# Team Composition and the Returns to Human Capital: Evidence from Nursing Teams

Elaine Kelly, Carol Propper and Ben Zaranko*

June 2023

**Abstract**

This paper examines the role of team production in healthcare, a sector where collaboration, communication and knowledge sharing are crucial. Employing novel high-frequency data we examine the impact of the size and human capital of hospital nursing teams on inpatient mortality. We find significant adverse mortality impacts of shortages of nurses with higher levels of general and firm-specific human capital but no clear returns to location- or team-specific experience. Adverse impacts of nurse shortages are particularly concentrated among patients with sepsis, a condition where early detection is key and nurses play a direct role in monitoring and assessing patients.

*JEL classification: I11, J24, J45, M50.*

*The research was enabled by the Imperial Clinical Analytics Research and Evaluation (iCARE) environment and used the iCARE team and data resources. Data management was provided by the Big Data and Analytical Unit (BDAU) at the Institute of Global Health Innovation (IGHI). This paper has been screened to ensure no confidential information is revealed. All authors gratefully acknowledge financial support from the UK Economic and Social Research Council (ESRC) through the Centre for the Microeconomic Analysis of Public Policy (CPP) at the Institute for Fiscal Studies (ES/M010147/1) and "The impact of medical labour on variation in patient outcomes: evidence from English public hospitals" research grant (ES/S003118/1). Propper also acknowledges funding from ERC Advanced Investigator Grant Healthcare Labour. We would like to thank Eric French, Imran Rasul and seminar participants at the Institute for Fiscal Studies, University of York, Monash University, University of Sydney, Essen Health Conference 2022, European Association of Labour Economists 2022 Conference, European Health Economics Association 2022 Conference, the Royal Economic Society conference 2022, and the Allied Social Science Associations 2023 meeting for their helpful comments. Any errors are the authors' own. Author affiliations and contacts: Kelly (Health Foundation and Institute for Fiscal Studies, elaine.kelly@health.org.uk); Propper (Imperial College Business School, Monash University and Institute for Fiscal Studies, c.propper@imperial.ac.uk, corresponding author); Zaranko (Institute for Fiscal Studies, ben.zaranko@ifs.org.uk).

# 1 Introduction

Teams play a key role in the organisation of work and exist to exploit gains from specialisation, knowledge sharing, and complementarities between workers. They often contain individuals with different skills and knowledge and at different levels of the firm hierarchy (Lazear and Shaw, 2007). But while team output depends on all inputs, in many settings the output of individual inputs cannot be directly measured. And despite the pervasiveness of teams and team production in the workplace (Wuchty et al., 2007; Deloitte, 2016, 2021) studies of how team composition affects team output remain surprisingly sparse (Ali et al., 2021).

This paper addresses this issue in a setting where the output is collective and of considerable societal importance, healthcare production. We examine nursing teams in hospitals and use patient mortality, a stark but frequently used measure of the quality of production, as our measure of output.[1] Nursing teams play a critical role in the delivery of healthcare in hospitals, with responsibility for monitoring patients, implementing and adjusting treatment plans and the co-ordination of overall care. The tasks that need to be undertaken vary in the number of people required and the skills required. The composition of the team in terms of its human capital, as well as its size, is therefore likely to be an important determinant of the quality of patient care.

We examine the impact on the probability of a patient death of the absence of individuals with different levels and types of human capital, different levels of seniority, and different levels of familiarity with the firm (hospital), plant (hospital ward) and team (co-workers). We identify this by exploiting daily shortfalls from target staffing levels that are determined several months in advance but may not be filled on the day, and from the rotation of staff into teams at the daily level that arises because care needs to be provided 24/7.[2] This allows us to study the effect of an absence of a more senior or more qualified staff member, compared to that of a more junior or less qualified staff member, as well as the absence of a staff member who is more familiar with her employer, her physical location or her fellow team members than one who is less familiar. Our setting allows us to abstract away differences in financial incentives as in the context we examine, the English National Health Service, pay is determined exogenously to the

[1] Mortality is a commonly used as a proxy for the quality of nursing care (e.g. Needleman et al. (2011); Griffiths et al. (2019)) where potential mechanisms include missed care (Ball et al., 2014, 2018), delayed monitoring of vital signs (Redfern et al., 2019), and failure to respond to deterioration in a patient's condition (Smith et al., 2020).

[2] This variation in size is different from contexts where the size of the team is fixed and only the composition changes such as professional basketball (Arcidiacono et al., 2017) or two-person physician teams within hospitals (Chen, 2021).

organisation we study, is unrelated to output, and individuals of the same grade and contract type are paid the same amount, regardless of their contribution to the output of their team.

Our data, from one large hospital in 2017, utilises around 19,000 teams defined at the 24 hour-ward level, containing just under 4,500 unique staff, responsible for 44,000 patients and providing 3.5 million hours of patient care. We control for differences in patient severity and differences across hospital sites, wards and days. Our identification comes from within-ward variation in staffing levels and composition across days and we explicitly address concerns that team staffing, despite being set 3 months in advance, is endogenous to mortality, patient severity and the number of patients that the team is treating. We undertake a large set of tests to examine these key assumptions and find no evidence of such endogeneity. We find no clear relationship between staffing levels and observable patient severity, which indicates that staffing does not respond to day-to-day variation in patient mix. We show that deaths on previous days have no clear impact on either subsequent staffing levels or subsequent staff sickness absences, indicating that staffing does not respond to adverse events. We find no evidence that low staffing leads to an outflow of patients from the ward, nor that sickness absences lead to such an outflow, suggesting that responses to staffing shocks do not lead to patient reallocation across wards. We show that average patient severity does not depend on staffing levels. We also find no serial correlation in patient deaths across days, indicating that our observable measures of patient severity do a good job of controlling for the probability of patient mortality.

We find that both the size and composition of the nursing team matters for patient mortality. For the average team the absence of a nurse with university degree-equivalent level training increases the odds of a patient death by approximately 10%. In contrast, the lack of one nursing assistant, who does not have the formal training of a registered nurse, has no impact on patient mortality. Within nurses with degree level training, the effect of absence on mortality increases with seniority. We also find evidence of substantial returns to firm-specific experience. Conditional on the quantity and skill-mix of the team, increasing the average firm-specific experience among degree qualified nurses by one year reduces the odds of a patient death by 7.2%, the equivalent to adding two-thirds of an extra nurse to the team. However, there is again no evidence of such an effect for nursing assistants. We find little returns to location-specific or team-specific experience, and no evidence of lagged effects of short-staffing or spillovers onto other teams.

Our results are driven by mortality impacts on patients of relatively low, rather than

high, clinical severity. The largest impacts are for patients with a diagnosis of sepsis. This is a condition which can develop suddenly and where early detection is of critical importance for survival, and for which nurses have a central and direct role in detection of the early signs and its subsequent control. This suggests that less intensive patient monitoring and assessment by teams that are lacking degree-qualified nurses is likely to be a key mechanism driving our results.

We contribute to three literatures. First, we contribute to the literature on the relationship between team composition and team productivity. In education, Herrmann and Rockoff (2012) show that less experienced teachers had less impact on pupil learning. In sport, Stuart (2017) evaluates the team productivity effects of injuries of different types of hockey players, Fischer et al. (2021) investigate the productivity effects of COVID-19 on the performance of top league soccer players in Germany and Italy, and Hoey et al. (2022) examine the effect of absences of hockey players on team output. In the context of patent production, Jaravel et al. (2018) show that the premature death of an inventor significantly lowers the subsequent innovation and earnings of their co-inventors. In healthcare Agha et al. (2021) show that teams composed of a smaller number of specialists have better outcomes and Chen (2021) finds that higher levels of shared work experience in two person doctor teams is associated with reduced patient mortality.

Our paper is most closely related to Bartel et al. (2014), who examine the impact of month-to-month variation in general and ward-specific human capital on the productivity of nursing teams in the USA. They find that experience working in the specific hospital unit (ward) is more important than general human capital (as measured by education). We extend this research by exploiting the higher granularity and higher frequency of our data and by focusing on a measure of clinical quality, patient mortality, as an outcome, while Bartel et al. only examine patient length of stay. The high frequency of our data enables us to examine changes in teams at the level at which teams are planned (the shift level). It also enables us to examine possible dynamic relationships between absences and patient outcomes as well as contemporaneous effects. The granularity of our data on the staffing side allows us to go beyond Bartel et al. in distinguishing the impact on team productivity of different levels and types of human capital. These include seniority *within* qualification group, contract type (agency vs regular workers), and experience at the hospital, ward and team (co-worker) levels.[3] On the patient side, using data on

[3]There is an extensive literature on interactions between team members, including peer effects in the workplace (Mas and Moretti, 2009; Guryan et al., 2009; Cornelissen et al., 2017; Silver, 2021), learning from co-workers (Jarosch et al., 2021), credit-sharing within teams (Jones, 2021), productivity spillovers (Arcidiacono et al., 2017) and team incentives (Hamilton et al., 2003; Bandiera et al., 2009, 2013; Burgess et al., 2010, 2017; Friebel et al., 2017). With the exception of the last (as workers are not paid for team

diagnoses enables us to assess which types of patient are most affected.

Second, we contribute to the literature on the role of nurses in health care production, building on papers such as Friedrich and Hackmann (2021), who exploit the labour supply effects of a parental leave program in Denmark to quantify the impacts of nurse employment on patient outcomes; Gruber and Kleiner (2012), who examine the impacts of nursing strikes; Propper and Van Reenen (2010), who exploit variation in outside wage options for nurses induced by centralised pay regulation to examine the impacts on hospital performance; and Lin (2014), who exploits legislative changes to minimum staffing requirements in nursing homes. Our data, which is both more granular and higher frequency than that used in most previous studies, provides new evidence on the impacts of both the quantity and quality of nurses on patient outcomes and of the potential mechanisms behind those effects.

A third related literature is concerned with worker human capital and productivity. These issues have received growing attention in the healthcare setting (e.g., Doyle et al. (2010); Currie and MacLeod (2017, 2020); Chen (2021); Chan et al. (2022); Chan and Chen (2022)). More widely, our paper relates to the relationship between human capital and productivity (e.g. Moretti (2004); Gennaioli et al. (2013)).

The paper proceeds as follows. Section 2 discusses team production in a nursing setting and motivates the human capital concepts we focus on in the paper. Section 3 describes the institutional context and our data. Section 4 presents descriptive statistics. Section 5 sets out our empirical strategy and considers possible threats to identification. Section 6 presents our results, looking first at the impact of overall staffing and general human capital and the returns to more specific forms of experience and then examining potential pathways for our findings. Section 7 concludes.

## 2 The team in a nursing setting

Our empirical analysis focuses on how nursing team composition affects production of healthcare. A team is a set of nursing staff who work together in a particular location (hospital ward) to treat the patients in that ward. While there are some tasks where a single nurse can safely care for multiple patients, other tasks require team members to work together over the course of a nursing shift, which requires collaboration and coordination of team actions. In addition, there is a need to convey information both within teams and across teams when shifts change. In some settings, the output of each team member can be observed and recorded (for example, as in Bandiera et al. (2009,

output in our setting) these mechanisms may underlie our results, but we do not directly examine them.

2013)). But in the nursing setting the output is a team output and cannot be directly attributed to any single individual. Providing high-quality patient care is a team effort, as is winning a match in a team sport such as basketball. What matters, therefore, is *team* productivity, which will depend on the size and composition of the team.

We estimate production of the nursing team as a function of four types of human capital, each accumulated through different types of experience. These are general, firm-specific (hospital), plant-specific (hospital ward) and team-specific experience (team of nursing co-workers). General human capital – such as overall education and experience – should raise an individual's productivity (Becker, 1964; Mincer, 1974) in nursing, as in other forms of production that require individuals with detailed knowledge and the ability to take decisions without high levels of supervision. Firm- or task-specific human capital may also be important (Parsons, 1972). Experience at a particular firm means workers develop firm-specific skills and knowledge that make them more productive at that firm, but not necessarily at other firms. Returns to firm-specific human capital are likely in the nursing context: different hospitals often employ their own systems, processes and protocols. The longer an individual has worked at a particular hospital, the more familiar she will be with those systems, processes and protocols. Huckman and Pisano (2006) provide evidence of returns to hospital-specific experience for surgeons carrying out cardiac surgery.

Similarly, different wards within a hospital may employ different systems and/or require nurses to perform different tasks, in which case some portion of accumulated human capital would be ward-specific. In other words, it may be that experience in a particular location or in performing a particular set of tasks yields productivity benefits only in that location or when performing those tasks. Finally, if nurses have worked together before they will be familiar with how other members of the team work and therefore may be more effective in carrying out joint tasks and in communication of tacit knowledge and time critical information.

To identify the impact of these different inputs we exploit the relative effects of shortages of different types of human capital on the output of the team. We posit that the smaller the size of the team, relative to its target or planned size, and lower team levels of human capital – whether general, firm-specific, plant-specific or team-specific – will be associated with lower team output.[4] The precise form of the production function

[4]An absent team member has a both a direct impact on team production by causing a team to be one member short, but may also have an indirect effect if they are replaced by someone with lower skills or familiarity with the production process (Herrmann and Rockoff, 2012; Bartel et al., 2014). Hoey et al. (2022) distinguish between the two in a setting where teams are always fully staffed. In our context they are not and therefore we do not make a distinction between direct and indirect effects.

and our identification strategy are discussed in Section 5. Before that we present the institutional context and our data.

# 3 Institutional background and data

Nursing teams in English hospitals are composed of two broad types of staff, differentiated by their general human capital. The first are registered nurses. These are fully qualified nurses, registered with the Nursing and Midwifery Council, who have completed formal training and hold a university diploma or degree-level qualification. The second are nursing assistants, who are personnel without formal training or registration requirements and who have responsibility only for basic patient tasks. We refer to the former as qualified nurses or RNs, and the latter as nursing assistants or NAs. Nurses are paid according to a national set of pay bands. Nursing assistants are typically employed on a single pay band, with multiple spine points. Within qualified nurses, there are multiple pay bands, each corresponding to a different level of responsibility. Each pay band has multiple spine points to account for experience. Promotion between pay bands is based on experience and typically requires specific training or qualifications.

## 3.1 Data sources

We exploit a novel data linkage between electronic staff roster data and electronic patient records for one NHS hospital, comprising a small number of spatially closely located acute hospital sites.[5] This hospital provides a full range of emergency and elective services, has a medical school attached, and is located in a major city with an ethnically and socioeconomically diverse population.[6] We combine routinely recorded administrative information on the size and composition of nursing teams with information on the characteristics and outcomes of patients under their care. Our data covers 1st January to 31st December 2017.

[5]Hospitals are known as NHS Trusts in England: for simplicity we refer to the Trust we study as a hospital.

[6]It is a relatively large in terms of admissions volumes and beds, but is not an outlier in this respect, and is representative of other English hospitals with similar characteristics. There are 25 medical schools in England, all of which are attached to one or more hospital Trusts. There were 31 hospital Trusts in England with 1,000 beds or more in the period January to March 2019. The Trust studied here belongs to both categories.

### 3.1.1 Teams

The hospital we examine has three large sites providing short term general acute care. These sites contain 89 wards in total, but to reduce heterogeneity we focus on inpatient wards that provide 24/7 hour care for adult patients.[7] For each of these 52 wards, we construct measures of staffing, patient characteristics and patient outcomes at the 24 hour level. We define each team to be a day-ward observation. Our final sample contains 18,922 observations.[8] Each of these is a distinct team comprising the nursing staff members who worked together on a given calendar day (including the handover between those doing the night shift and those who take over the following morning).[9]

Analysis at the team level allows us to study the different impact of team members with different levels of human capital and also has the advantage that concerns over unobserved patient severity are mitigated by averaging over all patients under the care of the team. As a robustness check in Section 6.3 we repeat our main analysis using the patient-day as the unit of observation.

### 3.1.2 Staffing

Nursing shifts are generated automatically by the e-rostering system several (generally three) months ahead of time in order to provide staff with at least eight weeks notice of their shift patterns. There are no mandated nurse:patient ratios in the UK. Instead the hospital uses an algorithm that generates the required number of shifts. This is based on the 'Safer nursing care tool' (SNCT), a national, evidence-based tool endorsed by the National Institute for Health and Care Excellence. The tool is based on expected levels of patient acuity and dependency.[10] The mix of staff across bands or grades is

[7] We exclude: 18 wards which do not record a patient death at any point in the year (which includes administrative wards and those that treat outpatients); 14 wards with non-continuous care, such as those that are closed on weekends or which regularly close to patients (defined as having an entire day with no patients in the ward 30 times or more over the year); and 5 maternity, paediatric and emergency department wards, which have very different staffing models and operational procedures.

[8] The sample does not contain 365 observations for each of the 52 wards (18,980 observations) because on 58 days (across 26 wards) no patients were treated, and these days are excluded from our sample.

[9] As we define a team as a 'hospital ward-day', a night shift that spans midnight will contribute to the staffing of more than one team. For instance a shift that starts at 20:00 on day t and ends at 08:00 on day t+1 will contribute 4 hours to the staffing of the first team and 8 hours to the staffing of the second. The mean length of a nursing shift is 11.5 hours and more than 70% of shifts are 12 hours long. 4,484 unique staff members worked in these teams, working 3,533,583 hours and 294,044 shifts over the course of the year. The teams were responsible for treating 44,485 unique patients with 66,662 separate hospital spells.

[10] Each January and June, all wards place patients in each bed into one of 5 groups based on levels of care required. This data is collected at 15:00 daily, Monday to Friday, for a minimum of 20 days. Each of the patient acuity groups has a different staff multiplier (0.99 for the lowest category, rising to 5.96

determined by the hospital. Shifts are then allocated to individual staff members either by the system or manually by ward managers.

Importantly, not all planned shifts are ultimately worked because of staff absences (due to sickness or annual leave, for instance), because there is a staff vacancy, or for another reason such as attendance at a training course. If there are no absences from the planned rota, a shift will be worked by 'local' nurses. These are employees of the hospital, working their contracted hours in the wards to which they are attached. In the event of an absence, the senior manager of the ward can try and fill this with a 'bank' nurse (nurses employed by the hospital but doing extracontractual or overtime work via the hospital 'staff bank', in their usual ward or another ward) or an agency nurse (nurses employed via an external agency and generally contracted at short notice).

The NHS, and the hospital we study, has experienced a shortage of nurses for many years. There are therefore not a large number of the hospital's directly employed staff (i.e. bank staff) who can frequently step in at short notice to cover unplanned absences. In addition, to control costs, the hospital has measures to limit the use of (expensive) agency workers. This means in the case of an unplanned absence, there may be no bank staff available and permission to use an agency worker may not be granted in time to bring the staffing up to the planned levels.[11]

NHS employees are paid according to a national pay structure, known as 'Agenda for Change' (AfC). The electronic staffing records include information on the AfC pay band of the nurse who worked each shift. We use this to identify skill levels and seniority. Staff in AfC pay bands 2–4 are nursing assistants and those in pay bands 5–8 are degree level qualified nurses. Individual pay bands also identify staff of different levels of seniority within each group. Within the category of qualified nurses, Band 5 contains newly qualified and less experienced nurses; Band 6 contains more senior and experienced nurses, including deputy ward managers; and Band 7 and 8 contains the most senior nurses and ward managers with the most responsibility. For those nurses employed by the hospital, the staffing records include the month and year in which their employment commenced, which we use to construct a measure of experience within the hospital.[12]

for the highest). Total staffing requirements are given by multiplying the number of patients in each category by their multiplier and then adding across the five groups. There are adjustments for patient flow (rates of admission and discharge) and measures of nursing quality or 'nurse sensitive indicators', which include drug errors, pressure ulcers and falls. The output then serves as a baseline against which nurse staffing levels are set.

[11] Actual staffing could still respond to patient severity if ward managers faced with shortages and/or sicker patients make greater efforts to fill shifts. In Section 5.2 we test whether staffing responds to patient composition and find no evidence that it does.

[12] Agency workers are not employed by the hospital and are given a value of 0 on this measure. This

### 3.1.3 Patient outcomes and severity

We use an important and high-stakes outcome: inpatient mortality. Patient mortality is a commonly used performance measure for hospitals and also for nursing (Needleman et al., 2011; Griffiths et al., 2019) and has the advantage of being unambiguous and of being accurately and consistently recorded across wards. Death at discharge is recorded in electronic patient records, which we link to the team responsible for treating them at the point of death. Our outcome is binary, with an indicator coded as 1 if a team (ward-day) had a patient death occur, and coded as 0 otherwise.[13]

In analyses of causes of death we use the International Classification of Diseases (10th edition; henceforth ICD-10) codes contained in the electronic patient records to construct binary indicators of the types of patient deaths that are most common in our sample and those which have been shown to be most sensitive to adequate provision of nursing care.

The electronic patient records contain information on a range of characteristics related to the patients' acuity and mortality risk. These include age, sex, ethnicity, and primary and secondary ICD-10 diagnosis codes, which we use to calculate a standard measure of clinical severity (the Elixhauser Comorbidity Index score (Elixhauser et al., 1998)) for each patient. We also calculate the length of hospital stay for each patient. We aggregate up to the team level, weighting by the amount of time that each patient spent under the care of the team.

## 3.2 Team staffing measures

### 3.2.1 Quantity of nurses

We measure the quantity of staffing in terms of deviations from the target staffing set by the hospital management for the ward. Specifically, we define the 'fill-rate' as the fraction of planned hours that were actually worked in ward $i$ on day $t$:

$$FillRate_{it} = 100 * \frac{\text{Hours worked in ward } i \text{ on day } t}{\text{Planned hours in ward } i \text{ on day } t} \tag{1}$$

where a team with a fill-rate of 100 is fully staffed, with staffing equal to target.[14]

is conservative if an agency worker is frequently employed by the same hospital.

[13] Deaths occur at all hours of the day (Appendix Figure A7), with the only noticeable spike occurring between 19:00 and 20:00 (when the handover between day and night shifts takes place) which suggests that there is no systematic misallocation of patient deaths across calendar days.

[14] Staffing cannot exceed the target level in this setting, so the fill-rate has an upper bound of 100. The NHS runs very close to capacity at all times: bed occupancy is over 95% and there are waiting lists for planned (elective) care to ensure beds are not left empty.

The ratio between actual and planned staffing – the 'fill-rate' – has a number of appealing features for our analysis. First, it is not sensitive to the absolute size of the ward in the question. A shortage of one staff member is a much more significant event for a team of five nurses than it is to a team of thirty: using the fill-rate captures this, because deviations are measured in percentage terms. Second, deviations are measured relative to a locally determined, ward-specific target which is considered by the hospital management to be the 'safe' level of staffing (informed by the SNCT), rather than relative to a mandated national-level minimum nurse:patient ratio. This allows for heterogeneity in what is medically considered to be safe staffing across different hospital wards. Third, we can define the fill-rate separately for different groups of nurses, which allows us to study the impact of both size and composition of the team in a comparable way across wards.

### 3.2.2 Human capital measures

**General human capital:** We measure the amount of general human capital in the team by the pay band of each team member. We group together all nursing assistants (NAs, those in AfC bands 2–4), and all registered nurses (RNs, those in AfC bands 5–8). We construct the fill rate (Equation 1) separately for NAs and RNs: the share of planned NA hours filled by NAs, $FillRateNA_{it}$, and the share of planned RN hours filled by RNs, $FillRateRN_{it}$.

We separately identify the different grades (bands) within the set of RNs to distinguish between qualified nurses of different levels of seniority and experience. We construct the fill rate for each band as a share of total planned RN hours. The sum of the fill rates for bands 5-8 will sum to $FillRateRN_{it}$.[15] This approach allows us to compare the impact, for instance, of a one percentage point increase in the RN fill-rate from a newly qualified (band 5) nurse versus a senior nurse manager (band 7 and 8). Such direct comparisons allow us to estimate the the relative returns to different amounts of general human capital within the team.

**Firm (hospital) specific human capital:** We define firm-specific human capital as the amount of time an individual has been employed in the hospital at any of its sites.[16]

[15]For example, the band 5 fill rate $FillRateRN_{it}^{Band5}$ is equal to 100 $*$ (Hours worked by Band 5 RNs in ward $i$ on day $t$)/(Planned RN hours in ward $i$ on day $t$). A full set of definitions is provided in the appendix.

[16]For 89% of individuals in our sample (who work 87% of all shifts) tenure is directly observable in staff records, as these contain the date at which the individual's employment at the hospital started. For the remaining 11% of individuals, for whom this information is missing, experience in the hospital

**Plant (hospital ward) specific human capital:** We measure plant (ward)-specific human capital in three ways. First, as the number of times that an individual has worked in a ward over the past 90 days.[17] We average this over the members of each team in our sample. Second, as the fraction of the individual's shifts in the last 90 days that were in the ward, again averaged over the members of each team. Third, we identify those who are less familiar with the ward. If an individual is within their first 30 days on the ward, having not worked there in the preceding 90 days, we define these individuals as 'new joiners' and construct a binary variable to indicate the presence of a 'new joiner' in a team.

**Team-specific human capital:** We define team-specific human capital in three ways. First, as the average number of shifts that an individual has worked together with their team mates, in any ward, over the past 90 days. Second, as the fraction of their fellow team members an individual has worked with at least once in the past 90 days. Third, as the fraction of fellow team members an individual has worked with at least 10 times in the past 90 days. In our analysis, we use the average of each of these measures across the members of each team.

# 4 Summary statistics

## 4.1 Staffing and team composition

The average team (ward-day) had 194.8 rostered hours. Within that, an average 53.4 hours were rostered for nursing assistants and 141.4 for qualified nurses (see Table A1). This is equivalent to around 16 staff members working 12-hour shifts, with around 4 assistants and 12 qualified nurses in the average team. Not all of these staff members will be working at any one time and shifts overlap in order to ensure continuous care.

Our analysis exploits variation in staffing hours as a percentage of planned (i.e. the fill rate) and the composition of the team. Table 1 provides the summary statistics for

is imputed based on their pay band, the % of shifts they work as overtime via the staff 'bank', the % of shifts they work at night, and whether they are an agency worker (agency workers are assigned a value of zero, as they are not employed by the hospital).

[17] We measure ward-specific experience based on the past 90 days for two reasons. One, we might expect the effects of experience to decay over time. Two, our sample covers a single calendar year and in order to construct a backwards-looking measure of experience, we have to discard a portion of the data at the start of the period (for which such a backwards-looking measure cannot be calculated). Using ward-specific experience over the previous 90 days strikes a balance between capturing how familiar workers will be with the ward, and having to discard no more than a quarter of our sample.

fill rates by staff group and average levels of different measures of human capital at the team level.

In just over half (51.2%) of teams, the actual number of hours worked was less than the number of planned/rostered hours, so just over half of teams had a fill-rate of less than 100. The mean fill rate for the average team, as shown in Table 1, was 95.7 (i.e. 95.7% of planned hours were actually worked), equivalent to an average of 8 hours of planned staffing unfilled, or two-thirds of an average shift unfilled. The fill rate was slightly lower for healthcare assistants (94.8%) than for qualified nurses (96.4%). There is little correlation between the NA and RN fill-rates (coefficient=0.066). The distribution of and variation in the RN and NA fill-rate across wards is summarised in Figure 1.

The majority of hours – 67.9% of all those planned – were worked by local staff in their regular ward to which they are attached and in which they do their contracted hours. A further 20.6% were worked by bank staff and a further 7.2% by agency workers. Nurses in pay band 5 comprise the largest staff group. Band 5 nurses work almost half (47.1%) of all hours and more than two-thirds (68.4%) of rostered RN hours. Agency work is also concentrated in band 5. Almost all hours worked by the more senior Band 6–8 nurses are in their regular ward. Band 7 and 8 staff – the most senior nurses responsible for managing their ward – work around 5% of all rostered RN hours.

In the average team, the average staff member had 6.3 years of experience within the hospital (5.6 years for nursing assistants and 6.7 for qualified nurses). The average team member had done 30.3 shifts in the ward in the past 90 days (27.8 for assistants; 31.5 for qualified nurses) and the average team member had worked 14.5 shifts with their fellow team members in the previous 90 days. A little over a tenth (12.2%) of all planned hours were worked by 'new joiners' who were within their first month on a ward, having not worked there for the previous 90 days or more.[18]

Importantly, there is considerable day-to-day within-ward variation in both the level and composition of staffing. This includes variation in the degree of hospital-, ward- and team-specific human capital amongst the team. This within-ward variation is key to our identification strategy. In part, this variation is because of staff rotations, arrivals and departures which contribute to turnover. But in addition, most staff members work across multiple wards, as shown in Figure 2. Junior staff, in particular, tend to work across multiple sites and wards within the hospital. The average nursing assistant worked 76.3 shifts across 6.8 different wards over the course of the year. The average band 5

[18]Summary statistics for all staffing measures are in Table A1.

nurse worked 51.6 shifts across 4.5 unique wards.[19] In contrast, the most senior nurses in band 7 and 8 move around little and did 96.7% of their shifts in their usual ward (compared to 72.4% for band 5s).

Table 1: Summary statistics: within-ward and between-ward standard deviation (SD)

| | Mean | SD | Within SD | Between SD |
|---|---|---|---|---|
| Share (%) of all planned hours worked | 95.7 | 5.4 | 4.9 | 2.2 |
| Share (%) of planned NA hours worked | 94.8 | 10.9 | 10.3 | 3.8 |
| Share (%) of planned RN hours worked | 96.4 | 5.8 | 5.3 | 2.3 |
| Share (%) of all planned hours worked by: | | | | |
| Band 5 staff | 47.1 | 11.5 | 8.2 | 8.1 |
| Band 6 staff | 17.0 | 12.5 | 6.3 | 10.8 |
| Band 7 and 8 staff | 3.8 | 4.5 | 3.3 | 3.1 |
| Local staff | 67.9 | 15.7 | 11.7 | 10.6 |
| Bank staff | 20.6 | 12.7 | 10.4 | 7.4 |
| Agency staff | 7.2 | 8.2 | 6.5 | 5.1 |
| Firm-specific experience: | | | | |
| Mean years of Trust experience (all) | 6.3 | 2.0 | 1.2 | 1.6 |
| Mean years of Trust experience (NAs) | 5.6 | 3.2 | 2.1 | 2.4 |
| Mean years of Trust experience (RNs) | 6.7 | 2.5 | 1.5 | 2.0 |
| Ward-specific experience: | | | | |
| Mean shifts in the ward in the past 90 days (all) | 30.3 | 5.2 | 4.2 | 3.2 |
| Mean shifts in the ward in the past 90 days (NAs) | 27.8 | 9.7 | 8.1 | 7.1 |
| Mean shifts in the ward in the past 90 days (RNs) | 31.5 | 5.9 | 4.7 | 3.5 |
| Team-specific experience: | | | | |
| Mean shared work experience (shifts) with other team members (in any ward) in the past 90 days (all) | 14.5 | 5.0 | 3.9 | 3.2 |
| Mean shared work experience (shifts) with other RNs (in any ward) in the past 90 days (RNs) | 15.2 | 5.7 | 4.5 | 3.5 |
| Patient death (binary) | 0.06 | 0.24 | 0.24 | 0.06 |
| Death rate per 1,000 patients | 3.58 | 16.20 | 15.67 | 4.15 |

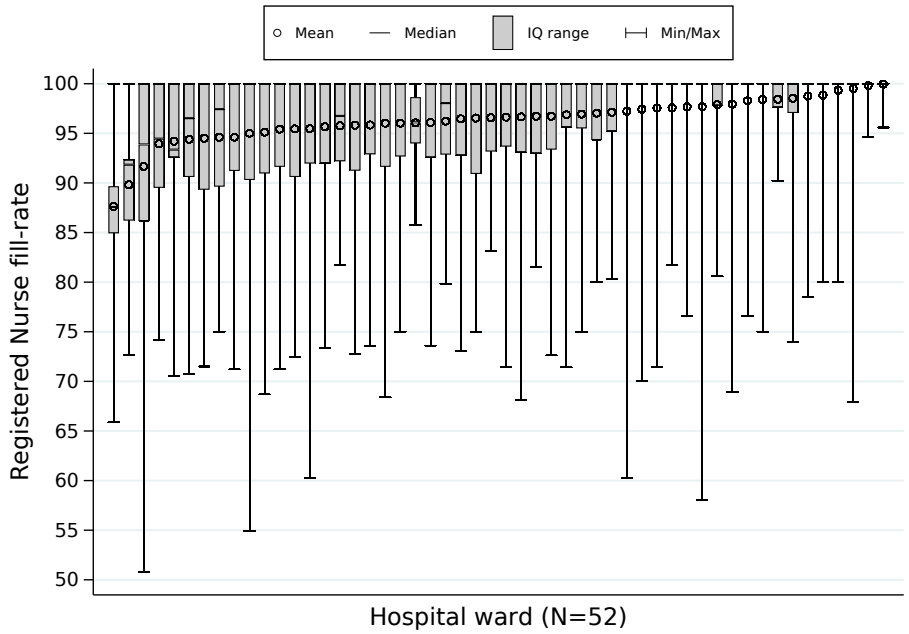
## 4.2 Patients and patient outcomes

The average team was responsible for treating 20.6 patients over the course of the 24-hour ward-day (Table A1). The average patient was 64.1 years old, 46% were female,

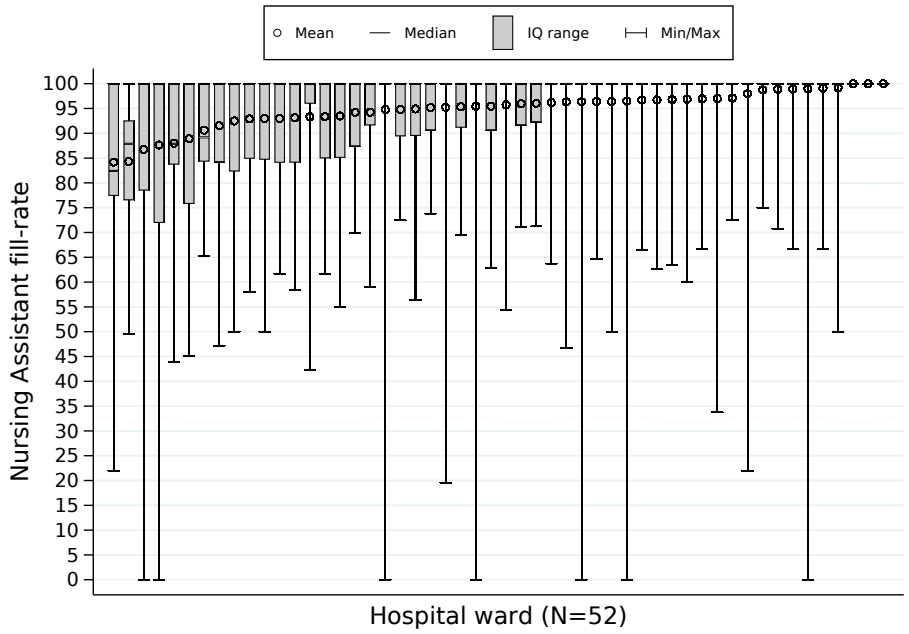
[19] The average band 5 works fewer shifts because they are more likely to be part-time, and more likely to be an agency worker (who could appear as little as once in our sample). After excluding agency workers, the mean number of shifts worked over the year by band 5 nurses was 89.1 (with a median of 92).

Figure 1: Distribution of and variation in fill-rate across wards

(a) Registered Nurses



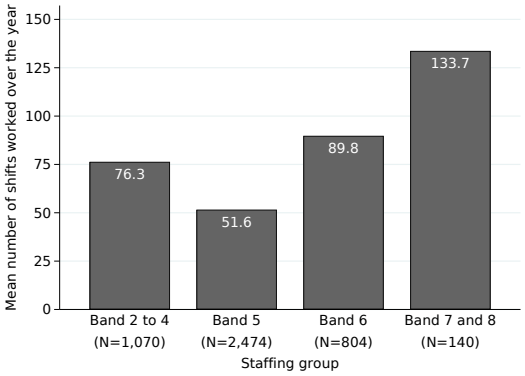

(b) Nursing Assistants



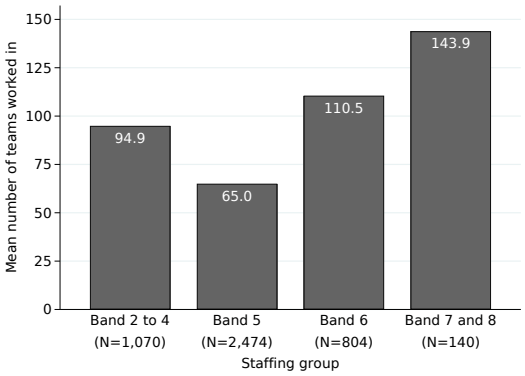

Note: Figures are calculated with respect to the 2017 calendar year (the period covered by our sample). The ordering of wards differs between panel a) and panel b).

Figure 2: Staff movements across teams, wards and hospitals

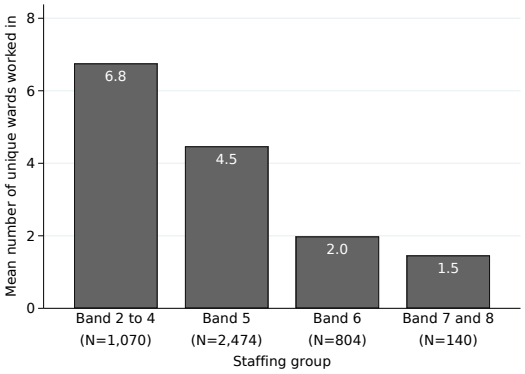
(a) Number of shifts (mean)



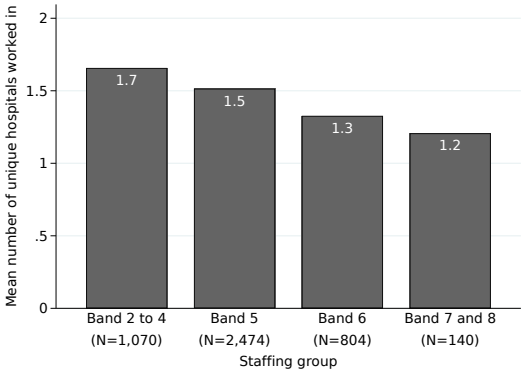

(b) Unique teams (mean)



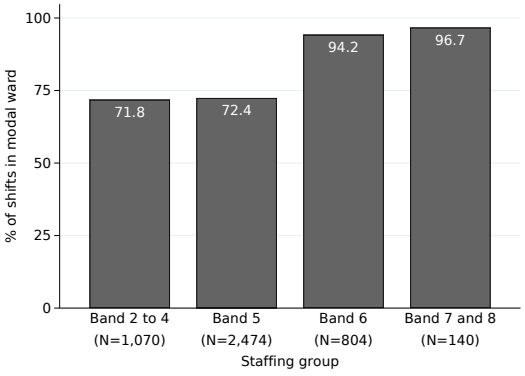

(c) Unique wards (mean)



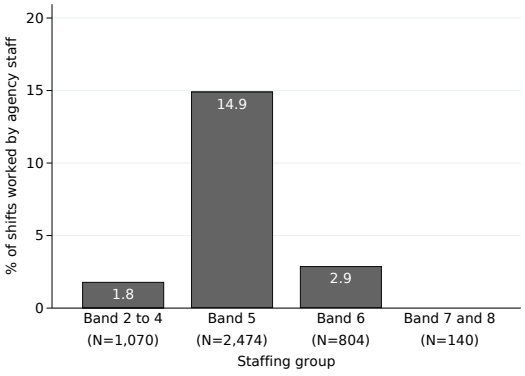

(d) Unique hospitals (mean)




(e) Share (%) of shifts in usual (modal) ward




(f) Share (%) of shifts worked by agency staff




Note: Figures are calculated with respect to the 2017 calendar year (the period covered by our sample).

and 52% were recorded with an ethnicity other than White (including 'unknown'). The average severity, as measured by the Elixhauser Comorbidity Score (Elixhauser et al., 1998), was 3.69. 6.3% of the teams in our sample (1,197 of 18,922) recorded a patient death. The mean mortality rate was 3.58 per 1,000 patients with considerable variation between and, importantly for identification, within wards (Appendix Figure A8 and Table 1).

# 5 Empirical strategy

## 5.1 Baseline model

We begin by estimating a simple production model in which the probability of a patient death is a function of a shortage of human capital. We allow for different effects of shortages of higher (qualified nurses) and lower (nursing assistants) human capital. We begin by assuming production is additively separable in the two types of human capital, with no explicit interactions or complementarities between different types of capital.

In this simplest model, we assume that healthcare $G$ is produced according to $G(L_{NA}, L_{RN}) = \beta_1 L_{NA} + \beta_2 L_{RN}$. To implement this, we estimate the following baseline fixed-effects logit model:

$$logit(Pr(y_{it} = 1|x_{it})) = \beta_1 FillRateNA_{it} + \beta_2 FillRateRN_{it} + \delta PatChars_{it} + \lambda_t + \eta_i \tag{2}$$

where $y_{it}$ is a binary indicator of a patient death in ward $i$ on day $t$ (where the ward-day is the team), $x_{it}$ is a vector of explanatory variables, and the function $logit(z)$ is defined as $log(\frac{z}{1-z})$.

Staffing quantity is measured by the fraction of planned nursing assistant and nursing hours that were actually worked, denoted by $FillRateNA_{it}$ and $FillRateRN_{it}$, respectively (as formally defined in Section 3.2). $PatChars_{it}$ measures time-varying characteristics of patients in ward $i$ on day $t$. We include the mean age of patients; the mean age squared; the female share; the share of patients who are non-white; the mean patient Elixhauser Comorbidity Index score; pairwise interactions between each of these variables; and the average length of hospital stay for the patients in the ward. We also include the number of patients treated in ward $i$ on day $t$ to control for variation in workload. $\lambda_t$ captures time effects (day of week; month; and a dummy variable indicating a public holiday). Ward fixed effects ($\eta_i$) capture time-invariant differences in physical

capital and the type of patient across wards (such as their medical specialty, patient mix, hospital management quality, and physical layout). Standard errors are clustered at the ward level.

$\beta_1$ and $\beta_2$ are the coefficients of interest, and represent the impact of a change in the quantity of lower and higher human capital, respectively, on the probability of a team experiencing a patient death. We present our results as odds ratios for ease of interpretation.

We extend this model to additionally estimate the other dimensions of human capital discussed above, including the average amount of firm-specific experience in the team, the average amount of recent experience working in the ward, and the average amount of recent experience working with other team members.[20]

## 5.2 Threats to identification

Identification in Equation (2) requires that that staffing is exogenous to outcomes. The institutional set-up makes such an assumption plausible. Nurse rosters are generated 2–3 months in advance to ensure that each ward contains the appropriate number and mix of staff. Deviations arise because the appropriate number and mix of staff cannot be found due to absences, departures, and more general staffing shortages. However, there are a number of possible threats to this key assumption and we discuss these here.

### 5.2.1 Staffing levels and patient mix

*Planned* staffing levels are exogenous to eventual deviations from expected patient volumes and severity, as they are determined several months before actual staffing takes place. But *actual* staffing could be endogenous to patient severity if managers respond to an influx of sicker patients by adjusting either the quantity or quality of staff on the ward. By definition, a more severe patient mix means a greater risk of a death. If staffing levels or quality adjust upwards in response to patient severity and if there is a negative relationship between staffing and the risk of death, this would bias our estimates towards zero and mean that we underestimate the true effect of higher nurse staffing on deaths.

The institutional context and the way in which shifts are allocated and filled lessen such concerns. First, staff shortages are extremely common in the English NHS. Nurses account for around 25% of all NHS staff (NHS Digital, 2021b), but almost 40% of all vacancies (NHS Digital, 2021a). More than half of the teams in our sample were under-staffed relative to target. The primary means of finding an additional nurse at

[20] See Appendix Sections A.1 and A.2 for a full set of summary statistics and definitions.

short notice is via an external agency, but budgetary constraints mean that this requires approval from a manager external to the ward. As noted above, this approval may not granted in time to fill an unexpected vacancy, or may not be granted at all.

However, we undertake a number of tests to examine the assumption that staffing is not related to patient severity. First, we examine whether staffing levels on day $t$ in ward $i$ are associated with observable patient severity. We find no correlation (Figure A5 and Table A2).[21] We examine whether the margin that could be most easily used to respond to changes in severity – agency staffing – is a function of patient severity. We find no evidence that agency staffing is higher when average patient severity is greater (Table A3). If anything, greater patient severity – as measured by the Elixhauser Comorbidity Index – is associated with *lower* levels of agency staffing (Figure A6). And, more generally, we are not able to predict the level of RN staffing using the observable characteristics of the patients under their care, which is consistent with our asssumption of random assignment (column 2 of Table A3). If staffing levels are not responding to changes in *observable* patient severity, this suggests they do not respond to changes in other, less easily observable, patient characteristics associated with mortality.

There is some evidence that agency staffing is higher when there is a greater *volume* of patients in the ward, but we control for numbers of patients in our analysis. Patient volumes in a particular ward could potentially be adjusted in response to staffing. For example, patients could be moved to other wards within the hospital if one ward were short staffed. We examine this and find no statistically significant relationships between any of our staffing measures and patients outflow to other wards (Table A4).

### 5.2.2 Staff shortages and ward performance

A second threat to our identification strategy would arise if staff shortages are endogenous to recent or expected ward performance. That is, staffing is lower because of 'no-shows' that are themselves a response to recent patient outcomes. Table A5 examines the effect of deaths in the ward on day $t-1$, $t-2$ and $t-3$ on the level and composition of staffing on day $t$. There is some weak evidence that a death on the previous day is associated with a lower RN fill-rate (column 2) at the 10% level. However, the effect is very small: a death on day $t-1$ is associated with a 0.23 percentage point reduction, which is equivalent to only 4% of a standard deviation. No such associations are evident for deaths two or three days previously (and the coefficient changes sign in the case of the latter). There are no statistically significant effects on NA staffing, nor on

[21] We similarly find no correlation between patient severity and the level of firm-specific, ward-specific, or team-specific experience in the team (results available on request).

the composition of RN staffing, nor on the level of agency staffing. We additionally look at whether a recent death in the ward is associated with a greater number of subsequent sickness absences (Table A6). For qualified nurses we find no statistically significant relationship between deaths in the previous three days and subsequent sickness. The coefficients are also small and vary in sign. For NAs, there is a statistically significant increase in sickness two days after a death on the ward, but the coefficients for a day on $t-1$ and $t-3$ are negative (with an odds ratio below 1) and not statistically significant. We also do not find any evidence of serial correlation in deaths, which lessens concerns that nursing staff might be able to predict a difficult shift, as well as concerns about serial correlation in the error term (Table A7). Taken together these tests strongly suggest that staff shortages are not a response to recent or expected patient mortality outcomes.

#### 5.2.3 Other staffing adjustments

A final concern is that because our data only includes information on nursing staff, we are unable to capture any endogenous responses from other staff groups. For instance, if a team is short-staffed of nurses, the hospital might respond by bringing in more doctors. In fact, the rostering system for doctors is separate and orthogonal to the rostering system for nurses. Further, a junior doctor would not fill in for a missing nurse: if anything, a senior nurse is more likely to step into the shoes of a missing or inexperienced junior doctor and so we would underestimate the effect of any shortage of nurses. Our ward and month fixed effects capture permanent differences in doctor quality across wards and any effects of seasonal doctor rotation, respectively.

## 6 Results

### 6.1 Overall staffing and general human capital

Table 2 reports our baseline results for quantity separated into higher and lower general human capital. These are shown graphically in Appendix Figure A1. The first column regresses a binary indicator of a patient death on the nursing assistant (NA) and registered nurse (RN) fill-rate (as defined in Section 3.2), with the controls discussed above. A one-unit (one percentage point) increase in each of these variables is equivalent to an additional 0.53 and 1.41 hours, respectively, for the average team (as per Table 1), though this will vary according to the size of the ward.

An increase in nursing assistant staffing has no statistically significant impact on the odds of a patient death. In contrast, higher levels of qualified nurse staffing are

associated with lower patient mortality: a one percentage point increase in the RN fill-rate reduces the odds of a patient death by 1.3% (with an odds ratio of 0.987). Using these results, an extra RN doing a 12-hour shift would reduce the odds of experiencing a patient death by around 10% in the average team.[22]

The second column distinguishes between qualified nurses of different seniority. Nurses in higher pay bands have more general human capital. Our point estimates indicate that a one percentage point increase in the share of planned RN (band 5–8) hours worked by a band 5 nurse would reduce the odds of a patient death by 1.1%, increasing to 1.6% for band 6, and 2.4% for band 7 and 8. While these estimates are not statistically significantly different from one another, the ordering and relative magnitudes of the point estimates are indicative of greater returns to higher skills. For the average team, adding an extra nurse in the most senior pay bands reduces the odds of a patient death by more than twice as much as adding a newly qualified (band 5) nurse.

The third column of Table 2 distinguishes between qualified nurses of the three different contract types, local, bank and agency. The results suggest that nurses directly employed by the hospital group – local and bank staff – have a bigger impact than do agency nurses, though the coefficients are again not statistically significantly different from one another. However, agency RNs are also almost exclusively employed in band 5, so this could just be a reflection of the skill and seniority mix of the team, as a team with a greater share of agency RNs would also be a team with a greater share of band 5s. To address this, the fourth column controls for band 6, 7 and 8 staffing, but separates band 5 staffing into agency and non-agency. The coefficients on agency and non-agency staffing are very similar, suggesting that it is not contract type that matters but the general human capital of the team.

[22]The average team has 141.4 rostered RN hours (Table 1). A one percentage point increase would therefore add approximately 1.4 hours, and an extra staff member working a 12-hour shift would increase staffing by $12/1.4 \approx 8.6$ percentage points of planned RN hours. Using the results from Table 2, $0.9874^{8.6} \approx 0.9$, equivalent to a 10% reduction. It is difficult to directly compare the magnitude of our estimate to previous studies in the nursing literature, as most have been conducted at the patient-level rather than the team-level. Past research has found that the hazard of death increased by 2% for patients exposed to shifts with RN staffing 8 hours or more below target (Needleman et al., 2011) and by 3% for every day patients experienced treatment from a team with below ward-mean RN staffing (Griffiths et al., 2019). Our estimates are broadly in line with these.

Table 2: Impact of nurse staffing levels, contract type and seniority on patient mortality

| | **Outcome: patient death** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.003<br>(0.004)<br>[0.360] | 1.003<br>(0.004)<br>[0.356] | 1.003<br>(0.004)<br>[0.348] | 1.003<br>(0.004)<br>[0.348] |
| Share (ppt) of planned Registered Nurse (RN) hours filled | 0.987**<br>(0.006)<br>[0.038] | | | |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | | | |
| Band 5 nurses | | 0.989*<br>(0.006)<br>[0.087] | | |
| Band 6 nurses | | 0.984**<br>(0.007)<br>[0.020] | | 0.984**<br>(0.007)<br>[0.020] |
| Band 7 and 8 nurses | | 0.976**<br>(0.011)<br>[0.025] | | 0.976**<br>(0.011)<br>[0.026] |
| Local (regular) nurses | | | 0.987**<br>(0.006)<br>[0.037] | |
| Bank nurses | | | 0.986**<br>(0.007)<br>[0.038] | |
| Agency nurses | | | 0.991<br>(0.006)<br>[0.140] | |
| Band 5 Local and Bank nurses | | | | 0.989*<br>(0.006)<br>[0.084] |
| Band 5 Agency nurses | | | | 0.991<br>(0.006)<br>[0.139] |
| Patient volume and characteristics | ✓ | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 18,922 | 18,922 | 18,922 | 18,922 |
| Clusters | 52 | 52 | 52 | 52 |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. Nursing assistants (NAs) are personnel without formal training or registration requirements; Registered Nurses (RNs) are fully qualified nurses on the Nursing and Midwifery Council register, who have completed formal training and hold a university diploma or degree-level qualification. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

## 6.2 Firm-specific, ward-specific and team-specific human capital

### 6.2.1 Firm-specific human capital

Table 3 focuses on the impact of firm-specific experience, after controlling for the quantity and seniority mix of staffing. All estimates include the NA fill-rate and the RN fill-rate, the latter separated by pay band, so these results are conditional on the quantity and general human capital (seniority mix) of the team. Estimates are shown graphically in Appendix Figure A2.

The first column shows that a greater amount of firm-specific experience among the team as a whole – as measured by mean years of experience working in the hospital – is associated with better team performance. The estimate is statistically significant at the 10% level. Adding firm-specific experience does not change the results for quantity of general human capital as the results for general human capital are similar to those in the second column of Table 2. The second column separates returns from firm specific experience for degree qualified nurses from that of nursing assistants. Increasing the average firm-specific experience among RNs in the team by one year reduces the odds of a patient death by 7.2% (equivalent to roughly two-thirds of the impact of adding an extra nurse). Increasing average firm-specific experience by one standard deviation (2.5 years, see Table 1) reduces the odds of a patient death by 17.0%. In contrast, a greater level of firm-specific experience among nursing assistants has no statistically significant impact. The third column of Table 3 indicates that the returns to firm-specific experience amongst degree qualified nurses increase more or less monotonically with experience, though the point estimates are not statistically significantly different from each other.

### 6.2.2 Ward-specific human capital

The first column of Table A11 examines the impact of recent work experience in the ward, conditional on the size and general human capital of the team, patient characteristics and time trends. The coefficients indicate that recent ward-specific experience – as measured by the number of shifts in the ward in the past 90 days – has no statistically significant effect. The second column examines whether there are returns to the average *share* of recent shifts that were in the ward and similarly finds no significant effects.

The third and fourth columns examine whether teams containing a new joiner perform differently to those which do not have individuals unfamiliar the ward.[23] There is

[23] A new joiner is an individual who is within their first 30 days on the ward, having not worked there in at least the previous 90 days. See Section 3.2.2 and the Appendix for a full set of definitions for all

Table 3: Impact of firm-specific experience on patient mortality

| | **Outcome: patient death** | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.003<br>(0.003)<br>[0.386] | 1.003<br>(0.003)<br>[0.331] | 1.003<br>(0.003)<br>[0.365] |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | | |
| Band 5 nurses | 0.989*<br>(0.006)<br>[0.082] | 0.989*<br>(0.006)<br>[0.062] | 0.988*<br>(0.006)<br>[0.067] |
| Band 6 nurses | 0.986**<br>(0.007)<br>[0.037] | 0.986**<br>(0.007)<br>[0.036] | 0.987**<br>(0.007)<br>[0.049] |
| Band 7 and 8 nurses | 0.979*<br>(0.011)<br>[0.066] | 0.980*<br>(0.011)<br>[0.078] | 0.980*<br>(0.011)<br>[0.088] |
| Average experience in the Trust (years) | 0.939*<br>(0.034)<br>[0.079] | | |
| Average experience in the Trust for NAs (years) | | 1.017<br>(0.015)<br>[0.246] | |
| Average experience in the Trust for RNs (years) | | 0.928***<br>(0.021)<br>[0.001] | |
| Percent of Registered Nurses (RNs) with: † | | | |
| 1-3 years Trust experience | | | 0.991**<br>(0.004)<br>[0.047] |
| 3-7 years Trust experience | | | 0.990**<br>(0.005)<br>[0.034] |
| 7-11 years Trust experience | | | 0.987***<br>(0.004)<br>[0.003] |
| 11-15 years Trust experience | | | 0.988**<br>(0.005)<br>[0.016] |
| 15+ years Trust experience | | | 0.985***<br>(0.005)<br>[0.005] |
| Patient volume and characteristics | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ |
| N | 18,922 | 18,922 | 18,922 |
| Clusters | 52 | 52 | 52 |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. Nursing assistants (NAs) are personnel without formal training or registration requirements; Registered Nurses (RNs) are fully qualified nurses on the Nursing and Midwifery Council register, who have completed formal training and hold a university diploma or degree-level qualification. See notes to Table 2 for a full list of controls.

† Reference group is those with less than one year of experience in the Trust.

a little evidence that teams containing a Band 5 new joiner have a higher risk of patient mortality (column 4: an odds ratio of 1.195, statistically significant at the 10% level) but the results broadly indicate that a lack of recent ward experience does not hinder team performance.

### 6.2.3 Team-specific experience

Table A10 examines the effect of familiarity of team members. The results in the first column indicate that the average amount of shared experience in the team has no statistically significant effect on team performance, for either nursing assistants or qualified nurses. The second column examines whether the average amount of shared experience between qualified nurses affects performance. Again there is no statistically significant effects on patient mortality. It is possible that the returns to team familiarity are non-linear, such that the negative effects from a lack of familiarity rapidly fall away with shared work experience. To test this, we examine the effects of the average fraction of the team who staff have ever worked with before (column 3) and the fraction of the team who staff have worked with 10 times or more (column 4). In both cases we find no statistically significant effect on mortality.

## 6.3 Robustness tests

We undertake several robustness tests. First, our main results condition on separate ward and month fixed effects. We also allow for ward-specific time trends by including ward-month fixed effects (Table A8) and ward-quarter fixed effects (Table A9). Our results are robust to these alternative specifications and the magnitude of our main estimates are unchanged. Second, we address the concern that we might not be adequately capturing characteristics of patients on the ward by repeating our main analyses at the patient-day level. This enables us to control for individual patient characteristics, rather than the ward-day average. Staffing measures are at the ward level and are calculated as the average across the wards in which the patient was treated that day, weighted by the number of minutes spent in each ward.[24] Table A13 presents the results. These show the same patterns as our team level analysis. Higher levels of qualified nurse staffing are associated with a lower probability of patient death, though this is only statistically significant (at the 10% level) for senior nurses in bands 7/8. Higher levels of firm-

variables used in the analysis.

[24] As the only variation in staffing across patients is at the ward level in our context, analysis at the patient level will provide little additional insight provided our controls for patient severity at the ward level are appropriate.

specific experience among qualified nurses is associated with reduced patient mortality, again statistically significant at the 10% level.

As a third test, we examine potential spillovers. Our baseline model includes only the contemporaneous impacts of team composition on the outcomes of patients. However, there may be dynamic effects within a team or spillovers onto other teams. One possible response to a team being short-staffed may be to send patients to other wards in the hospital, creating staffing problems for those wards. We show in Table A4 that there is no evidence that short staffing is associated with an increase in patient outflow. It could be that the negative effects of being treated by a short-staffed team do not materialise immediately and instead lead to the patient dying some time later under the care of a different team. We examine this possibility in Table A16 in which the outcome variable is a binary indicator equal to one if a patient treated in ward $i$ on day $t$ dies under the care of a different team within 1, 2 or 3 days. We find no evidence of an impact of team size or composition on subsequent deaths among patients treated by the team, indicating that there are limited spillovers of short-staffing onto other teams. Conducting the same analysis with longer time periods gives very similar results. We also test for effects of lagged staffing at both the team-level (Table A17) and patient-day level (Table A18), in each case finding no evidence that staffing levels and composition 1, 2, or 3 days previously had any statistically significant impact on patient mortality. This suggests no significant role for dynamic effects of staffing shortages.

Finally, we consider an alternative specification with a different implied team production function. Equation 2 implicitly assumes a Cobb-Douglas production function, where each type of labour – NA and RN, in the baseline model – is additively separable. We consider an alternative model of the form $G(L_{NA}, L_{RN}) = \beta_1 L_{NA} + \beta_2 L_{RN} + g(L_{NA}, L_{RN})$. To capture possible interactions between shortages in the different types of labour we let $g(L_{NA}, L_{RN}) = \gamma_1 L_{NA} L_{RN}$. We find no statistically significant interactions between shortages of the two broad nursing labour types.

In a further analysis, given our finding of higher returns to individuals in bands 7 and 8, who are also the ward managers, we allow for the absence of these senior nurse managers to have an additional effect on top of a shortage of (band 5 and 6) qualified nurses, by including an interaction term between the band 5 and 6 RN fill-rate and a dummy indicating the presence of a senior nurse on the ward. The coefficient on this interaction term is close to zero and not statistically significant. We therefore conclude that production is separable in the different types of nursing labour.

## 6.4 Potential pathways

### 6.4.1 Nursing sensitive patients and conditions

We do not have information on cause of death or nurse-patient interactions. We do, however, observe the severity, and the diagnosis, of those patients who die and use this to provide indirect evidence of potential pathways for our observed effects.

We first examine whether the impacts vary by the clinical severity of the patient. Finding a relationship could indicate where resource-constrained teams focus their efforts. We measure severity by the Elixhauser score. The median Elixhauser Comorbidity Index score among patients who die in hospital in our sample was 5. We define 'low-severity' patients as patients with an Elixhauser score of less than 5, and 'high-severity' patients as those with those with a score of more than 5.[25]

We then examine whether the impacts vary by patient diagnosis as this hints at the likely cause of death, which may be linked to (lack of) patient care. More than 80% of the patients who die in hospital in our sample are recorded with a primary diagnosis in one of four ICD-10 categories: infectious diseases, cancer, diseases of the circulatory system, or diseases of the respiratory system. At the more granular level, the modal primary diagnosis among patients who died was pneumonia (which falls within the category of diseases of the respiratory system), followed by sepsis (which falls within the category of infectious diseases).[26]

Sepsis is a common and potentially life-threatening condition, driven by an extreme immune response to an infection. It claims more lives in England each year than lung cancer (Health Education England, 2023). In the US, one-third of people who die in hospital had sepsis during that hospitalisation (Centre for Disease Control and Protection, 2023). In our sample, the equivalent figure is 26.3%. Early identification is key to successful treatment of sepsis (Torsvik et al., 2016), and English hospitals have widely adopted a National Early Warning Score (NEWS) aimed at improving detection of pa-

[25] 3.6% of teams experienced the death of a low-severity patient, and 3.0% experienced the death of a high-severity patient (Table A1). Low-severity patients account for 54.4% of deaths in our sample.

[26] 81.9% of patients who died had a primary diagnosis of an infectious disease (ICD-10 chapter 1), cancer (chapter 2), a disease of the circulatory system (chapter 9) or the respiratory system (chapter 10). Specifically, 13.8%, 17.4%, 27.2% and 23.5% of patients recorded as having died had a primary diagnosis in ICD-10 chapters 1, 2, 9 and 10, respectively. The modal *primary* disease diagnosis among patients who died was pneumonia (13.5% of the total), followed by sepsis (13.0%).Other common primary diagnoses for patients who died in our sample include pneumonitis due to inhalation of food and vomit (6.7% of deaths); acute myocardial infarction (heart attack, 5.7%); cerebral infarction (stroke, 4.8%); cranial injury (4.5%); heart failure (4.1%); and lung cancer (2.7%). 36.3% and 26.3% of patients who died had a primary *or* secondary diagnosis of pneumonia and sepsis, respectively and 2.4% and 1.8% of teams experienced a death of one of these patients, respectively (Table A1).

tient deterioration, with a particular focus on sepsis (Inada-Kim, 2022; NHS England, 2023). It often presents through rapid falls in blood pressure and increases in heart rate, both of which qualified nurses have a direct role in monitoring and, if necessary, escalating to a ward manager or doctor (Kleinpell et al., 2013). Shortages of nurses would therefore be expected to have greater impacts on the mortality outcomes of patients with sepsis, rather than those in hospital with a diagnosis of a circulatory condition, such as a heart attack or a stroke, where rapid clinical attention is required upon onset but mortality outcomes are less sensitive to subsequent nurse monitoring. We might also expect to find limited mortality impacts on cancer patients, for whom deaths typically represent the end of a longer pre-planned treatment pathway and are plausibly less sensitive to nursing care on any given day.

### 6.4.2 Results

In Table 4 we examine the impact of nurse staffing on deaths of low- and high-severity patients. The results (displayed graphically in Appendix Figure A3) indicate that the impact of a shortage of qualified nurses, and in particular of senior RNs, falls almost entirely on low-severity patients. This is consistent with short-staffed teams rationing care by focusing on the most (observably) sick patients (Bail and Grealish, 2016; Jones et al., 2015; Hegney et al., 2019). Other, less observably sick, patients then receive less care as a result, increasing their mortality rate. The finding that senior nurses are more important is consistent with those nurses being better at identifying patients with a greater mortality risk not predicted by readily observable characteristics such as age and recorded comorbidities. Table A19 similarly shows that the impacts of lower hospital-specific experience within the team also fall predominantly upon low-severity patients.

In Table 5 we examine the impacts on patients with different diagnoses. These results are also presented graphically in Appendix Figure A4. Panel A of the table presents the results for the disease categories where we might expect mortality outcomes to be particularly sensitive to nursing care and monitoring. We find evidence of a statistically significant impact on mortality outcomes for patients with an infectious disease (column 1) and within this category, our results are concentrated in, and of a greater magnitude for, patients with sepsis (column 2). For respiratory diseases (column 3), a shortages of the most senior staff has a statistically significant impact on mortality, though only at the 10% level. Further analysis (available upon request) suggests that this result is driven by patients with pneumonia, rather than long term chronic conditions such as

emphysema and bronchitis, which also fall within the category of respiratory diseases. Pneumonia is a condition for which the initial presentation can be more complex than in the case of sepsis, plausibly making experience more valuable.

Panel B of Table 5 presents the results for the disease categories where we expect nursing care to have a less immediate effect on death rates. These are deaths of patients with cancer (column 4), diseases of the circulatory system (column 5, which includes strokes and heart attacks), and other conditions (column 6). In each case, we find no statistically significant effects of absences of either qualified or unqualified nurses, with odds ratios close to 1.

We also conduct these analyses at the patient-day level. These again corroborate the team level analyses. Short staffing of qualified nurses increases rates of mortality only for low-severity patients, although this is only statistically significant for the most senior of the qualified nurses (Table A14). Shortages of qualified nurses of all levels of seniority result in an increased death rate of patients with sepsis (Table A15). There is no impact of qualified nurses shortages on deaths among patients without a sepsis diagnosis.

These patterns of heterogeneity by patient severity and diagnosis shed light on the mechanisms that may underlie the relationship we find between short-staffing and increased mortality. Our results are consistent with resource-constrained teams having less time to monitor patients or to make assessments of the risks to patients based on that monitoring. Deaths emerge for conditions such as sepsis where this monitoring is crucial for diagnosis and survival is very time sensitive to diagnosis and treatment. The effects of short-staffing are concentrated amongst those patients who have lower levels of clinical severity, suggesting that teams which are short-staffed focus their more limited resources on those patients who are sicker.[27] Such rationing has been shown to occur in a variety of healthcare settings, including triage in the Emergency Department (Aacharya et al., 2011), and over the weekend when staffing levels are reduced (Sutton et al., 2018).

[27]This holds also within patients with sepsis: results available upon request.

Table 4: Impact of nurse staffing on patient mortality: heterogeneity by patient severity

| | **Outcome:** | |
|---|---|---|
| | Death of low-severity patient† | Death of high-severity patient† |
| | (1) | (2) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.003<br>(0.005)<br>[0.544] | 1.005<br>(0.004)<br>[0.171] |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | |
| Band 5 nurses | 0.984*<br>(0.009)<br>[0.063] | 0.998<br>(0.011)<br>[0.851] |
| Band 6 nurses | 0.980**<br>(0.009)<br>[0.028] | 0.991<br>(0.010)<br>[0.384] |
| Band 7 and 8 nurses | 0.964**<br>(0.015)<br>[0.018] | 0.993<br>(0.017)<br>[0.680] |
| Patient and time controls | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ |
| N | 18,567 | 18,563 |
| Clusters | 51 | 51 |

† The median Elixhauser Cormorbidity Index among patients who died in our sample was 5. A low-severity patient is defined as one with an Elixhauser score of less than 5; a high-severity patient as one with an Elixhauser score of 5 or more.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital ward-day. Nursing assistants (NAs) are personnel without formal training or registration requirements; Registered Nurses (RNs) are fully qualified nurses on the Nursing and Midwifery Council register, who have completed formal training and hold a university diploma or degree-level qualification. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table 5: Impact of nurse staffing on patient mortality: heterogeneity by diagnosis

| **Panel A** | **Death of patient with ICD-10 diagnosis:** | | |
|---|---|---|---|
| | in chapter 1 (infection) | of sepsis† | in chapter 10 (respiratory) |
| | (1) | (2) | (3) |
| Band 5 RN fill-rate | 0.970** | 0.964*** | 0.984 |
| | (0.015) | (0.010) | (0.011) |
| | [0.043] | [<0.001] | [0.148] |
| Band 6 RN fill-rate | 0.958** | 0.952*** | 0.988 |
| | (0.016) | (0.011) | (0.011) |
| | [0.011] | [<0.001] | [0.278] |
| Band 7–8 RN fill-rate | 0.973 | 0.966** | 0.957* |
| | (0.025) | (0.016) | (0.024) |
| | [0.277] | [0.043] | [0.078] |
| NA staffing levels | ✓ | ✓ | ✓ |
| Patient and time controls | ✓ | ✓ | ✓ |
| Ward fixed effects | ✓ | ✓ | ✓ |
| N | 14,566 | 17,840 | 14,928 |
| Clusters (wards) | 40 | 49 | 41 |
| % of deaths in sample | 13.8 | 26.3 | 23.5 |

| **Panel B** | **Death of patient with ICD-10 diagnosis:** | | |
|---|---|---|---|
| | in chapter 2 (cancer) | in chapter 9 (circulatory) | in another chapter |
| | (4) | (5) | (6) |
| Band 5 RN fill-rate | 0.994 | 1.016 | 0.999 |
| | (0.021) | (0.016) | (0.014) |
| | [0.776] | [0.320] | [0.919] |
| Band 6 RN fill-rate | 0.977 | 1.018 | 0.993 |
| | (0.019) | (0.016) | (0.018) |
| | [0.226] | [0.263] | [0.685] |
| Band 7–8 RN fill-rate | 0.986 | 0.988 | 1.000 |
| | (0.030) | (0.021) | (0.019) |
| | [0.655] | [0.550] | [0.991] |
| NA staffing levels | ✓ | ✓ | ✓ |
| Patient and time controls | ✓ | ✓ | ✓ |
| Ward fixed effects | ✓ | ✓ | ✓ |
| N | 13,100 | 12,378 | 15,296 |
| Clusters (wards) | 36 | 34 | 42 |
| % of deaths in sample | 17.4 | 27.2 | 18.2 |

†A patient is coded as having sepsis if they have a primary or secondary ICD-10 diagnosis code of A400, A403, A409, A410, A411, A414, A415, A418, or A419.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital ward-day. Nursing assistants (NAs) are personnel without formal training or registration requirements; Registered Nurses (RNs) are fully qualified nurses on the Nursing and Midwifery Council register, who have completed formal training and hold a university diploma or degree-level qualification. All regressions control for: NA fill-rate; the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

# 7 Conclusion

This paper examines the role of team production in healthcare, a labour- and teamwork-intensive sector that accounts for a substantial fraction of employment in many high income economies. We focus on teams of nurses, for whom collaboration, communication and knowledge sharing are crucial. Employing a novel high-frequency dataset that links electronic staffing rotas to inpatient mortality records in a large NHS hospital group in England, we examine the impact of both the size and composition of nursing teams on team productivity, as measured by inpatient mortality rates.

We use qualifications and grade as a proxy for general human capital, and find that teams with higher levels of general human capital are more productive. An additional nursing assistant has no impact on the likelihood of a patient death, but a greater number of hours worked by degree-qualified nurses is associated with lower inpatient mortality, with the magnitude of these effects increasing with qualifications and experience. The most senior nurses, who are responsible for managing their ward, are around twice as productive as their newly qualified counterparts. Notably, this differential is somewhat wider than the pay differential between these groups.[28]

Furthermore, our results suggest that teams composed of workers with more firm-specific human capital perform better (consistent with Huckman and Pisano (2006)) and support the idea that familiarity with hospital-specific systems and processes can yield substantial benefits. The fact that there are returns to firm-specific experience for degree-qualified nurses but not nursing assistants suggests that that general and firm-specific human capital are complementary. We find no clear evidence of an effect of either plant-specific (ward-specific) experience or familiarity with team members. This latter result may be due to the fact that assignment to different wards is relatively frequent and so individuals work in various settings and with different team members relatively often. It may also indicate that hospital level protocols have been developed to allow for these frequent rotations and therefore what is important is knowing these.

Our results are largely driven by mortality impacts on relatively low-severity, rather than high-severity, patients and by impacts on patients with sepsis, a condition for which nurse monitoring and assessment are crucial for timely diagnosis and treatment. This is consistent with the idea that short-staffed teams or teams lacking in senior staff are able

[28] According to the published annual pay scales, a Band 7 or Band 8a nurse in the English NHS 2022/23 earned, on average, around £50,000, depending on experience (with the lowest-paid Band 7 earning around £42,000 and the highest-paid Band 8a nurse earning around £55,000). A Band 5 nurse earned between £27,000 and £33,000 (depending on experience – a Band 5 nurse with two years of experience could expect to earn around £29,000). This implies an average pay differential substantially lower than 2.

to monitor patients less intensively and/or focus their efforts on patients who are less (observably) sick. Hospitals implement protocols such as the NEWS score to improve early detection of patients whose condition is deteriorating: our analysis suggests that insufficient levels of nurse staffing could undermine the effectiveness of such protocols.

Our finding that less sick patients are more likely to die when teams are short-staffed or lack the most qualified nurses has implications for the total costs (in terms of lost life) from nursing shortages. In particular, our results do not suggest that shortages of nurses are simply bringing forward deaths of patients that had a high chance of dying in the near future in any case. Instead, deaths are occurring among those who (on observable characteristics at least) were less severely ill and had a lower risk of dying soon, or who might have survived with better monitoring and early detection. This implies a greater overall cost to social welfare.

A simple back-of-the-envelope calculation, provided in Appendix A.5, combines our estimates with the value of a (quality-adjusted) life year, as used by the body in the UK which assesses whether medicines should be provided by the NHS (the National Institute for Health and Care Excellence, NICE), to assess the cost-effectiveness of higher nurse staffing as a mortality-reduction strategy. This suggests that each patient who died in our sample – among whom the median age was 77 – who might have otherwise survived with higher levels of nursing staff, would need to have otherwise lived for a further three or four years for the elimination of all short staffing to be a cost-effective intervention. However, the value of life implicitly used by NICE is quite low (£20,000–30,000): using a higher value comparable to that often used in the USA (100,000 USD) would suggest that this would be a cost-effective strategy if patients who would have died in the absence of the intervention instead go on to live for an additional two or so years.

A more general conclusion of our analysis is that nurse staffing shortages, increasingly common in many healthcare systems, have adverse consequences for patient care but that efforts to substitute less skilled individuals for more highly skilled individuals are unlikely to boost team performance. This echoes a recent findings for the substitution of nurse practitioners for doctors in the VA system in the US (Chan and Chen, 2022). Given the returns to both general and hospital experience that we find, one alternative strategy to increase the quality of care provided by nursing teams would be to increase efforts to retain their more skilled staff members.

## References


Aacharya, R. P., C. Gastmans, and Y. Denier (2011, December). Emergency department triage: An ethical analysis. BMC Emergency Medicine 11(1).

Agha, L., K. M. Ericson, K. H. Geissler, and J. B. Rebitzer (2021, September). Team Relationships and Performance: Evidence from Healthcare Referral Networks. Management Science, mnsc.2021.4091.

Ali, A. J., J. Fuenzalida, M. Gómez, and M. J. Williams (2021, June). Four lenses on people management in the public sector: An evidence review and synthesis. Oxford Review of Economic Policy 37(2), 335–366.

Arcidiacono, P., J. Kinsler, and J. Price (2017, January). Productivity Spillovers in Team Production: Evidence from Professional Basketball. Journal of Labor Economics 35(1), 191–225.

Bail, K. and L. Grealish (2016, November). 'Failure to Maintain': A theoretical proposition for a new quality indicator of nurse care rationing for complex older people in hospital. International Journal of Nursing Studies 63, 146–161.

Ball, J. E., L. Bruyneel, L. H. Aiken, W. Sermeus, D. M. Sloane, A. M. Rafferty, R. Lindqvist, C. Tishelman, and P. Griffiths (2018, February). Post-operative mortality, missed care and nurse staffing in nine countries: A cross-sectional study. International Journal of Nursing Studies 78, 10–15.

Ball, J. E., T. Murrells, A. M. Rafferty, E. Morrow, and P. Griffiths (2014, February). Care left undone' during nursing shifts: Associations with workload and perceived quality of care. BMJ Quality & Safety 23(2), 116–125.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social Connections and Incentives in the Workplace: Evidence From Personnel Data. Econometrica 77(4), 1047–1094.

Bandiera, O., I. Barankay, and I. Rasul (2013). Team Incentives: Evidence from a Firm Level Experiment. Journal of the European Economic Association 11(5), 1079–1114.

Bartel, A. P., N. D. Beaulieu, C. S. Phibbs, and P. W. Stone (2014, April). Human Capital and Productivity in a Team Environment: Evidence from the Healthcare Sector. American Economic Journal: Applied Economics 6(2), 231–259.

Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition. NBER.

Burgess, S., C. Propper, M. Ratto, S. v. H. Kessler Scholder, and E. Tominey (2010, September). Smarter Task Assignment or Greater Effort: The Impact of Incentives on Team Performance. The Economic Journal 120(547), 968–989.

Burgess, S., C. Propper, M. Ratto, and E. Tominey (2017, October). Incentives in the Public Sector: Evidence from a Government Agency. The Economic Journal 127(605), F117–F141.

Centre for Disease Control and Protection (2023). What is Sepsis?

Chan, D. C., M. Gentzkow, and C. Yu (2022, April). Selection with Variation in Diagnostic Skill: Evidence from Radiologists. The Quarterly Journal of Economics 137(2), 729–783.

Chan, David C, J. and Y. Chen (2022, October). The Productivity of Professions: Evidence from the Emergency Department. Working Paper 30608, National Bureau of Economic Research.

Chen, Y. (2021, December). Team-Specific Human Capital and Team Performance: Evidence from Doctors. American Economic Review 111(12), 3923–3962.

Cornelissen, T., C. Dustmann, and U. Schönberg (2017, February). Peer Effects in the Workplace. American Economic Review 107(2), 425–456.

Currie, J. and W. B. MacLeod (2017). Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians. Journal of Labor Economics 35(1), 18977.

Currie, J. M. and W. B. MacLeod (2020). Understanding Doctor Decision Making: The Case of Depression Treatment. Econometrica 88(3), 847–878.

Deloitte (2016). Global Human Capital Trends 2016. https://www2.deloitte.com/us/en/insights/focus/human-capital-trends/2016.html.

Deloitte (2021). 2021 Global Human Capital Trends: Special report. https://www2.deloitte.com/us/en/insights/focus/human-capital-trends.html.

Doyle, J. J., S. M. Ewer, and T. H. Wagner (2010, December). Returns to physician human capital: Evidence from patients randomized to physician teams. Journal of Health Economics 29(6), 866–882.

Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey (1998, January). Comorbidity Measures for Use with Administrative Data. Medical Care 36(1), 8–27.

Fischer, K., J. J. Reade, and W. B. Schmal (2021). The Long Shadow of an Infection: COVID-19 and Performance at Work. Working Paper.

Friebel, G., M. Heinz, M. Krueger, and N. Zubanov (2017, August). Team Incentives and Performance: Evidence from a Retail Chain. American Economic Review 107(8), 2168–2203.

Friedrich, B. U. and M. B. Hackmann (2021, October). The Returns to Nursing: Evidence from a Parental-Leave Program. The Review of Economic Studies 88(5), 2308–2343.

Gennaioli, N., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer (2013, February). Human Capital and Regional Development. The Quarterly Journal of Economics 128(1), 105–164.

Griffiths, P., A. Maruotti, A. Recio Saucedo, O. C. Redfern, J. E. Ball, J. Briggs, C. Dall'Ora, P. E. Schmidt, and G. B. Smith (2019, August). Nurse staffing, nursing assistants and hospital mortality: Retrospective longitudinal cohort study. BMJ Quality & Safety 28(8), 609–617.

Gruber, J. and S. A. Kleiner (2012, February). Do Strikes Kill? Evidence from New York State. American Economic Journal: Economic Policy 4(1), 127–157.

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009, September). Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments. American Economic Journal: Applied Economics 1(4), 34–68.

Hamilton, B. H., J. A. Nickerson, and H. Owan (2003, June). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. Journal of Political Economy 111(3), 465–497.

Health Education England (2023). Sepsis Awareness.

Hegney, D. G., C. S. Rees, R. Osseiran-Moisson, L. Breen, R. Eley, C. Windsor, and C. Harvey (2019, March). Perceptions of nursing workloads and contributing factors, and their impact on implicit care rationing: A Queensland, Australia study. Journal of Nursing Management 27(2), 371–380.

Herrmann, M. A. and J. E. Rockoff (2012, October). Worker Absence and Productivity: Evidence from Teaching. Journal of Labor Economics 30(4), 749–782.

Hoey, S., T. Peeters, and J. van Ours (2022). The Impact of Absent Coworkers on Productivity in Teams. SSRN Electronic Journal.

Huckman, R. S. and G. P. Pisano (2006, April). The Firm Specificity of Individual Performance: Evidence from Cardiac Surgery. Management Science 52(4), 473–488.

Inada-Kim, M. (2022, November). NEWS2 and improving outcomes from sepsis. Clinical Medcine 22(6), 514–517.

Jaravel, X., N. Petkova, and A. Bell (2018, April). Team-Specific Capital and Innovation. American Economic Review 108(4-5), 1034–1073.

Jarosch, G., E. Oberfield, and E. Rossi-Hansberg (2021). Learning From Coworkers. Econometrica 89(2), 647–676.

Jones, B. F. (2021, May). The Rise of Research Teams: Benefits and Costs in Economics. Journal of Economic Perspectives 35(2), 191–216.

Jones, T. L., P. Hamilton, and N. Murry (2015, June). Unfinished nursing care, missed care, and implicitly rationed care: State of the science review. International Journal of Nursing Studies 52(6), 1121–1137.

Kleinpell, R., L. Aitken, and C. A. Schorr (2013, May). Implications of the New International Sepsis Guidelines for Nursing Care. American Journal of Critical Care 22(3), 212–222.

Lazear, E. P. and K. L. Shaw (2007, December). Personnel Economics: The Economist's View of Human Resources. Journal of Economic Perspectives 21(4), 91–114.

Lin, H. (2014). Revisiting the relationship between nurse staffing and quality of care in nursing homes: An instrumental variables approach. Journal of Health Economics, 12.

Mas, A. and E. Moretti (2009, February). Peers at Work. American Economic Review 99(1), 112–145.

Mincer, J. A. (1974). Schooling, Experience, and Earnings. NBER.

Moretti, E. (2004, May). Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions. American Economic Review 94(3), 656–690.

Needleman, J., C. L. Leibson, and M. Harris (2011). Nurse Staffing and Inpatient Hospital Mortality. The New England Journal of Medicine, 10.

NHS Digital (2021a). NHS Vacancy Statistics - September 2021.

NHS Digital (2021b). NHS Workforce Statistics - October 2021.

NHS England (2023). NHS England - National Early Warning Score.

Parsons, D. O. (1972, November). Specific Human Capital: An Application to Quit Rates and Layoff Rates. Journal of Political Economy 80(6), 1120–1143.

Propper, C. and J. Van Reenen (2010, April). Can Pay Regulation Kill? Panel Data Evidence on the Effect of Labor Markets on Hospital Performance. Journal of Political Economy 118(2), 222–273.

Redfern, O. C., P. Griffiths, A. Maruotti, A. Recio Saucedo, and G. B. Smith (2019, September). The association between nurse staffing levels and the timeliness of vital signs monitoring: A retrospective observational study in the UK. BMJ Open 9(9), e032157.

Silver, D. (2021, May). Haste or Waste? Peer Pressure and Productivity in the Emergency Department. The Review of Economic Studies 88(3), 1385–1417.

Smith, G. B., O. Redfern, A. Maruotti, A. Recio-Saucedo, and P. Griffiths (2020, April). The association between nurse staffing levels and a failure to respond to patients with deranged physiology: A retrospective observational study in the UK. Resuscitation 149, 202–208.

Stuart, H. C. (2017, April). Structural Disruption, Relational Experimentation, and Performance in Professional Hockey Teams: A Network Perspective on Member Change. Organization Science 28(2), 283–300.

Sutton, E., J. Bion, C. Aldridge, A. Boyal, J. Willars, and C. Tarrant (2018, December). Quality and safety of in-hospital care for acute medical patients at weekends: A qualitative study. BMC Health Services Research 18(1).

Torsvik, M., L. T. Gustad, A. Mehl, I. L. Bangstad, L. J. Vinje, J. K. Damås, and E. Solligård (2016, December). Early identification of sepsis in hospital inpatients by ward nurses increases 30-day survival. Critical Care 20(1).

Wuchty, S., B. F. Jones, and B. Uzzi (2007, May). The Increasing Dominance of Teams in Production of Knowledge. Science 316(5827), 1036–1039.

# A Appendix

## A.1 Summary statistics

Table A1: Summary statistics: all variables used in team-level analysis

| | Mean | SD |
|---|---|---|
| **Staffing: quantity and seniority** | | |
| Fill-rate (all) | 95.7 | 5.4 |
| NA fill-rate | 94.8 | 10.9 |
| RN fill-rate | 96.4 | 5.8 |
| Local RN fill-rate | 70.5 | 16.9 |
| Bank RN fill-rate | 16.3 | 12.7 |
| Agency RN fill-rate | 9.5 | 11.0 |
| Band 5 RN fill-rate | 68.4 | 15.8 |
| Band 5 non-agency RN fill-rate | 59.2 | 16.2 |
| Band 5 agency RN fill-rate | 9.2 | 11.0 |
| Band 6 RN fill-rate | 22.8 | 13.3 |
| Band 7–8 RN fill-rate | 5.2 | 5.8 |
| **Staffing: firm-specific experience** | | |
| Mean years of experience in the hospital group (all) | 6.3 | 2.0 |
| Mean years of experience in the hospital group (NAs) | 5.6 | 3.2 |
| Mean years of experience in the hospital group (RNs) | 6.7 | 2.5 |
| % of RNs with < 1 year of experience in hospital group | 23.1 | 15.5 |
| % of RNs with 1–3 years of experience in hospital group | 19.5 | 14.1 |
| % of RNs with 3–7 years of experience in hospital group | 16.9 | 12.4 |
| % of RNs with 7-11 years of experience in hospital group | 11.3 | 11.6 |
| % of RNs with 11–15 years of experience in hospital group | 14.3 | 13.7 |
| % of RNs with 15+ years of experience in hospital group | 14.9 | 12.7 |
| **Staffing: ward-specific experience** | | |
| Mean shifts in ward in past 90 days (NAs) | 27.8 | 9.7 |
| Mean shifts in ward in past 90 days (RNs) | 31.5 | 5.9 |
| Mean % of shifts in past 90 days in the ward (NAs) | 74.6 | 21.5 |
| Mean % of shifts in past 90 days in the ward (RNs) | 83.4 | 12.2 |
| NA new joiner in the team | 0.57 | 0.50 |
| RN new joiner in the team | 0.71 | 0.46 |
| Band 5 RN new joiner in the team | 0.68 | 0.47 |
| Band 6 RN new joiner in the team | 0.08 | 0.28 |
| Band 7–8 RN new joiner in the team | 0.01 | 0.12 |

| | Mean | SD |
|---|---|---|
| **Staffing: team-specific experience** | | |
| Mean shifts with other team members in past 90 days (NAs) | 13.5 | 6.3 |
| Mean shifts with other team members in past 90 days (RNs) | 14.9 | 5.1 |
| Mean shifts with other RNs in past 90 days (RNs) | 15.2 | 5.7 |
| Mean % of team members worked with before (NAs) | 82.5 | 16.3 |
| Mean % of team members worked with before (RNs) | 86.0 | 11.9 |
| Mean % of team members worked with 10+ times before (NAs) | 52.8 | 24.8 |
| Mean % of team members worked with 10+ times before (RNs) | 59.5 | 20.2 |
| **Patient characteristics** | | |
| Age | 64.18 | 9.86 |
| Female % | 0.46 | 0.18 |
| Non-white %† | 0.52 | 0.16 |
| Elixhauser Comorbidity Index | 3.69 | 1.17 |
| Length of hospital stay (days)‡ | 28.53 | 17.08 |
| Number of patients treated in the ward | 20.61 | 8.46 |
| **Patient outcomes** | | |
| Death (binary) | 0.063 | 0.243 |
| Death rate per 1,000 patients | 3.58 | 16.20 |
| Death of low-severity patient (binary)* | 0.036 | 0.185 |
| Death of high-severity patient (binary)* | 0.030 | 0.170 |
| Death with primary diagnosis in ICD-10 chapter 1 (binary) | 0.009 | 0.096 |
| Death with primary diagnosis in ICD-10 chapter 2 (binary) | 0.009 | 0.106 |
| Death with primary diagnosis in ICD-10 chapter 9 (binary) | 0.018 | 0.132 |
| Death with primary diagnosis in ICD-10 chapter 10 (binary) | 0.015 | 0.123 |
| Death with primary diagnosis in other ICD-10 chapter (binary) | 0.012 | 0.109 |
| Death of patient with sepsis diagnosis (binary)** | 0.018 | 0.131 |
| Death of patient with pneumonia diagnosis (binary) ** | 0.024 | 0.153 |
| Observations (teams) | 18,922 | |
| Clusters (hospital wards) | 52 | |

† Non-white includes patients with ethnicity recorded as 'unknown'. ‡ Calculated as the average for the patients treated in a given unit on a given day (so that patients with a stay of more than one day are counted multiple times), rather than the average for the universe of patients. * Low-severity patients are defined as those with an Elixhauser index score of less than 5; high-severity patients are defined as those with an Elixhauser index score of at least 5. ** A patient is coded as having sepsis if they have a three-digit ICD-10 primary or secondary diagnosis code of A40 or A41. A patient is coded as having pneumonia if they have a three-digit ICD-10 primary or secondary diagnosis of J18, J18, J18 or J18.

## A.2 Definitions

**Human capital in a nursing context**

Let $T_{kt}$ be a team consisting of $n$ members working in location (hospital ward) $k$ on day $t$. $|T_{kt}|$ is the number of staff in the team, which is not fixed and varies across both wards and time. $|T^*_{kt}|$ is the optimal size of the team. We allow the productivity of an individual $i$ to depend on their innate ability (denoted by $\theta_i$) and four types of human capital, each of which is accumulated through different types of experience: general, firm-specific, plant-specific and team-specific capital. Specifically, let the productivity $P$ of individual $i$, working in firm (hospital) $h$, in ward $k$ on day $t$ be equal to:

$$P_{ihkt} = \quad f(\theta_i, W_{it}, X_{iht}, Y_{ihkt}, Z_{-ihkt})$$

where:

- $W_{it}$ **general human capital**
- $X_{iht}$ **firm-specific human capital**: the individual's $i$'s experience working in firm (hospital) $h$ at date $t$.
- $Y_{ihkt}$ **plant-specific human capital**: individual $i$'s experience working in plant (ward) $k$ within hospital $h$ at date $t$.
- $Z_{-ihkt}$ **team-specific human capital**: individual $i$'s shared work experience with the other individuals working in ward $k$ within hospital $h$ at date $t$ (i.e. with the other team members present on that date, in that hospital ward).

**Staffing quantity and general human capital definitions**

We measure the amount of general human capital, $W_{it}$, in the team by the pay band of each team member, and construct a range of measures to capture this:

$$FillRateNA_{it} = 100 * \frac{\text{Hours worked by NAs in ward } i \text{ on day } t}{\text{Planned NA hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it} = 100 * \frac{\text{Hours worked by RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{Band5} = 100 * \frac{\text{Hours worked by Band 5 RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{Band6} = 100 * \frac{\text{Hours worked by Band 6 RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{Band7and8} = 100 * \frac{\text{Hours worked by Band 7 and 8 RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{Local} = 100 * \frac{\text{Hours worked by Local (regular) RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{Bank} = 100 * \frac{\text{Hours worked by Bank RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{Agency} = 100 * \frac{\text{Hours worked by Agency RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{LocalBand5} = 100 * \frac{\text{Hours worked by Local Band 5 RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{BankBand5} = 100 * \frac{\text{Hours worked by Bank Band 5 RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

$$FillRateRN_{it}^{AgencyBand5} = 100 * \frac{\text{Hours worked by Agency Band 5 RNs in ward } i \text{ on day } t}{\text{Planned RN hours in ward } i \text{ on day } t}$$

**Firm-specific human capital**

$X_{iht}$ is defined (as per Section 2) as the amount of time (in years) that individual $i$ has been employed in the NHS Trust at date $t$. We then define:

$$\text{Average experience in the hospital group} = \overline{X}_{kt} = \frac{1}{|T_{kt}|} \sum_{i \in T_{kt}} X_{iht}$$

$$\text{Average experience in the hospital group for RNs} = \overline{X}_{kt}^{RN} = \frac{1}{|T_{kt}^{RN}|} \sum_{i \in T_{kt}^{RN}} X_{iht}$$

$$\text{Average experience in the hospital group for NAs} = \overline{X}_{kt}^{NA} = \frac{1}{|T_{kt}^{NA}|} \sum_{i \in T_{kt}^{NA}} X_{iht}$$

where $T_{kt}$ is the team working in ward $k$ on day $t$, $T_{kt}^{RN}$ is the set of RNs working in ward $k$ on day $t$, and $T_{kt}^{NA}$ is the set of NAs working in ward $k$ on day $t$.

**Ward-specific human capital**

Let $T_{kt}$ be the team working in ward $k$ on day $t$, $T_{kt}^{RN}$ be the set of RNs working in ward $k$ on day $t$, and $T_{kt}^{NA}$ be the set of NAs working in ward $k$ on day $t$. Let $K(h)$ be the set of wards in hospital $h$. Then for all $k \in K(h)$ we define $Y_{ihkt}$ as the ward-specific human capital for worker $i$ in ward $k$ on day $t$ as:

$$Y_{ihkt} = \sum_{\tau=t-90}^{t-1} \mathbb{1}(i \in T_{\tau k}) \quad \forall k \in K(h)$$

where $\mathbb{1}(i \in T_{\tau k})$ is an indicator equal to one if individual $i$ was part of team $T_{\tau k}$ and zero otherwise. $Y_{ihkt}$ therefore represents the number of times the individual has worked in ward $k$ in the previous 90 days. We then further define:

$$\text{Individual } i \text{ all shifts in the past 90 days} = \widetilde{Y}_{it} = \sum_{k \in K(h)} \sum_{\tau=t-90}^{t-1} \mathbb{1}(i \in T_{\tau k})$$

$$\text{Mean shifts in ward in the past 90 days} = \overline{Y}_{kt} = \frac{1}{|T_{kt}|} \sum_{i \in T_{kt}} Y_{ihkt}$$

$$\text{Mean NA shifts in ward in the past 90 days} = \overline{Y}_{kt}^{NA} = \frac{1}{|T_{kt}^{NA}|} \sum_{i \in T_{kt}^{NA}} Y_{ihkt}$$

$$\text{Mean RN shifts in ward in the past 90 days} = \overline{Y}_{kt}^{RN} = \frac{1}{|T_{kt}^{RN}|} \sum_{i \in T_{kt}^{RN}} Y_{ihkt}$$

$$\text{Mean NA \% of shifts in ward in the past 90 days} = \frac{1}{|T_{kt}^{NA}|} \sum_{i \in T_{kt}^{NA}} 100 * (Y_{ihkt} / \widetilde{Y}_{it})$$

$$\text{Mean RN \% of shifts in ward in the past 90 days} = \frac{1}{|T_{kt}^{RN}|} \sum_{i \in T_{kt}^{RN}} 100 * (Y_{ihkt} / \widetilde{Y}_{it})$$

We then define a 'new joiner' as someone who is within their first month on a ward:

$$FirstShift_{ihkt} = \mathbb{1}(Y_{ihkt} = 0 \wedge i \in T_{kt})$$

$$NewJoiner_{ihkt} = \mathbb{1}(\sum_{\tau=t-30}^{t} FirstShift_{ikh\tau} > 0 \wedge i \in T_{kt})$$

where $NewJoiner_{ihkt}$ is defined separately for NAs and RNs, and for RNs of different bands.

**Team-specific human capital**

Let $T_{kt}$ be the team working in ward $k$ on day $t$, $T_{kt}^{RN}$ be the set of RNs working in ward $k$ on day $t$, and $T_{kt}^{NA}$ be the set of NAs working in ward $k$ on day $t$. Let $K(h)$ be the set of wards in hospital $h$. We define team-specific human capital, $Z_{-ihkt}$ as the average number of shifts that individual $i$ has worked together with their team mates, in any ward, over the past 90 days prior to day $t$. Let $E(i,j,t)$ represent the (recent) shared work experience between workers $i$ and $j$, as of date $t$, where:

$$E(i,j,t) = \sum_{k \in K(h)} \sum_{\tau=t-90}^{t-1} \mathbb{1}(i \in T_{\tau k} \wedge j \in T_{\tau k})$$

where $\mathbb{1}(i \in T_{\tau k} \wedge j \in T_{\tau k})$ is an indicator that equals one if both workers $i$ and $j$ worked in ward $k$ on day $\tau \in [t-90, t-1]$ (i.e. equals one if both $i$ and $j$ were in team $T_{\tau k}$). $E(i,j,t)$ therefore measures the number of shifts $i$ and $j$ have worked together. We then define an individual's average shared work experience with fellow team members (team-specific human capital) as:

$$Z_{-ihkt} = \frac{1}{|T_{kt}| - 1} \cdot \sum_{j \in T_{kt}, j \neq i} E(i,j,t)$$

We then additionally define:

Individual $i$'s average shared work experience with RNs $= \widetilde{Z}_{-ihkt} = \dfrac{1}{|T_{kt}^{RN}| - 1} \cdot \displaystyle\sum_{j \in T_{kt}^{RN}, j \neq i} E(i,j,t)$

Mean shifts with other team members in past 90 days $= \overline{Z}_{kt} = \dfrac{1}{|T_{kt}|} \displaystyle\sum_{i \in T_{kt}} Z_{-ihkt}$

Mean NA shifts with other team members in past 90 days $= \overline{Z}_{kt}^{NA} = \dfrac{1}{|T_{kt}^{NA}|} \displaystyle\sum_{i \in T_{kt}^{NA}} Z_{-ihkt}$

Mean RN shifts with other team members in past 90 days $= \overline{Z}_{kt}^{RN} = \dfrac{1}{|T_{kt}^{RN}|} \displaystyle\sum_{i \in T_{kt}^{RN}} Z_{-ihkt}$

Mean RN shifts with other RNs in past 90 days $= \dfrac{1}{|T_{kt}^{RN}|} \displaystyle\sum_{i \in T_{kt}^{RN}} \widetilde{Z}_{-ihkt}$

$$WorkedWith_{-ihkt} = \mathbb{1}(Z_{-ihkt} \geq 1 \wedge i \in T_{kt})$$

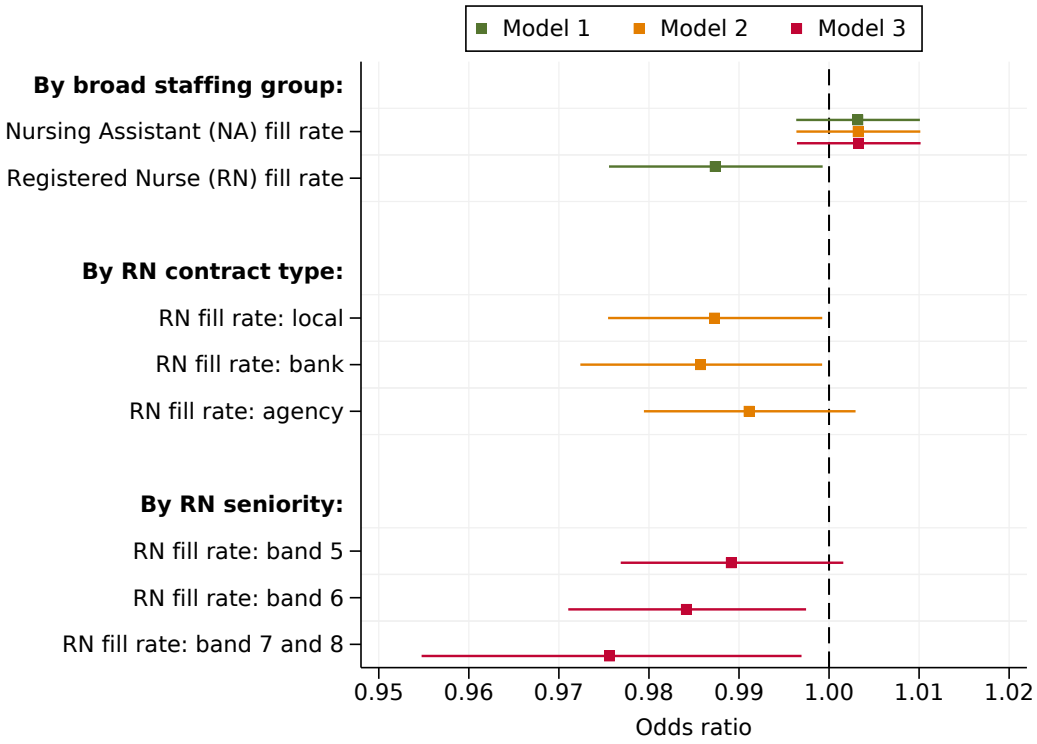
$$WorkedWith^{10}_{-ihkt} = \mathbb{1}(Z_{-ihkt} \geq 10 \wedge i \in T_{kt})$$

Mean % of team members worked with in past 90 days $= \frac{1}{|T_{kt}|} \sum_{i \in T_{kt}} WorkedWith_{-ihkt}$

Mean % of team members worked with 10+ times in past 90 days $= \frac{1}{|T_{kt}|} \sum_{i \in T_{kt}} WorkedWith^{10}_{-ihkt}$

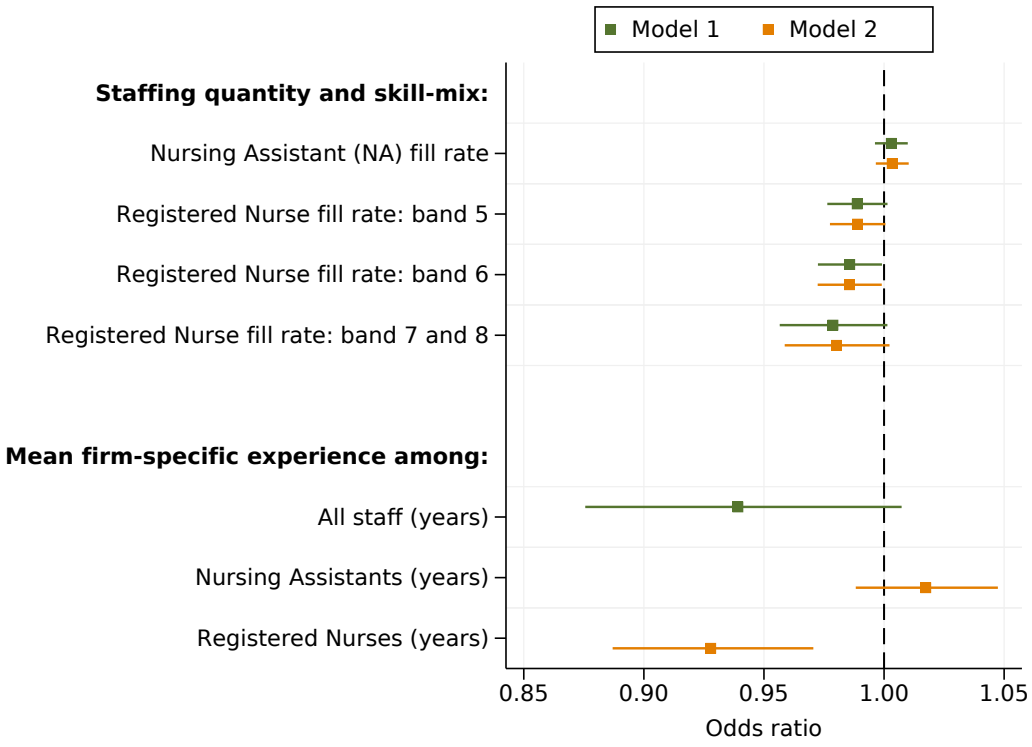
## A.3 Main results displayed as coefficient plots

Figure A1: Impact of nurse staffing levels, contract type and seniority on patient mortality



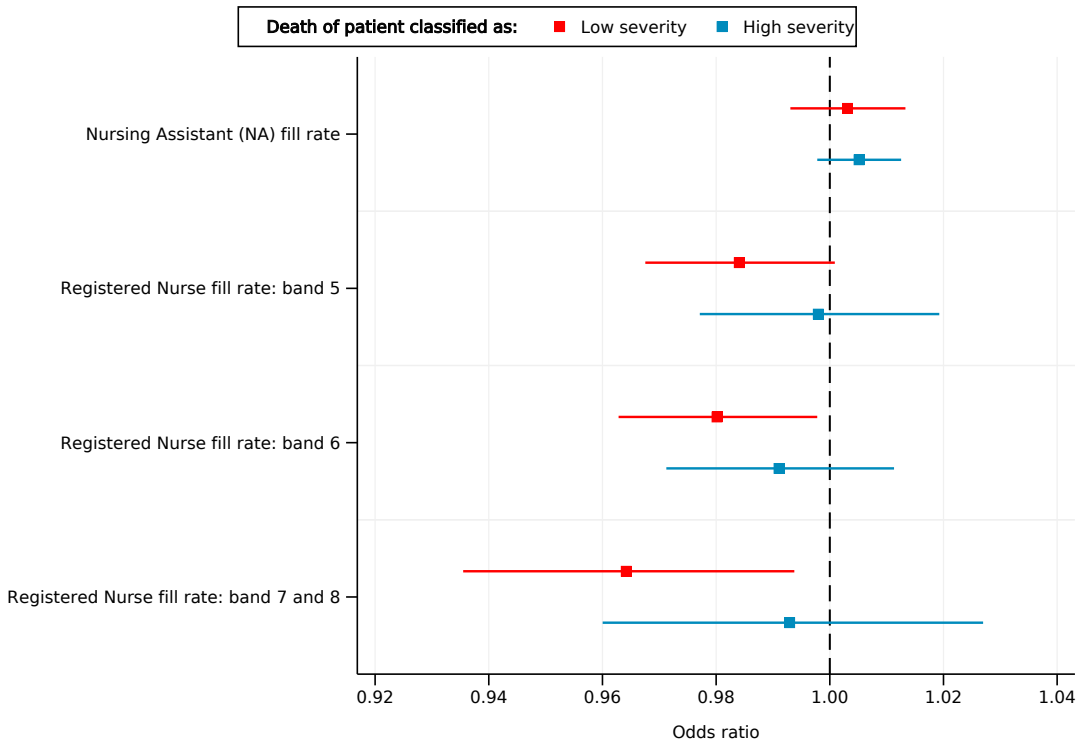

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. 95% confidence intervals (with robust standard errors) are shown. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Figure A2: Impact of firm-specific experience on patient mortality



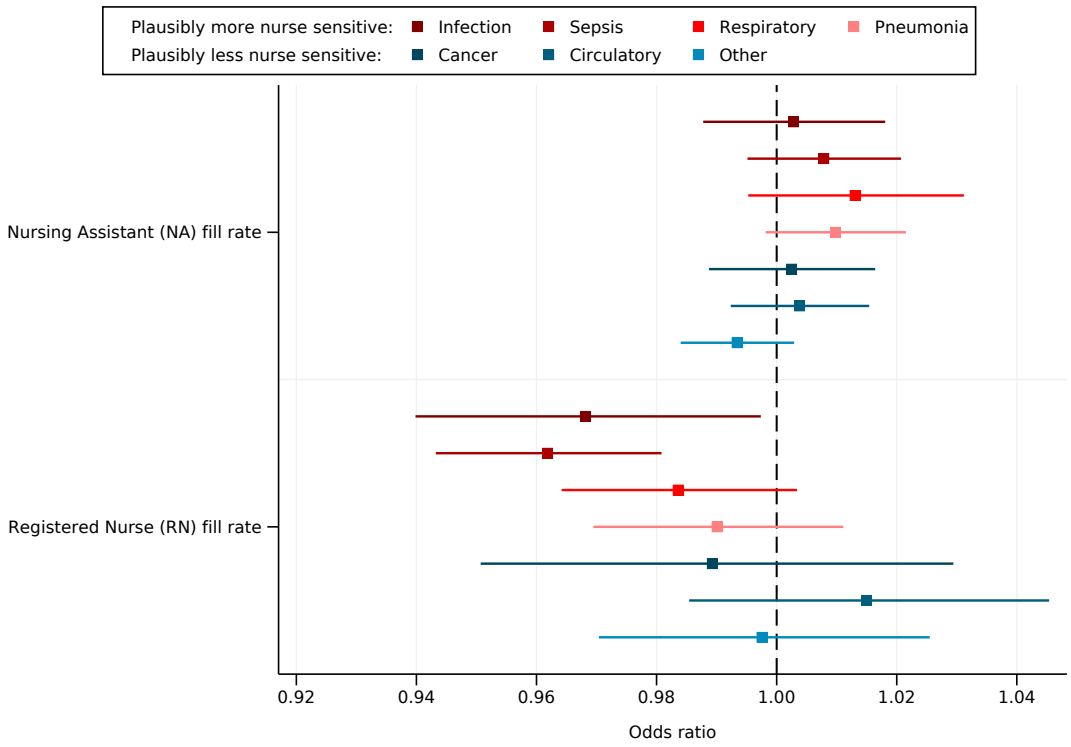

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. 95% confidence intervals (with robust standard errors) are shown. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Figure A3: Impact by patient severity




Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. 95% confidence intervals (with robust standard errors) are shown. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Figure A4: Impact by patient diagnosis




Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. 95% confidence intervals (with robust standard errors) are shown. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

## A.4 Additional results and figures

Table A2: Relationship between patient severity and team composition

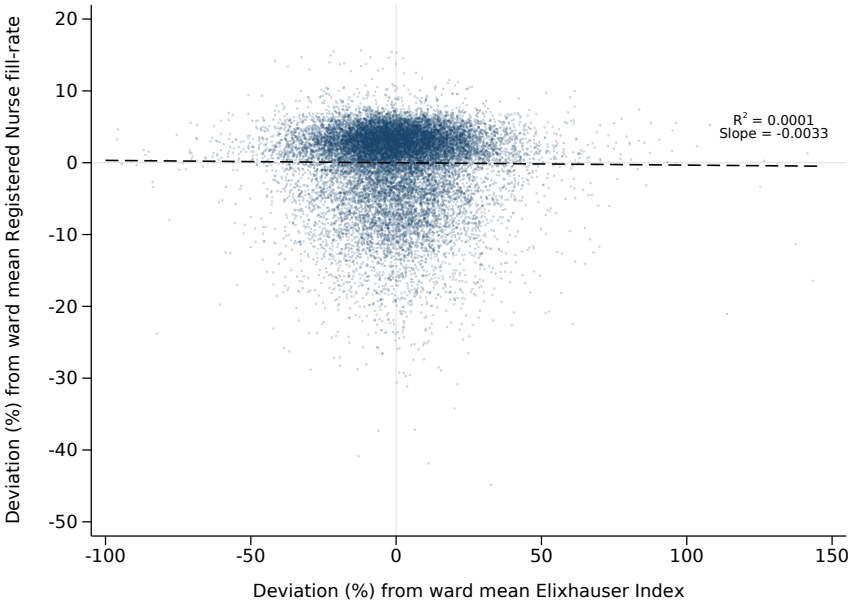
| | **Outcome: mean patient Elixhauser Cormorbidity Index** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | -0.001<br>(0.001)<br>[0.270] | -0.001<br>(0.001)<br>[0.275] | -0.001<br>(0.001)<br>[0.278] | -0.001<br>(0.001)<br>[0.277] |
| Share (ppt) of planned Registered Nurse (RN) hours filled | -0.000<br>(0.001)<br>[0.784] | | | |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | | | |
| Band 5 nurses | | -0.000<br>(0.001)<br>[0.732] | | |
| Band 6 nurses | | -0.001<br>(0.002)<br>[0.497] | | -0.001<br>(0.002)<br>[0.501] |
| Band 7 and 8 nurses | | 0.002<br>(0.003)<br>[0.418] | | 0.002<br>(0.003)<br>[0.414] |
| Local (regular) nurses | | | -0.000<br>(0.002)<br>[0.853] | |
| Bank nurses | | | -0.001<br>(0.002)<br>[0.538] | |
| Agency nurses | | | -0.000<br>(0.002)<br>[0.836] | |
| Band 5 Local and Bank nurses | | | | -0.001<br>(0.001)<br>[0.713] |
| Band 5 Agency nurses | | | | -0.000<br>(0.002)<br>[0.909] |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 18,922 | 18,922 | 18,922 | 18,922 |
| Clusters | 52 | 52 | 52 | 52 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

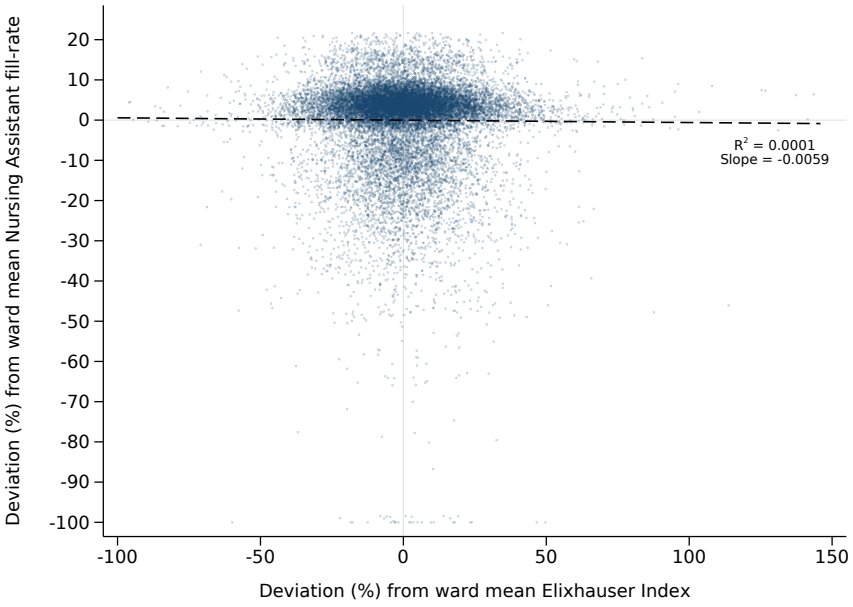
Coefficients shown are from a fixed-effects regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. Time controls include a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Figure A5: Relationship between staffing levels and observable patient comorbidity, as measured by Elixhauser Cormorbidity Index

(a) Registered Nurses

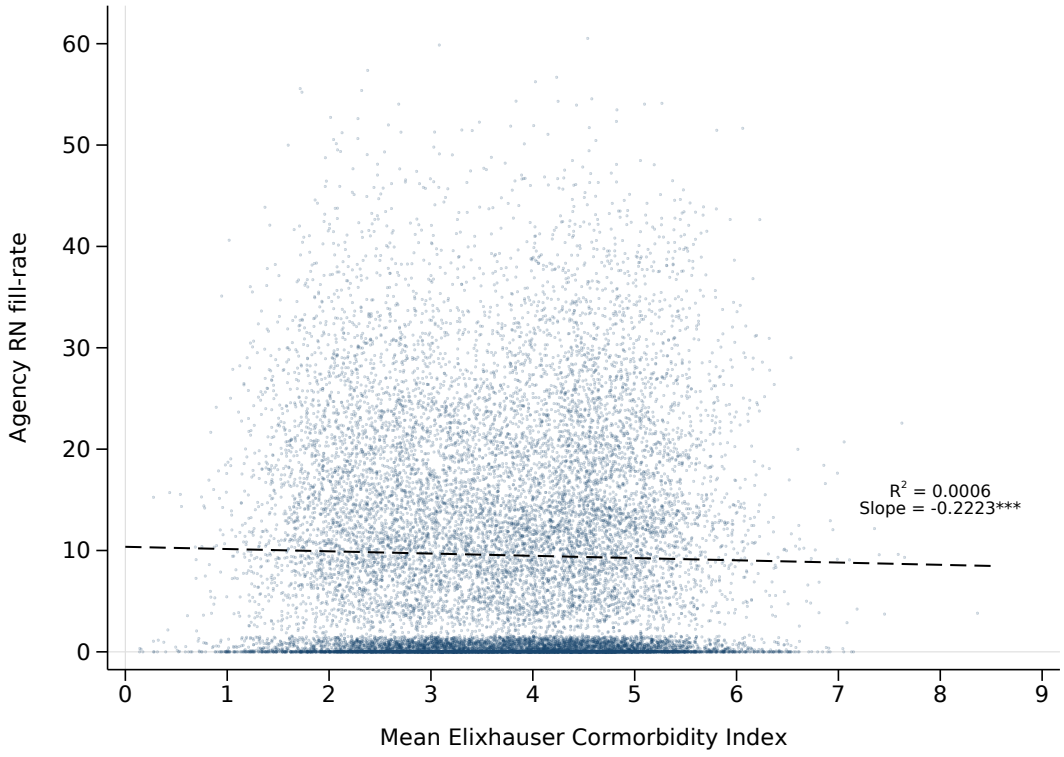



(b) Nursing Assistants




Note: each point denotes an observation (i.e. a team, defined as a hospital ward-day). Values are calculated relative to the unit ward, as our baseline specification includes ward fixed effects, and so all results are identified from within-ward variation.

Figure A6: Relationship between agency staffing and observable patient comorbidity, as measured by Elixhauser Cormorbidity Index




Note: each point denotes an observation (i.e. a team, defined as a hospital ward-day).

Table A3: Relationship between staffing and patient characteristics

| | **Staffing outcome:** | | |
|---|---|---|---|
| | Overall fill-rate | RN fill-rate | Agency RN fill-rate |
| | (1) | (2) | (3) |
| Mean age | -0.239* | -0.349** | -0.230* |
| | (0.129) | (0.171) | (0.121) |
| Mean age$^2$ | 0.002* | 0.002* | 0.001 |
| | (0.001) | (0.001) | (0.001) |
| Female share | -0.135 | -0.044 | 1.524 |
| | (3.456) | (4.647) | (3.397) |
| Non-white share | -1.226 | -0.817 | -1.159 |
| | (3.207) | (3.744) | (3.290) |
| Mean Elixhauser Comorbidity Index | 0.502 | 0.443 | -0.071 |
| | (0.632) | (0.808) | (0.570) |
| Age * Female | 0.026 | 0.034 | -0.004 |
| | (0.060) | (0.089) | (0.058) |
| Age * Non-white | 0.047 | 0.059 | 0.045 |
| | (0.042) | (0.053) | (0.051) |
| Age * Elixhauser | -0.004 | -0.003 | 0.002 |
| | (0.008) | (0.010) | (0.008) |
| Female * Non-white | -2.215 | -3.818 | -3.168 |
| | (1.960) | (2.704) | (2.177) |
| Female * Elixhauser | -0.084 | -0.042 | 0.048 |
| | (0.349) | (0.463) | (0.328) |
| Non-white * Elixhauser | -0.329 | -0.424 | -0.133 |
| | (0.300) | (0.365) | (0.305) |
| Number of patients in ward | 0.048 | 0.065 | 0.083*** |
| | (0.030) | (0.040) | (0.027) |
| Mean hospital length of stay | -0.001 | 0.002 | 0.013* |
| | (0.007) | (0.009) | (0.007) |
| NA fill-rate | × | ✓ | ✓ |
| Local and Bank RN fill-rate | × | × | ✓ |
| Time controls | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ |
| N | 18,922 | 18,922 | 18,922 |
| Adjusted R-squared | 0.386 | 0.042 | 0.728 |
| Joint F-statistic | 25.01 | 7.42 | 529.21 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Coefficients shown are from a fixed-effect regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses. The unit of analysis is the hospital ward-day. Time controls include a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A4: Impact of nurse staffing levels on patient outflow

| | **Outcome: proportion of patients moving ward** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 0.001<br>(0.005)<br>[0.827] | 0.001<br>(0.005)<br>[0.832] | 0.001<br>(0.005)<br>[0.800] | 0.001<br>(0.005)<br>[0.800] |
| Share (ppt) of planned Registered Nurse (RN) hours filled | 0.005<br>(0.009)<br>[0.600] | | | |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | | | |
| Band 5 nurses | | 0.005<br>(0.010)<br>[0.609] | | |
| Band 6 nurses | | -0.003<br>(0.011)<br>[0.781] | | -0.003<br>(0.011)<br>[0.769] |
| Band 7 and 8 nurses | | 0.023<br>(0.025)<br>[0.381] | | 0.022<br>(0.025)<br>[0.386] |
| Local (regular) nurses | | | 0.005<br>(0.009)<br>[0.595] | |
| Bank nurses | | | 0.007<br>(0.011)<br>[0.538] | |
| Agency nurses | | | -0.001<br>(0.010)<br>[0.937] | |
| Band 5 Local and Bank nurses | | | | 0.006<br>(0.010)<br>[0.563] |
| Band 5 Agency nurses | | | | -0.000<br>(0.010)<br>[0.982] |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 18,922 | 18,922 | 18,922 | 18,922 |
| Clusters | 52 | 52 | 52 | 52 |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are from a fixed-effect regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. All regressions control for a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A5: Impact of deaths on previous days on levels of nurse staffing

| | **Outcome:** | | |
|---|---|---|---|
| | NA fill-rate | RN fill-rate | Band 5 RN fill-rate |
| | (1) | (2) | (3) |
| Death in the ward on day T-1 | 0.054 | -0.232* | 0.094 |
| | (0.245) | (0.120) | (0.273) |
| | [0.827] | [0.058] | [0.731] |
| Death in the ward on day T-2 | -0.029 | -0.072 | -0.284 |
| | (0.394) | (0.154) | (0.343) |
| | [0.941] | [0.640] | [0.411] |
| Death in the ward on day T-3 | -0.119 | 0.239 | 0.195 |
| | (0.217) | (0.150) | (0.261) |
| | [0.585] | [0.118] | [0.458] |
| | **Outcome:** | | |
| | Band 6 RN fill-rate | Band 7-8 RN fill-rate | Agency RN fill-rate |
| | (4) | (5) | (6) |
| Death in the ward on day T-1 | -0.232 | -0.095 | 0.225 |
| | (0.262) | (0.095) | (0.262) |
| | [0.381] | [0.322] | [0.395] |
| Death in the ward on day T-2 | 0.268 | -0.057 | -0.018 |
| | (0.269) | (0.148) | (0.230) |
| | [0.323] | [0.703] | [0.938] |
| Death in the ward on day T-3 | 0.016 | 0.028 | 0.186 |
| | (0.190) | (0.135) | (0.246) |
| | [0.934] | [0.834] | [0.452] |
| Patient and time controls | ✓ | ✓ | ✓ |
| Ward fixed effects | ✓ | ✓ | ✓ |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are from a fixed-effects regression with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses. The unit of analysis is the hospital ward-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A6: Impact of deaths on previous days on levels of nurse staffing

| | **Outcome:** | |
|---|---|---|
| | NA new sickness absence | RN new sickness absence |
| | (1) | (2) |
| Death in the ward on day T-1 | 0.880<br>(0.082)<br>[0.171] | 0.959<br>(0.070)<br>[0.570] |
| Death in the ward on day T-2 | 1.364***<br>(0.140)<br>[0.002] | 1.025<br>(0.076)<br>[0.734] |
| Death in the ward on day T-3 | 0.925<br>(0.152)<br>[0.635] | 1.057<br>(0.122)<br>[0.630] |
| Patient and time controls | ✓ | ✓ |
| Ward fixed effects | ✓ | ✓ |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital ward-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A7: Serial correlation in patient deaths

| | **Outcome: patient death on day T** | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Patient death on day T-1 | 1.003<br>(0.157)<br>[0.986] | 1.004<br>(0.150)<br>[0.980] | 1.000<br>(0.149)<br>[0.999] |
| Time controls | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ |
| Patient volume and characteristics | × | ✓ | ✓ |
| Staffing levels and composition | × | × | ✓ |
| N | 18,870 | 18,870 | 18,870 |
| Clusters | 52 | 52 | 52 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. Time controls include a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday. Patient controls include the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; and the number of patients treated in the unit that day. Staffing controls include the fraction of planned NA hours actually worked (the NA fill-rate); the fraction of planned RN hours worked by band 5s; the fraction of planned RN hours worked by band 6s; and the fraction of planned RN hours worked by band 7 and 8s.

Table A8: Robustness of main results to inclusion of ward-month fixed effects

| | **Outcome: patient death** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.002<br>(0.004)<br>[0.533] | 1.002<br>(0.004)<br>[0.532] | 1.002<br>(0.004)<br>[0.525] | 1.002<br>(0.004)<br>[0.525] |
| Share (ppt) of planned Registered Nurse (RN) hours filled | 0.988*<br>(0.007)<br>[0.065] | | | |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | | | |
| Band 5 nurses | | 0.990<br>(0.007)<br>[0.117] | | |
| Band 6 nurses | | 0.984**<br>(0.008)<br>[0.040] | | 0.984**<br>(0.007)<br>[0.038] |
| Band 7 and 8 nurses | | 0.975**<br>(0.011)<br>[0.022] | | 0.975**<br>(0.011)<br>[0.021] |
| Local (regular) nurses | | | 0.987*<br>(0.007)<br>[0.053] | |
| Bank nurses | | | 0.987*<br>(0.007)<br>[0.063] | |
| Agency nurses | | | 0.993<br>(0.007)<br>[0.304] | |
| Band 5 Local and Bank nurses | | | | 0.989<br>(0.007)<br>[0.100] |
| Band 5 Agency nurses | | | | 0.992<br>(0.007)<br>[0.278] |
| Patient volume and characteristics | ✓ | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward – month fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 13,362 | 13,362 | 13,362 | 13,362 |
| Clusters | 440 | 440 | 440 | 440 |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward – month fixed effects. Robust standard errors, clustered at the hospital unit – month level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; and a dummy indicating whether the day was a bank holiday.

Table A9: Robustness of main results to inclusion of ward-quarter fixed effects

| | **Outcome: patient death** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.003<br>(0.004)<br>[0.454] | 1.003<br>(0.004)<br>[0.450] | 1.003<br>(0.004)<br>[0.445] | 1.003<br>(0.004)<br>[0.444] |
| Share (ppt) of planned Registered Nurse (RN) hours filled | 0.987**<br>(0.006)<br>[0.041] | | | |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | | | |
| Band 5 nurses | | 0.988*<br>(0.007)<br>[0.080] | | |
| Band 6 nurses | | 0.984**<br>(0.007)<br>[0.029] | | 0.984**<br>(0.007)<br>[0.029] |
| Band 7 and 8 nurses | | 0.975**<br>(0.011)<br>[0.019] | | 0.975**<br>(0.011)<br>[0.019] |
| Local (regular) nurses | | | 0.986**<br>(0.006)<br>[0.034] | |
| Bank nurses | | | 0.986**<br>(0.007)<br>[0.041] | |
| Agency nurses | | | 0.991<br>(0.007)<br>[0.210] | |
| Band 5 Local and Bank nurses | | | | 0.988*<br>(0.007)<br>[0.069] |
| Band 5 Agency nurses | | | | 0.991<br>(0.007)<br>[0.206] |
| Patient volume and characteristics | ✓ | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward – quarter fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 17,398 | 17,398 | 17,398 | 17,398 |
| Clusters | 191 | 191 | 191 | 191 |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward – quarter fixed effects. Robust standard errors, clustered at the hospital unit – month level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; and a dummy indicating whether the day was a bank holiday.

Table A10: Impact of team-specific experience on patient mortality

| | **Outcome: patient death** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Average number of shifts with other team members in past 90 days (NAs) | 1.004<br>(0.010)<br>[0.725] | | | |
| Average number of shifts with other team members in past 90 days (RNs) | 0.993<br>(0.012)<br>[0.552] | | | |
| Average number of shifts with other RNs in past 90 days (RNs) | | 0.998<br>(0.009)<br>[0.791] | | |
| Average share of team members worked with before (NAs) | | | 1.005<br>(0.004)<br>[0.155] | |
| Average share of team members worked with before (RNs) | | | 0.994<br>(0.005)<br>[0.236] | |
| Average share of team members worked with 10+ times before (NAs) | | | | 1.001<br>(0.002)<br>[0.547] |
| Average share of team members worked with 10+ times before (RNs) | | | | 0.996<br>(0.003)<br>[0.146] |
| Staffing quantity and seniority mix | ✓ | ✓ | ✓ | ✓ |
| Patient volume and characteristics | ✓ | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 13,598 | 13,598 | 13,598 | 13,598 |
| Clusters | 52 | 52 | 52 | 52 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital ward-day. The first two specifications have fewer observations owing to the exclusion of teams without both NAs and RNs in the team. Staffing quantity and seniority mix controls include: the NA fill-rate, the band 5 RN fill-rate, the band 6 RN fill-rate, and the band 7–8 RN fill-rate. All regressions additionally control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A11: Impact of ward-specific experience on patient mortality

| | **Outcome: patient death** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Average shifts in the ward in past 90 days (NAs) | 1.001<br>(0.005)<br>[0.865] | 1.000<br>(0.002)<br>[0.874] | | |
| Average shifts in the ward in past 90 days (RNs) | 0.998<br>(0.008)<br>[0.823] | | | |
| Average % of hours in the ward in past 90 days (NAs) | | 1.000<br>(0.002)<br>[0.874] | | |
| Average % of hours in the ward in past 90 days (RNs) | | 0.999<br>(0.004)<br>[0.844] | | |
| NA new joiner in the team (binary) † | | | 0.879<br>(0.072)<br>[0.116] | 0.880<br>(0.072)<br>[0.118] |
| RN new joiner in the team (binary) † | | | 1.131<br>(0.121)<br>[0.251] | |
| Band 5 RN new joiner in the team (binary) † | | | | 1.195*<br>(0.115)<br>[0.065] |
| Band 6 RN new joiner in the team (binary) † | | | | 1.021<br>(0.111)<br>[0.849] |
| Band 7-8 RN new joiner in the team (binary) † | | | | 0.793<br>(0.222)<br>[0.408] |
| Staffing quantity and seniority mix | ✓ | ✓ | ✓ | ✓ |
| Patient volume and characteristics | ✓ | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 13,598 | 13,598 | 14,290 | 14,290 |
| Clusters | 52 | 52 | 52 | 52 |

*** p < 0.01, ** p < 0.05, * p < 0.1

† New joiners are defined as those who are within 30 days of starting on the unit for the first time, having not done so in (at least) the past 90 days.

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital ward-day. The first two specifications have fewer observations owing to the exclusion of teams without both NAs and RNs in the team. Staffing quantity and seniority mix controls include: the NA fill-rate, the band 5 RN fill-rate, the band 6 RN fill-rate, and the band 7–8 RN fill-rate. All regressions additionally control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A12: Impact of firm-specific, plant-specific and team-specific experience on patient deaths

| | **Outcome: patient death** | |
|---|---|---|
| | (1) | (2) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.001 | 1.001 |
| | (0.004) | (0.004) |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | |
| Nurses of any band | 0.983** | |
| | (0.007) | |
| Band 5 nurses | | 0.984** |
| | | (0.008) |
| Band 6 nurses | | 0.981** |
| | | (0.008) |
| Band 7 and 8 nurses | | 0.982 |
| | | (0.011) |
| **Firm-specific experience:** | | |
| Average experience in the Trust for NAs (years) | 1.013 | 1.013 |
| | (0.018) | (0.018) |
| Average experience in the Trust for RNs (years) | 0.928*** | 0.932*** |
| | (0.021) | (0.023) |
| **Plant-specific experience:** | | |
| Average shifts in the ward in past 90 days (NAs) | 0.998 | 0.998 |
| | (0.016) | (0.016) |
| Average shifts in the ward in past 90 days (RNs) | 1.015 | 1.017 |
| | (0.021) | (0.021) |
| **Team-specific experience:** | | |
| Average number of shifts with other team members in past 90 days (NAs) | 1.011 | 1.011 |
| | (0.034) | (0.034) |
| Average number of shifts with other team members in past 90 days (RNs) | 0.976 | 0.975 |
| | (0.034) | (0.033) |
| Patient volume and characteristics | ✓ | ✓ |
| Time controls | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ |
| N | 13,598 | 13,598 |
| Clusters | 52 | 52 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses. The unit of analysis is the hospital unit-day. The unit of analysis is the hospital ward-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A13: Patient-day level analysis: impact of staffing fill-rate and firm-specific experiene on mortality

| | **Outcome: death (binary)** | |
|---|---|---|
| | (1) | (2) |
| Nursing Assistant (NA) fill-rate | 1.004<br>(0.004)<br>[0.262] | 1.004<br>(0.004)<br>[0.268] |
| Registered Nurse (RN) fill-rate: band 5 | 0.995<br>(0.007)<br>[0.453] | 0.994<br>(0.007)<br>[0.372] |
| Registered Nurse (RN) fill-rate: band 6 | 0.989<br>(0.007)<br>[0.138] | 0.990<br>(0.007)<br>[0.173] |
| Registered Nurse (RN) fill-rate: band 7 & 8 | 0.981*<br>(0.011)<br>[0.088] | 0.984<br>(0.011)<br>[0.149] |
| Average experience in the Trust for NAs (years) | | 1.023<br>(0.017)<br>[0.162] |
| Average experience in the Trust for RNs (years) | | 0.954*<br>(0.026)<br>[0.082] |
| Patient characteristics | ✓ | ✓ |
| Time controls | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ |
| N | 270,553 | 270,356 |
| Clusters (teams) | 18,564 | 18,544 |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are odds ratios from a logit regression. Robust standard errors clustered at the hospital unit-day level (i.e. team level) are shown in parentheses, and p-values are shown in square brackets. The unit of analysis is the patient-day. Staffing variables are calculated as the average across the wards the patient was treated in that day, weighted by the number of hours spent in each. All regressions control for patient age; patient age$^2$; patient age$^3$; sex; a dummy for being non-white; the patient's Elixhauser Cormorbidity Index; pairwise interaction terms between age, sex, non-white dummy and Elixhauser Cormorbidity Index; a set of dummies for 17 separate diseases, derived from diagnosis codes; the day within the patient's hospital spell; a dummy for each day of the week; a dummy for month of the year; a dummy indicating whether the day was a bank holiday; and a set of dummies for each hospital ward the patient was treated in that day.

Table A14: Patient-day level analysis: heterogeneity by patient severity

| | **Outcome: patient death** | |
|---|---|---|
| | Low-severity patients† | High-severity patients† |
| | (1) | (2) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 0.998<br>(0.005)<br>[0.648] | 1.008<br>(0.005)<br>[0.115] |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | |
| Band 5 nurses | 0.986<br>(0.012)<br>[0.233] | 1.000<br>(0.009)<br>[0.999] |
| Band 6 nurses | 0.984<br>(0.012)<br>[0.167] | 0.993<br>(0.009)<br>[0.422] |
| Band 7 and 8 nurses | 0.961**<br>(0.017)<br>[0.024] | 0.990<br>(0.014)<br>[0.497] |
| Patient characteristics | ✓ | ✓ |
| Time controls | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ |
| N | 138,649 | 120,771 |
| Clusters | 16,658 | 17,853 |

† The median Elixhauser Cormorbidity Index among patients who died in our sample was 5. A low-severity patient is defined as one with an Elixhauser score of less than 5; a high-severity patient as one with an Elixhauser score of 5 or more.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the patient-day. Staffing variables are calculated as the average across the wards the patient was treated in that day, weighted by the number of hours spent in each. All regressions control for patient age; patient age$^2$; patient age$^3$; sex; a dummy for being non-white; the patient's Elixhauser Cormorbidity Index; pairwise interaction terms between age, sex, non-white dummy and Elixhauser Cormorbidity Index; a set of dummies for 17 separate diseases, derived from diagnosis codes; the day within the patient's hospital spell; a dummy for each day of the week; a dummy for month of the year; a dummy indicating whether the day was a bank holiday; and a set of dummies for each hospital ward the patient was treated in that day.

Table A15: Patient-day level analysis: heterogeneity by diagnosis

| | **Outcome: patient death** | |
|---|---|---|
| | Without sepsis† | With sepsis† |
| | (1) | (2) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.003 | 1.008 |
| | (0.004) | (0.007) |
| | [0.503] | [0.284] |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | |
| Band 5 nurses | 1.005 | 0.972** |
| | (0.009) | (0.012) |
| | [0.583] | [0.019] |
| Band 6 nurses | 1.000 | 0.960*** |
| | (0.009) | (0.014) |
| | [0.975] | [0.004] |
| Band 7 and 8 nurses | 0.986 | 0.966* |
| | (0.013) | (0.019) |
| | [0.302] | [0.085] |
| Patient characteristics | ✓ | ✓ |
| Time controls | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ |
| N | 232,432 | 24,618 |
| Clusters | 17,069 | 11,144 |

† A patient is coded as having sepsis if they have a primary or secondary ICD-10 diagnosis code of A400, A403, A409, A410, A411, A414, A415, A418, or A419.

*** p < 0.01, ** p < 0.05, * p < 0.1. Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the patient-day. Staffing variables are calculated as the average across the wards the patient was treated in that day, weighted by the number of hours spent in each. All regressions control for patient age; patient age$^2$; patient age$^3$; sex; a dummy for being non-white; the patient's Elixhauser Cormorbidity Index; pairwise interaction terms between age, sex, non-white dummy and Elixhauser Cormorbidity Index; a set of dummies for 17 separate diseases, derived from diagnosis codes; the day within the patient's hospital spell; a dummy for each day of the week; a dummy for month of the year; a dummy indicating whether the day was a bank holiday; and a set of dummies for each hospital ward the patient was treated in that day.

Table A16: Nurse staffing, team spillovers and subsequent patient deaths

| | **Outcome: patient treated by the team dying under the care of another team within:** | | |
|---|---|---|---|
| | 1 day | 2 days | 3 days |
| | (1) | (2) | (3) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.001 | 0.998 | 0.998 |
| | (0.003) | (0.002) | (0.002) |
| | [0.638] | [0.496] | [0.282] |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | | |
| Band 5 nurses | 0.994 | 0.998 | 1.000 |
| | (0.005) | (0.005) | (0.005) |
| | [0.208] | [0.769] | [0.966] |
| Band 6 nurses | 0.998 | 0.999 | 1.001 |
| | (0.006) | (0.006) | (0.006) |
| | [0.654] | [0.872] | [0.870] |
| Band 7 and 8 nurses | 0.989 | 0.992 | 0.999 |
| | (0.010) | (0.008) | (0.008) |
| | [0.234] | [0.351] | [0.876] |
| Patient volume and characteristics | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ |
| N | 18,818 | 18,766 | 18,714 |
| Clusters | 52 | 52 | 52 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The outcome variable is defined as being equal to 1 if a patient under the care of a team subsequently dies under the care of another team within the specified number of days. The unit of analysis is the hospital unit-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A17: Impact of lagged staffing on patient mortality

| | **Outcome: patient death** | |
|---|---|---|
| | (1) | (2) |
| Nursing Assistant (NA) fill-rate on day T | 1.003<br>(0.004)<br>[0.359] | 1.003<br>(0.003)<br>[0.347] |
| Registered Nurse (RN) fill-rate on day T | 0.989*<br>(0.007)<br>[0.087] | |
| Registered Nurse (RN) fill-rate on day T-1 | 0.992<br>(0.006)<br>[0.177] | |
| Registered Nurse (RN) fill-rate on day T-2 | 0.997<br>(0.008)<br>[0.748] | |
| Registered Nurse (RN) fill-rate on day T-3 | 1.008<br>(0.008)<br>[0.347] | |
| Band 5 RN fill-rate on day T | | 0.989*<br>(0.006)<br>[0.088] |
| Band 6 RN fill-rate on day T | | 0.984**<br>(0.007)<br>[0.020] |
| Band 7 & 8 RN fill-rate on day T | | 0.977*<br>(0.012)<br>[0.057] |
| Band 7 & 8 RN fill-rate on day T-1 | | 0.994<br>(0.011)<br>[0.588] |
| Band 7 & 8 RN fill-rate on day T-2 | | 0.992<br>(0.008)<br>[0.301] |
| Band 7 & 8 RN fill-rate on day T-3 | | 1.014<br>(0.010)<br>[0.180] |
| Patient volume and characteristics | ✓ | ✓ |
| Time controls | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ |
| N | 18,766 | 18,766 |
| Clusters | 52 | 52 |

*** p < 0.01, ** p < 0.05, * p < 0.1

Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital ward-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Table A18: Patient-day level analysis: impact of lagged staffing

| | **Outcome:** | | | |
|---|---|---|---|---|
| | Death (binary) | | Deviation from predicted mortality | |
| | (1) | (2) | (3) | (4) |
| NA fill-rate on day T | 1.002<br>(0.004)<br>[0.661] | 1.003<br>(0.005)<br>[0.564] | 0.001<br>(0.002)<br>[0.736] | 0.001<br>(0.002)<br>[0.611] |
| RN fill-rate: band 5 on day T | 0.991<br>(0.008)<br>[0.256] | 0.990<br>(0.008)<br>[0.197] | -0.004<br>(0.003)<br>[0.269] | -0.004<br>(0.003)<br>[0.211] |
| RN fill-rate: band 6 on day T | 0.986*<br>(0.008)<br>[0.088] | 0.987<br>(0.008)<br>[0.115] | -0.006*<br>(0.003)<br>[0.070] | -0.006*<br>(0.003)<br>[0.095] |
| RN fill-rate: band 7 & 8 on day T | 0.973**<br>(0.013)<br>[0.044] | 0.974**<br>(0.013)<br>[0.046] | -0.012**<br>(0.006)<br>[0.041] | -0.012**<br>(0.006)<br>[0.028] |
| RN fill-rate: band 7 & 8 on day T-1 | 0.994<br>(0.012)<br>[0.635] | | -0.003<br>(0.006)<br>[0.532] | |
| RN fill-rate: band 7 & 8 on day T-2 | 0.992<br>(0.012)<br>[0.490] | | -0.004<br>(0.005)<br>[0.499] | |
| RN fill-rate: band 7 & 8 on day T-3 | 1.011<br>(0.011)<br>[0.309] | | 0.005<br>(0.005)<br>[0.336] | |
| Average Trust experience for NAs on day T | | 1.017<br>(0.019)<br>[0.373] | | -0.003<br>(0.006)<br>[0.532] |
| Average Trust experience for RNs on day T | | 0.949*<br>(0.029)<br>[0.082] | | -0.022**<br>(0.011)<br>[0.045] |
| Average Trust experience for RNs on day T-1 | | 0.991<br>(0.030)<br>[0.772] | | -0.005<br>(0.011)<br>[0.644] |
| Average Trust experience for RNs on day T-2 | | 0.994<br>(0.028)<br>[0.826] | | -0.003<br>(0.010)<br>[0.744] |
| Average Trust experience for RNs on day T-3 | | 1.029<br>(0.027)<br>[0.274] | | 0.010<br>(0.010)<br>[0.314] |
| Patient characteristics | ✓ | ✓ | ✓ | ✓ |
| Time controls | ✓ | ✓ | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ | ✓ | ✓ |
| N | 162,067 | 161,933 | 162,685 | 162,551 |
| Clusters (teams) | 17,472 | 17,452 | 17,743 | 17,723 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Coefficients shown in columns 1 and 2 are odds ratios from a logit regression. Coefficients shown in columns 3 and 4 are from an OLS regression. In both cases, robust standard errors clustered at the hospital unit-day level (i.e. team level) are shown in parentheses, and p-values are shown in square brackets. The unit of analysis is the patient-day. Staffing variables are calculated as the average across the wards the patient was treated in that day, weighted by the number of hours spent in each. All regressions control for patient age; patient age$^2$; patient age$^3$; sex; a dummy for being non-white; the patient's Elixhauser Cormorbidity Index; pairwise interaction terms between age, sex, non-white dummy and Elixhauser Cormorbidity Index; a set of dummies for 17 separate diseases, derived from diagnosis codes; the day within the patient's hospital spell; a dummy for each day of the week; a dummy for month of the year; a dummy indicating whether the day was a bank holiday; and a set of dummies for each hospital ward the patient was treated in that day.

Table A19: Impact of nurse staffing and firm-specific experience on patient mortality: heterogeneity by patient severity

| | **Outcome:** | |
|---|---|---|
| | Death of low-severity patient† | Death of high-severity patient† |
| | (1) | (2) |
| Share (ppt) of planned Nursing Assistant (NA) hours filled | 1.002<br>(0.006)<br>[0.722] | 1.007<br>(0.005)<br>[0.160] |
| Share (ppt) of planned Registered Nurse (RN) hours filled by: | | |
| Band 5 nurses | 0.984*<br>(0.008)<br>[0.060] | 0.996<br>(0.011)<br>[0.731] |
| Band 6 nurses | 0.984*<br>(0.009)<br>[0.071] | 0.990<br>(0.011)<br>[0.351] |
| Band 7 and 8 nurses | 0.977<br>(0.016)<br>[0.175] | 0.987<br>(0.016)<br>[0.443] |
| Average experience in the Trust for NAs (years) | 1.008<br>(0.021)<br>[0.683] | 1.018<br>(0.017)<br>[0.289] |
| Average experience in the Trust for RNs (years) | 0.909***<br>(0.028)<br>[0.002] | 0.952<br>(0.033)<br>[0.151] |
| Patient and time controls | ✓ | ✓ |
| Hospital ward fixed effects | ✓ | ✓ |
| N | 17,638 | 17,642 |
| Clusters | 50 | 51 |

† The median Elixhauser Cormorbidity Index among patients who died in our sample was 5. A low-severity patient is defined as one with an Elixhauser score of less than 5; a high-severity patient as one with an Elixhauser score of 5 or more.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficients shown are odds ratios from a fixed-effect logit regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital ward-day. All regressions control for the mean patient age; the mean patient age$^2$; the female share of patients; the non-white share of patients; the mean Elixhauser Comorbidity Index; pairwise interaction terms between age, female share, non-white share and Elixhauser Cormorbidity Index; the mean hospital length of stay of patients treated in the unit that day; the number of patients treated in the unit that day; a dummy for each day of the week; a dummy for month of the year; and a dummy indicating whether the day was a bank holiday.

Figure A7: Distribution of time of recorded patient death

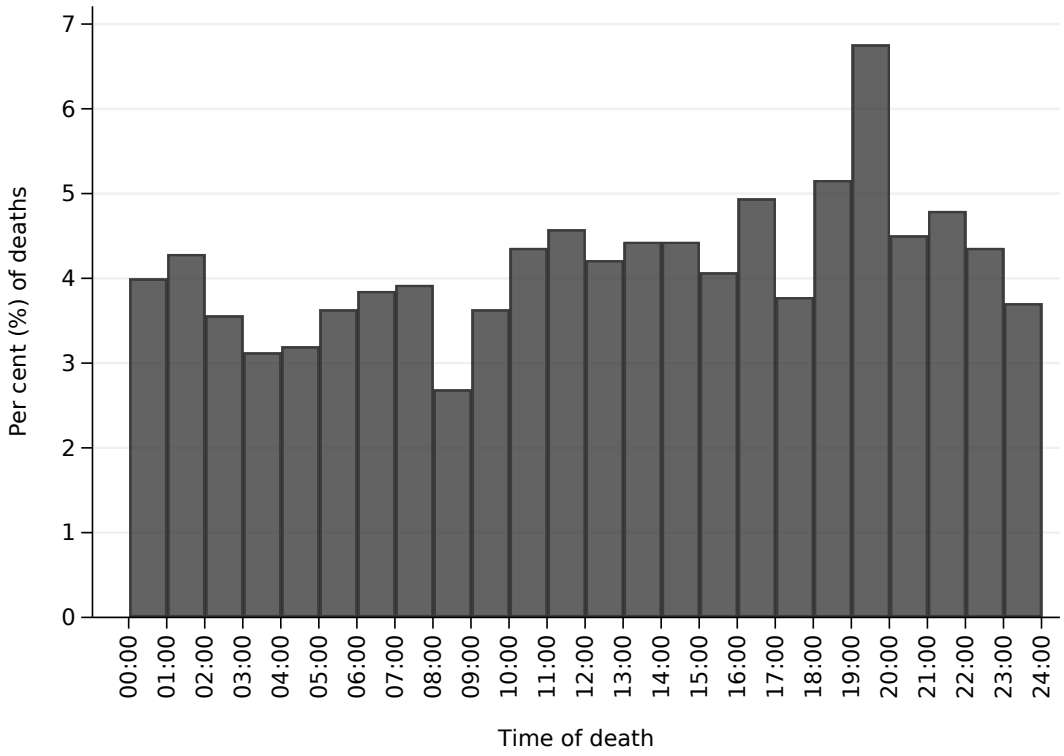
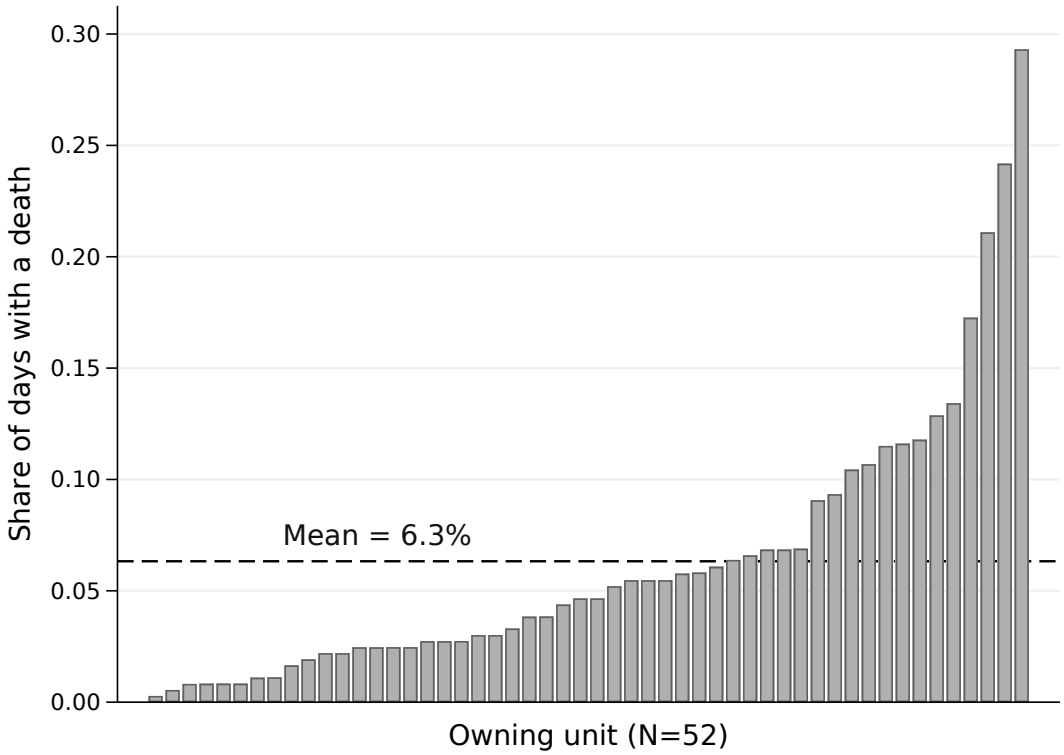


Figure A8: Frequency of patient deaths across wards



## A.5 Back of the envelope calculation of cost of staffing shortages

Our estimates suggest that adding a 12-hour shift from a degree-qualified nurse to the average team would reduce the odds of experiencing a patient death by around 10% (as per Section 6.1). This implies a reduction in the probability of a death of around 0.6 percentage points.[29] That suggests that it would require roughly 170 additional shifts in order to save one life. Based on the hourly rates for agency nurses – the marginal unit of nursing available in the short-term – used by the hospital in 2022/23, each additional shift would cost around £440.[30] A back-of-the-envelope calculation therefore suggests that each life saved would cost approximately £75,000. The cost per quality-adjusted life year (QALY) threshold used by the National Institute for Health and Care Excellence (NICE) for England and Wales is between £20,000 and £30,000. We are not able to estimate the number of QALYs saved by each additional nurse and so cannot conduct a direct cost-utility analysis of the effects of higher nurse staffing. Our estimates suggest, though, that each patient who died in our sample – among whom the median age was 77 – who might have otherwise survived with higher levels of nursing staff, would need to have otherwise lived for a further three or four years for this to be a cost-effective intervention.

[29] A patient death occurs in 6.3% of teams in our sample, equivalent to baseline odds of $0.063/(1 - 0.063) \approx 0.067$. A 10% reduction in those odds would take it to $\approx 0.060$, equivalent to a probability of $0.06/1.06 \approx 5.7\%$, around 0.6 percentage points lower than the sample mean probability of 6.3%.

[30] This is the mean hourly price cap for agency nurses, using the highest spine point within each AfC band and weighting by the proportion of hours worked by staff of different grades, on weekends, and at night. This suggests a mean hourly rate of around £36.70.