, we also include fixed effects for Month-by-Year, Integrated Care Board, Geographic Segment, Occupation, Pay Band, Practice Type, and Job Type. The Practice Type fixed effects include controls for whether the practice is a hospital, GP, or clinic. The Job Type fixed effects include controls for whether the job is permanent, temporary or an apprenticeship, and whether it is full-time or part-time.

7. Heterogeneous Effects by Occupation

The NHS hires many different types of workers ranging from cleaners to therapists and so there may be heterogeneity in the vacancy elasticity according to occupation. It may be that vacancies for lower skilled jobs paying lower wages respond differently to an increase in wage than vacancies for higher skilled jobs which already have a much higher base wage. In addition, it is possible that those occupations for which the NHS is the dominant employer, such as nurses, may have different elasticities than occupations for which there are many outside options. To check if this is the case, we take all occupation categories with at least 2,000 observations in our sample and estimate equation 1 separately for each. Table 6 reports the coefficient on the inner London dummy variable in each case.

Interestingly, we see that the effects are very similar across all job categories. The largest effect is for vacancies in administrative jobs and the smallest is for management vacancies, but the differences are minimal. The lack of variation in elasticities across the different job categories suggests that the differing availability of outside options does not affect how sensitive vacancies are to wages.

Table 6: The Effect of the London Premium on Vacancy Duration by Occupation

Variables	(1) Allied Health Professionals	(2) Management	(3) Nurse	(4) Administration	(5) Healthcare Support Staff	(6) Therapy
Inner	-0.231**	-0.212**	-0.283***	-0.323***	-0.243***	-0.221***
	(0.101)	(0.091)	(0.064)	(0.047)	(0.075)	(0.076)
R-Squared	0.218	0.211	0.182	0.247	0.195	0.221
Observations	2,509	2,946	12,458	4,026	2,096	6,249

Note: Standard errors are clustered by postcode and are reported in parentheses. *** p<0.01; ** p<0.05. A bandwidth of 2 km is used in all regressions. In all columns, we also include distance, inner London × distance, fixed effects for Month-by-Year, Integrated Care Board, Geographic Segment, Occupation, Pay Band, Practice Type, and Job Type. The Practice Type fixed effects include controls for whether the practice is a hospital, GP, or clinic. The Job Type fixed effects include controls for whether the job is permanent, temporary or an apprenticeship, and whether it is full-time or part-time.

8. Policy Implications

The results so far show that a wage increase of 4.35 percent leads to a reduction in vacancy duration by 6 days per week. But what does this imply for government costs and for health outcomes? While we cannot speak directly to the effect of paying more to NHS workers on the health of the population, there is evidence from Propper and Van Reenen (2010) that a 10 percent higher outside wage is associated with a 7 percent increase in death rates. They attribute part of this relationship to higher outside wages leading to recruitment difficulties as higher quality workers are attracted to the higher paying non-NHS jobs.

Moreover, recent work by Kelly et al. (2022) suggest that labour shortages of key workers weigh on the ability of the NHS to deliver effective health care which manifests itself in poorer health outcomes and longer waiting list times for patients. Buchan (2019) also suggests that labour shortages may increase physical and psychological pressures on the remaining workforce which could further reduce the ability of the NHS to retain staff, exacerbating the aforementioned challenges over time. Given our finding that higher wages result in lower vacancy durations which should in turn ease recruitment difficulties, these studies also suggest that higher wages should contribute to better health outcomes.

The NHS Pay Review Body, a quasi-independent board which advises the government on issues relating to NHS worker pay, acknowledges that increasing wages could act to improve staff retention rates and improve wellbeing (NHS Pay Review Body, 2022). Therefore, the Body consistently recommends pay increases on the basis that they act as a mechanism to reduce vacancy durations and labour shortages, at least in the short run. Our study provides support for this recommendation.

Our analysis is particularly pertinent for the issue of nursing shortages in the UK. Due to the chronic shortage of nurses, the NHS frequently hires agency workers to fill vacant positions. The cost of agency nurses is substantially higher than the cost of regular nurses. An investigation by the Royal College of Nursing (RCN) found that NHS trusts in England alone spent over £3 billion on agency staff between 2020 and 2022 and that the figure for London was £630 million (Nursing Times, 2023). The RCN argued that this large expenditure was mainly attributable to a lack of funding for the NHS. Our study suggests that increasing pay by 4.35% reduces vacancy durations by almost 1 week. Thus, while it is costly to pay nurses more, the direct financial benefits of reduced agency costs and indirect benefits of improved health outcomes may make this a worthwhile endeavour.

9. Conclusion

We examine the relationship between wages and vacancy duration, exploiting a spatial discontinuity in pay in the UK's National Health Service. Due to government regulations, workers in inner London must be paid 4.35% more than workers in outer London. We use a sharp regression discontinuity design to compare the length of time for which vacancies are advertised on either side of the inner-outer London border.

We find that the wage premium in inner London reduces vacancy duration by almost a week. Our estimates imply that the elasticity of vacancy duration with respect to wages is -6.3, which is considerably larger than most previous estimates. We find similar estimates for all occupation groups hired by the NHS, including administrative roles, where the NHS competes for workers with private sector employers. Our results suggest that firms can use wages to help fill vacancies quickly. The results also provide evidence that increasing the wages offered to NHS workers may be a way to tackle the persistent NHS staff shortages as well as reducing the amount of money spent on agency staff and contributing to better health outcomes.

References

Abi Adams-Prassl, Maria Balgova, and Matthias Qian., (2020). Flexible work arrangements in low wage jobs: Evidence from job vacancy data. IZA discussion paper 13691.

Albrecht, J., Decreuse, B. and Vroman, S. (2020), "Directed Search and Phantom Vacancies" (Discussion Paper Series 13704, IZA September).

Banfi, S. and Villena-Roldan, B., (2019). Do high-wage jobs attract more applicants? directed search evidence from the online labor market. *Journal of Labor Economics*, 37(3), pp.715–746.

Bassier, I., Manning, A. and Petrongolo, B., (2023). Vacancy duration and wages. IZA discussion Papers, (16371).

Buchan, J., A. Charlesworth, B. Gershlick and I. Seccombe, (2019). A critical moment: NHS staffing trends, retention and attrition. Available: https://www.health.org.uk/sites/default/files/upload/publications/2019/A%20Critical%20Moment_1.pdf

Calonico, Sebastian & Cattaneo, Matias & Farrell, Max. (2019). Optimal Bandwidth Choice for Robust Bias Corrected Inference in Regression Discontinuity Designs1. The Econometrics Journal. 23. 10.1093/ectj/utz022.

Eeckhout, J. and Kircher, P., 2010. Sorting and decentralized price competition. *Econometrica* 78(2), pp.539–574.

Faberman, R.J. and Menzio, G., (2018). Evidence on the relationship between recruiting and the starting wage. *Labour Economics*, 50, pp.67–79.

Greaves, Ellen & Sibieta, Luke, (2019). "Constrained optimisation? Teacher salaries, school resources and student achievement," *Economics of Education Review*, Elsevier, vol. 73(C).

Kelly, E., Stoye, G. and Warner, M., (2022). Factors associated with staff retention in the NHS acute sector [Online]. Institute for Fiscal Studies. Available from: https://doi.org/10.1920/re.ifs.2022.0216.

Kuhn, P. and Mansour, H., 2014. Is internet job search still ineffective? The economic journal, 124(581),pp.1213–1233.

Lindholm, A., Wahlström, N., Lindsten, F. & Schön, T. B. (2022), Machine Learning - A First Course for Engineers and Scientists, Cambridge University Press. Available from: https://smlbook.org [Accessed 15th February 2023].

Marinescu, I. and Wolthoff, R., 2020 Opening the Black Box of the Matching Function: The Power of Words. *Journal of Labor Economics* 2020 38:2, 535-568

Manning, Alan. (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press, 2003. https://doi.org/10.2307/j.ctt5hhpvk.

Morgan, B., 2022. NHS staffing shortages. The kings fund [Online]. Available from: https://www.kingsfund.org.uk/publications/nhs-staffing-shortages.

Mueller, Andreas, Damian Osterwalder, Josef Zweimüller, and Andreas Kettemann. (2024). Vacancy Durations and Entry Wages: Evidence from Linked Vacancy–Employer–Employee Data, *The Review of Economic Studies*, Volume 91, Issue 3, May 2024, Pages 1807–1841

NHS, 2023. NHS terms and conditions of service handbook. Available from: https://www.nhsemployers.org/publications/tchandbook.

NHS Pay Review Body, 2022. NHS pay review body thirty-fifth report 2022. Available from:www.gov.uk/government/publications/nhs-pay-review-body-thirty-sixth-report-2023.

Nursing Time, 2023. New RCN analysis reveals cost of short-term approach to filling staffing gaps. Available here: https://www.rcn.org.uk/news-and-events/news/uk-32bn-agency-spend-could-have-paid-salaries-of-31000-nurses-051223#:~:text=Findings%20from%20a%20Freedom%20of,shortages%20to%20keep%20wards%20open.

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, no. 2 (2010): 222–73. https://doi.org/10.1086/653137.

Shields, M.A., (2004). Addressing nurse shortages: what can policy makers learn from the econometric evidence on nurse labour supply? *Economic Journal* 114(499), November, pp.464–498.

Shimer, R., (2005). The assignment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy*, 113(5), pp.996–1025.

Statista, 2022. The world's biggest employers. Available from: https://www.statista.com/chart/3585/the-worlds-biggest-employers/.

The Health Foundation, (2020). 2019 annual report and financial statements. Available from: https://www.health.org.uk/news-and-comment/charts-and-infographics/how-reliant-is-the-nhs-in-england-on-international-nurse-recruitment.

Wolthoff, R., (2018). Applications and interviews: Firms' recruiting decisions in a frictional labour market. *The Review of Economic Studies*, 85(2), pp.1314–1351.

Appendix

Figure A1: Job Vacancy Advertisement

Clinical Research Nurse - Critical Care | Guy's and St Thomas' NHS Foundation Trust

Posting date:	08 August 2023
Salary:	Not specified
Additional salary information:	£42,471 - £50,364 p.a. inclusive of HCA
Hours:	Full time
Closing date:	07 September 2023
Location:	London, SE1 7EH
Company:	Guys and St Thomas NHS Foundation Trust
Job type:	Permanent
Job reference:	5424427/196-BRC1992

Summary

Are you looking for a new challenge to progress you career within Research nursing? Do you think you could contribute to a busy and dynamic clinical research team? We are looking for a dynamic and highly motivated nurse to join our friendly Critical Care research team. You should have excellent communication and interpersonal skills and have the ability and initiative to work independently. You should be able to manage your workload flexibly across a range of projects and deadlines in a fast-pace working environment.

You will be an essential member of the Critical Care research team, leading a growing portfolio of observational research and clinical trials of investigational medicines & devices within this specialty. The role will include supporting the delivery of commercial and non-commercial studies. You will work closely with a dynamic multi-disciplinary research team to deliver our portfolio of studies. You will be required to liaise with trial site coordinators for site initiation visits, site visits and site closures as appropriate. You will be required to attend research meetings on a regular basis.

There are excellent opportunities for further education and research development as part of these roles, with access to CPD funding and revalidation support. Opportunities for developing new research and clinical skills are available and actively encouraged.

Note: This figure shows an example of a job vacancy posted on the government find a job website

NHS Job Categories Description

In this section, we explain the roles of each of the job categories that we use in the estimation. Jobs are assigned to particular categories according to their job title and job description.

- *Mental Well-being Professionals* focus on providing mental health and wellbeing support to patients. They include Psychologists, Psychiatrists, Psychotherapists, and Counsellors who work with individuals experiencing mental health issues.
- *Domestic Services* categorise staff that are responsible for maintaining cleanliness and hygiene in healthcare facilities. They ensure that hospitals and clinics are kept clean, sanitised, and well-maintained to provide a safe environment for patients and staff.
- *Estate services* staff manage and maintain the physical infrastructure of healthcare facilities, such as hospitals and clinics. They handle tasks related to building maintenance, repair, security, and logistics.

- *Nurses* are responsible for providing medical care, administering medications, monitoring patients' conditions, and supporting doctors and other healthcare professionals in various settings.
- *Midwives* specialise in providing care to expectant mothers before, during, and after childbirth. They assist with prenatal care, childbirth, and postnatal support to ensure the well-being of both the mother and the baby.
- *Administrators* handle tasks related to scheduling appointments, managing patient records, billing, and other administrative duties.
- *Support Services Staff* assist in various areas of patient care and hospital operations, providing help with non-medical tasks like patient transportation, portering, and general assistance. This category also includes Technical Engineers and Information Technology (IT) Assistants.
- *Allied Health Professionals* encompasses a diverse group of specialised professionals who support patient care in various ways, such as Physiotherapists, Occupational Therapists, radiographers, and audiologists.
- *Clinical Support Staff* assist Allied Health Professionals in delivering patient care. They may include Medical Assistants, Phlebotomists, and other roles that provide support during medical procedures and tests.
- *Healthcare Support Staff* assist patients. This category may include Healthcare Assistants, Nursing Assistants, and other roles that support patients' daily needs.
- *Pharmacists* are responsible for dispensing medications, providing medication-related advice to patients and healthcare professionals, and ensuring the safe and appropriate use of drugs.

- *Research and Science* involves professionals engaged in medical and scientific research, including clinical researchers, laboratory technicians, and scientists. They contribute to advancing medical knowledge and developing new treatments.
- *Corporate Services and Management* handle various administrative, financial, and strategic functions within the NHS. They oversee policy development, resource allocation, and overall management of healthcare organisations.

Machine Learning Approach for Job Titles

In order to assign each job vacancy to a particular job category, we follow, Adams-Prassl et al. (2022) and take a machine learning approach. Specifically, we take the following steps:

- Manually assign job vacancies to the defined job categories according to key words and phrases included in the job title.
- Split the manually assigned job vacancies into an 80\20 train\test split.
- Train a machine learning model on the manually assigned data to classify job vacancies that cannot be assigned manually to each job category.
- Run the model on the job vacancies that were not assigned manually.

We label job vacancies manually using key terms in the job title of each vacancy. We use the *strpos* command in Stata to assign job vacancies to the job categories. Given the broad scope of job titles in the NHS, manual assignment only accounts for approximately 85% of the observations. Therefore, we employ a machine learning model to assign the job vacancies that are not matched manually to the predefined job category, using the job description. We divide the manually assigned job vacancies, along with their corresponding job description, into a 80\20 train\test split. We then train a Bayesian classifier machine learning text classification model on 80% of the manually assigned job vacancies.

We employ a Multinominal Logistic Regression (MNLR) text classification model for the purpose of assigning job descriptions to job categories. Consider two types of job category $k \in \{1,2\}$. We denote y_k as a binary indicator that takes a value of 1 if the job category is k and 0 otherwise. The MNLR model starts by vectorising each job description, $d \in D$, into individual words. Defining n as the number of unique words that are collected following vectorisation, then $\boldsymbol{x_d} = (x_{1d}, .., x_{id}, .., x_{nd})$ is a vector containing the frequency of occurrence for every word in a given job description d. Note that we exclude "stop words" such as "the" and "and" to improve the model's predictive power. This model is interested in determining the probability that a given job vacancy is classified by a particular job category k, given the collection of words in the job description x_d, $p(y_k|x_d)$. The logistic regression model is linear in log odds such that:

$$\log \frac{p(y_1 = 1|\boldsymbol{x_d})}{p(y_1 = 0|\boldsymbol{x_d}).} = \boldsymbol{x'_d \beta_d}$$

This can be rewritten as the logistic regression function:

$$g(\boldsymbol{x}) = p(y_k = 1|\boldsymbol{x_d}) = \frac{e^{\boldsymbol{x'_d \beta_d}}}{1 + e^{\boldsymbol{x'_d \beta}}}$$

The estimated coefficients $\hat{\beta}$ are calculated via maximum likelihood using the manually assigned training data. This approach can be generalised to consider the case where there are $K > 2$ job categories such that $k \in \{1, .., K\}$. In this case define a set of parameters $\boldsymbol{\beta_k}$ for each job category k. Lindholm et al. (2022) explain that, when using the softmax function, the individual class probabilities $p(y = k|\boldsymbol{x})$ can be expressed as:

$$g_k(\boldsymbol{x}) = p(y = k|\boldsymbol{x}) = \frac{e^{\boldsymbol{x'_d \beta_k}}}{\sum_{j=1}^{K} 1 + e^{\boldsymbol{x'_d \beta_j}}}$$

The MNLR then requires a decision rule that assigns classes to a particular job vacancy. A common approach is to simply assign a class to the job vacancy which has the highest probability of being correct, given the vectorised words in a given job description $\boldsymbol{x_d}$. After being trained on the 80% split of the manually assigned data, the model is then tested on the remaining 20% of the remaining manually assigned data. Following Adams-Prassl et al. (2020), we report three statistics, namely the Precision, Recall and F1 score that is derived from the tested data. Noting that TP, FP and FN denote the number of true positives, true negatives and false negatives respectively, the equations for each of these metrics is given below:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The respective accuracy, precision and F1 score for the test data is 0.87, 0.76, 0.81 respectively. Once the model has been trained and tested, the MNLR is used to classify the jobs, that were not assigned manually, according to the decision rule.