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

Elaine Kelly, Carol Propper and Ben Zaranko*

November 2022

Abstract

Teams are an important form of work organisation. We examine the impact of team composition on collective output. Our focus is the in-hospital nursing team. We exploit daily rotation of nursing staff across teams to examine the effect on patient mortality outcomes of the quantity of staff and different levels and types of human capital. We find evidence of significant returns to the quantity of registered nurses with university-level qualifications, and to firm-specific human capital. The mortality impacts are concentrated in low-severity patients. We find no clear evidence of returns to location-specific experience or familiarity with other team members.

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 HealthcareLabour. We would like to thank Eric French, Imran Rasul and seminar participants at the Institute for Fiscal Studies, University of York, Essen Health Conference 2022, European Association of Labour Economists 2022 Conference, European Health Economics Association 2022 Conference, Royal Economic Society conference 2022, Monash University and University of Sydney 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 an essential role in the organisation of work. Teams are formed because they make possible gains from specialisation, knowledge sharing, and complementarities between the skills of different workers. They typically contain individuals with different skills and knowledge and at different levels of the firm hierarchy (Lazear and Shaw, 2007). However, while the composition of the team matters, team output is typically a function of all inputs and the output of individual inputs cannot be directly measured. This raises important identification questions about the contribution of different inputs. 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).

The aim of this paper is to study the impact of team composition in a setting where the output is collective and of considerable societal importance. Our context is the hospital nursing team and the measure of output we examine is (in)patient mortality. Production in healthcare is widely characterised by team production and nurses make up a large proportion of the medical labour force, as well as being a sizeable group in the overall labour force. Nursing teams play a critical role in the delivery of healthcare, 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 needed. 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 productivity and the quality of patient care.

Our focus is the impact of individuals with different levels and types of human capital, as determined by qualifications, seniority and the level of familiarity with the firm, location of production and the team. We identify this by exploiting deviations from staffing levels that were planned several months in advance and from the rotation of staff into teams at the daily level, which arises because care needs to be provided 24/7 but nurses only work 12-hour shifts and on average three shifts per week.[1] This allows us to study the effect of an absence of a more senior or more qualified staff member, compared to 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. In our specific context, the English National Health Service, pay is determined exogenously to the organisation we study, is unrelated to output, and individuals of the

[1]The variation in size is different to 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).

same grade and contract type are paid the same amount, regardless of their contribution to the output of their team. This allows us to abstract away from any role for individual financial incentives.

Our data utilises around 19,000 teams containing just under 4,500 unique staff, responsible for 44 thousand patients and providing 3.5 million hours of patient care. In our analysis of the impact of team composition on team output, we control for patient acuity and firm (hospital), plant (ward) and day fixed effects. Identification comes from within-ward variation in staffing levels and composition over time. We explicitly address concerns that team staffing, despite being set 3 months in advance, is endogenous to our outcome variable, mortality, and to patient severity and volumes. We undertake a large set of tests to examine these key assumptions and find no evidence of such sources of endogeneity. Specifically, 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 the patient mix. We show that deaths on previous days have no clear impact on either subsequent staffing levels or subsequent staff sickness absences. 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. 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 morality. Finally, allowing for hospital ward-specific time trends does not change our results.

Our results show that both the size and composition of the nursing team matters for patient mortality.[2] For the average team an additional registered nurse (one with university degree-equivalent level training) reduces the odds of a patient death by approximately 10%. In contrast, the addition of an extra nursing assistant, who does not have the formal training of a registered nurse, has no impact on patient mortality. Within the category of registered nurses, more senior nurses have a larger impact. We also find evidence of substantial returns to firm-specific experience (in our setting, experience working within the hospital group). Conditional on the quantity and skill-mix of the team, increasing the average firm-specific experience among registered nurses by one year reduces the odds of a patient death by 7.2% , which is equivalent to roughly two-thirds of the impact of adding an extra registered nurse to the team. However, there is again no evidence of such an effect for nursing assistants. We also find little

[2]Mortality is commonly used in the medical and nursing literature as a proxy for quality of nursing care (e.g. Needleman et al. (2011); Griffiths et al. (2019)), with potential mechanisms including 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).

returns to location-specific or team-specific experience. Finally, we show that our results are driven by mortality impacts on patients of relatively low, rather than high, clinical severity. Whilst we do not observe nurse-patient interactions, it is plausible that in teams that are short-staffed or missing senior members effort is focused on the more severely ill.

Our work contributes 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 for different types of injured 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 variations 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).[3] It also enables us to examine possible dynamic relationships between absences and patient outcomes as well as contemporaneous effects. The granularity of our data allows us to go beyond Bartel et al. in distinguishing the impact on team productivity of levels of seniority *within* qualification group and the impact of temporary (agency) workers compared to regular workers, and allows us to separately identify the effects of location and hospital (firm) experience, and the effects of familiarity of team members with each other.[4]

[3]Rotas are planned at the shift level. We aggregate two 12 hour shifts to the 24 hour level to allow for both the night and the day shifts as these have different tasks and different staffing levels and composition.

[4]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),

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 more high 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.[5]

A third related literature is concerned with worker human capital and productivity. These issues have received growing attention in the health care 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 describes the institutional context, outlines a simple model of team production in a nursing setting, and introduces the data and staffing measures used in our analysis. Section 3 presents descriptive statistics. Section 4 presents our empirical strategy and considers possible threats to identification. Section 5 sets out our results, looking first at the impact of overall staffing and general human capital, before turning to the returns to more specific forms of experience and heterogeneity across different types of patient. Section 6 concludes.

2 The team in a nursing setting, data, and variable definitions

2.1 Nursing teams in the English National Health Service

Nursing teams in public hospitals in the the UK contain two broad groups of workers. Registered nurses (RNs) are fully qualified nurses, registered in the UK with the Nursing and Midwifery Council, who have completed formal training and hold a university

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

[5] Our research also contributes to research on the value of agency workers: see for example Autor and Houseman (2010); Bryson (2013); Jahn and Pozzoli (2013).

diploma or degree-level qualification. Nursing assistants (NAs) are personnel without formal training or registration requirements, who are typically employed as health care assistants or other support workers, with responsibility only for basic patient tasks. For RNs, there are several levels of seniority and responsibility, organised and paid according to a national set of pay bands.[6] Promotion between these pay bands is based on experience and typically requires specific training or qualifications. Each ward is led by a senior nurse manager. NAs and RNs are typically contracted to work a set number of hours in a particular ward. Part time contracted hours are common, particularly for nurses with young children. Shifts are rostered 2-3 months in advance. Staff may also choose to work additional shifts either in their usual ward or elsewhere. Hospitals can also employ staff via an external agency if there are insufficient directly employed staff members available.[7]

2.2 Team production in nursing

We motivate and structure our empirical analysis by specifying how team composition may affect nursing production. A team is a set of nursing staff who work together in a particular location (hospital ward) to treat the patients in that ward. The output of the team is the medical care provided for those patients. Nursing care is delivered in teams because there are some tasks where a single nurse can safely care for multiple patients, while others require multiple team members to work together. Nurses and nursing assistants will typically perform both types of task over the course of a nursing shift. Tasks may be separated by type of worker (RN or NA), and there may be gains from specialisation for tasks that are repetitive, such as checking blood pressure. In addition, patients require round-the-clock care and so the sharing of information with the next team when nurses are handing over between shifts is of critical importance.

In some settings, the output of each team member can be observed and recorded separately (for example, as in the teams of fruit pickers in Bandiera et al. (2009, 2013)). But in the healthcare 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.

[6] Nursing assistants are typically employed on a single pay band, with multiple spine points. For RNs, there are multiple pay bands, each corresponding to a different level of responsibility. Each pay band has multiple spine points to account for experience.

[7] Nursing teams work alongside other healthcare professionals, including doctors. We do not observe these staff in our data and so cannot consider their interactions with nursing teams. The rosters of these other staff groups are, however, determined completely separately and just as far in advance.

Let T_{kt} be the team of n nurses $T = \{2, ..., n\}$ working in a location (a hospital ward) (k) on a particular day (t). Team size is not fixed and will vary across both wards and time. Let T^*_{kt}be the optimal size of the team.[8] The productivity of an individual i will depend on their innate ability (which we do not observe) and four types of human capital (which we do observe). Each of these is accumulated through different types of experience: general, firm-specific, plant-specific and team-specific capital.

- 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).

General human capital - such as overall education and experience - should raise an individual's productivity (Becker, 1964; Mincer, 1974). Human capital theory typically distinguishes between this and firm or task specific human capital, the last two not being portable across jobs (Parsons, 1972). The idea is that with experience at a particular firm, 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 (firms) often employ their own systems, processes and protocols. The longer an individual has worked at a particular hospital, the more familiar they will be with those systems, processes and protocols.[9] Similarly, different wards within a hospital (analogous to different plants within a firm) may employ different systems and/or require nurses to perform different tasks, in which case some portion of accumulated human capital would be ward- or plant-specific. Finally, if nurses have worked together before they

[8]In some nursing contexts, nursing/patient ratios are mandated. In our setting the optimal or 'target' team size is determined by hospital managers at the ward day of the week level based on factors such the ward's size, medical specialty (for example, intensive care units require more staff per patient), physical layout, and so on.

[9]Huckman and Pisano (2006) provide evidence of returns to hospital-specific experience for surgeons carrying out cardiac surgery.

will be familiar with how other members of the team work and therefore may be more effective.[10]

Given this, we posit that the size of the team, relative to its optimum, and higher team levels of human capital – whether general, firm-specific, plant-specific or team-specific – will be associated with higher output. This relationship is the empirical question that we address here.[11]

2.3 Data sources and sample construction

2.3.1 Data sources

We exploit a novel data linkage between electronic staff roster data and electronic patient records for a single NHS hospital group (known as an NHS Trust) in England. 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 the period between 1 January and 31 December 2017. Here we detail the sources for our key variables from these data. In Section 2.4 we present the precise definitions of the key variables used in estimation.

2.3.2 Teams

The NHS hospital used in our study includes three large hospital sites comprising of 89 wards. The three sites are physically separate but within the same area of a major city. To reduce heterogeneity we focus on inpatient wards with responsibility for providing continuous, round-the-clock treatment for treating adult patients.[12] For each of the remaining 52 hospital wards, we construct various measures of staffing, patient characteristics and patient outcomes at the daily level. The unit of analysis is the hos-

[10]An individual may also have a direct effect on the productivity of another worker. We do not estimate such peer effects explicitly in our analysis, though they may be picked up by the team familiarity component of human capital.

[11]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.

[12]We exclude: 18 wards which do not record a patient death at any point in the year (which includes administrative wards); 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.

pital ward-day. The final sample contains 18,922 observations.[13] Each of these can be thought of as a distinct team: 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).[14] 4,484 unique staff members worked across these teams, working 3,533,583 hours and 294,044 shifts over the course of the year. These teams were responsible for treating 44,485 unique patients with 66,662 separate hospital spells.

2.3.3 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 8 weeks notice of their shift patterns. The roster is generated so as to ensure that each team contains the appropriate number and mix of staff, with the ward-specific appropriate ('target') level of staffing as determined by hospital managers.[15] These shifts are then allocated to individual staff members either by the system or manually by managers. The mean length of a nursing shift is 11.5 hours and more than 70% of shifts are 12 hours long.[16]

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 a attendance at a training course. If there are no absences from the planned rota, a shift will be worked by a 'local' staff member (one who is directly employed by the NHS Trust, working their contracted hours in the ward 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' staff member (one who is employed by the NHS Trust but doing extracontractual or overtime work via the NHS 'staff bank', in their usual ward or elsewhere) or an agency worker (who is employed via an external agency and is generally contracted at short notice). To control costs, despite being in a situation of high vacancies of nursing staff, the Trust seeks to limit the use of (expensive) agency workers. This means in the case

[13]Note that 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.

[14]Note that because 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.

[15]There are no mandated national nurse to patient staffing ratios in the UK.

[16]Our data does not contain any information on doctors. The rostering system for doctors is completely separate and is independent of nurse staffing decisions. Our empirical strategy (Section 4) controls for permanent differences in doctor quality across wards and for rotation of doctors to wards during the year. Student nurses are treated as supernumerary and are therefore excluded from our analysis.

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.

To identify skill levels and seniority, we use the fact that 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 staff member who worked each shift. Staff in AfC pay bands 2–4 are NAs and those in pay bands 5–8 are RNs. Individual pay bands also identify staff of different levels of seniority within each group. Within the category of RNs, Band 5 includes newly qualified and less experienced nurses; Band 6 includes more senior and experienced nurses, including deputy ward managers; and Band 7 and 8 includes the most senior nurses and ward managers with the most responsibility.

For those staff members employed by the Trust, the staff records include the month and year in which their employment commenced, which we use to construct a measure of experience within the NHS Trust. Agency workers are not employed by the Trust and are given a value of 0 on this measure.[17]

2.3.4 Outcome measure

We use an important and high-stakes outcome: inpatient mortality. Death at discharge is recorded in electronic patient records, which we link to the team responsible for treating them at the point of death. Patient mortality is a commonly used performance measure in the medical literature and in a nursing context (e.g. Needleman et al. (2011); Griffiths et al. (2019)) and has the advantage of being unambiguous and of being accurately and consistently recorded across wards. Deaths occur at all hours of the day (Appendix Figure A5), 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. The outcome is treated as binary, with an indicator coded as 1 if a team (ward-day) had a patient death occur, and coded as 0 otherwise.

2.3.5 Patient severity and volumes

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 diagnosis codes, which we use to calculate a standard measure of clinical severity (the Elixhauser Comorbidity Index score (Elixhauser et al., 1998)) for each

[17] This is conservative if an agency worker is frequently employed by the same Trust.

patient. We also calculate the length of hospital stay for each patient. Our analysis is at the team level, so we aggregate up to the team level, weighting by the amount of time that each patient spent under the care of the team. In our main specification we include the following patient characteristics: 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.

2.4 Team staffing measures

2.4.1 Quantity of staff

We measure the quantity of staffing in terms of deviations from the target staffing set by the hospital management for the ward.[18] 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.[19]

The ratio between actual and planned staffing – the 'fill-rate' – has a number of appealing features for our analysis. First, it is relative 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, rather than relative to a mandated national-level minimum. 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 staff, which allows us to study the impact of both size and composition of the team in a comparable way across wards.

2.4.2 Human capital definitions

General human capital: We measure the amount of general human capital, W_{it}, in the team by the pay band of each team member. We start by grouping together all

[18] The NHS runs very close to capacity at all times: bed occupancy is over 95% and there are waiting lists for elective care to ensure beds are not left empty.

[19] Staffing cannot exceed the target level in this setting, so the fill-rate has an upper bound of 100.

NAs (those in AfC bands 2–4), and all registered nurses (those in AfC bands 5–8). We then construct the fill rate (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 (a newly qualified nurse is pay band 5, while the most senior nurses and ward managers are in pay bands 7 or 8). We then 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 $FillRateRN_{it}$.[20] 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-specific human capital: Firm-specific human capital, X_{iht}, represents the experience of individual i in firm (hospital) h at time t. We define this as the amount of time an individual has been employed in the firm, which in this context is the NHS Trust (at any one of the 3 constituent hospital sites).[21]

Plant-specific (ward-specific) human capital: We define plant-specific, or ward-specific, human capital, Y_{ihkt}, in several ways. First, 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:

$$Y_{ihkt} = \sum_{\tau=t-90}^{t-1} \mathbb{1}(i \in T_{\tau k}) \quad \forall k \in K(h) \tag{2}$$

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.[22] We then similarly construct the fraction of individual

[20] 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).

[21] 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 Trust started. For the remaining 11% of individuals, for whom this information is missing, experience in the Trust 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 Trust).

[22] We measure ward-specific experience based on the past 90 days for two reasons. One, one might

i's recent shifts that took place in ward k, by dividing Y_{ihkt} by the number of shifts individual i has done anywhere in the hospital group in the previous 90 days.

Finally, we identify those who are not familiar with the ward. If an individual is within their first 30 days on the ward, we define these individuals as 'new joiners':

$$\begin{aligned} FirstShift_{ihkt} &= \mathbb{1}(Y_{ihkt} = 0 \wedge i \in T_{kt}) \\ NewJoiner_{ihkt} &= \mathbb{1}(\sum_{\tau=t-30}^{t} FirstShift_{ikh\tau} > 0 \wedge i \in T_{kt}) \end{aligned} \tag{3}$$

where $\mathbb{1}(Y_{ihkt} = 0 \wedge i \in T_{kt})$ is an indicator equal to one if an individual is working in team T_{kt}, having not worked in ward k in the previous 90 days. $\mathbb{1}(\sum_{\tau=t-30}^{t} FirstShift_{ikh\tau} > 0 \wedge i \in T_{kt})$ is therefore an indicator equal to one if an individual in team T_{kt} did their first shift (in at least 90 days) in this ward on date $\tau \in [t-30, t]$ – i.e. at some point in the last 30 days. A 'new joiner' is therefore someone who is within their first month on the ward.

Team-specific human capital: 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.[23] Let $K(h)$ be the set of wards in hospital h. Let $E(i, j, t)$ represent the (recent) shared work experience between workers i and j, as of date t, where:

$$E(i, j, t) = \sum_{k \in K(h)} \sum_{\tau=t-90}^{t-1} \mathbb{1}(i \in T_{\tau k} \wedge j \in T_{\tau k}) \tag{4}$$

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 can then define an individual's average shared work experience with fellow team members (team-specific human capital) as:

$$Z_{-ihkt} = \frac{1}{|T_{kt}| - 1} \cdot \sum_{j \in T_{kt}, j \neq i} E(i, j, t) \tag{5}$$

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 shared work 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.

[23] As above, we measure shared work experience based on the past 90 days to allow for the effects of experience to decay over time, and to strike a balance between defining a suitably backwards-looking measure and having to discard a substantial portion of our sample.

where T_{kt} is the set of team members (in ward k on date t), $|T_{kt}|$ is the size of the team, and $E(i,j,t)$ is defined as above. Z_{-ihkt} therefore captures the average number of shifts that individual i has worked with their team members in the most recent 90 days. We also use this measure to create an indicator when for when $Z_{-ihkt} \geq 10$, i.e when an individual in team T_{kt} has worked with their team members 10 or more times, on average, over the last 90 days.

3 Summary statistics

3.1 Staffing and team composition

Table 1 presents summary statistics for the quantity and composition of staffing for the 18,922 teams in our sample. The average team (ward-day) had 194.8 rostered hours. Within that, an average 53.4 hours were rostered for NAs, and 141.4 for RNs. This is equivalent to around 16 staff members working 12-hour shifts, with around 4 NAs and 12 RNs 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 over the full 24 hours.

In just over half (51.2%) of teams, the actual number of hours worked was less than the number of planned/rostered hours. Equivalently, just over half of teams had a fill-rate (as defined above) of less than 100. In the average team, 8 hours of planned staffing was 'unfilled', equivalent to two-thirds of an average shift. This is equivalent to saying that the average team had a fill-rate of 95.7: 95.7% of planned hours were actually worked. This was slightly lower for NAs (94.6%) than for RNs (96.4%).[24] The distribution of and variation in the RN and NA fill-rate across wards is summarised in Figure 1.

Table 1 also shows the share of planned hours worked by staff of different contract type and seniority. 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 (as defined in Section 2.3) and a further 7.2% by agency workers.

Nurses in NHS AfC pay band 5 are the largest staffing 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

[24]The NA and RN fill-rate are positively correlated in our data, but only very weakly (coefficient=0.066).

responsible for managing their ward – work around 5% of all rostered RN hours.

A little over a tenth (12.2%) of all planned hours are worked by 'new joiners': those who are within their first month on a ward, having not worked there for the previous 90 days or more. In the average team, the average staff member had 6.3 years of experience within the NHS Trust (5.6 years for NAs; 6.7 for RNs); the average team member had done 30.3 shifts in the ward in the past 90 days (27.8 for NAs; 31.5 for RNs); and the average team member had worked 14.5 shifts with their fellow team members in the previous 90 days.[25]

Importantly, there is considerable day-to-day variation in both the level and composition of staffing. This includes variation in the degree of firm (hospital group)-, ward- and team-specific human capital amongst the team. This within-ward variation is key to our identification strategy and is documented in Table 3. 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 hospitals and wards within the hospital group. The average Nursing Assistant worked 76.3 shifts across 6.8 different wards over the course of the year. The average band 5 RN worked 51.6 shifts across 4.5 unique wards .[26] In contrast, the most senior nurses in band 7 and 8 move around little, and did 96.7% of their shifts in their usual unit (versus 72.4% for band 5s).

3.2 Patients and patient outcomes

The average team in our sample was responsible for treating 20.6 patients over the course of the 24 hours (Table 2). The average patient was 64.1 years old, 46% were female, and 52% were recorded with an ethnicity other than White (including 'unknown'). The average Elixhauser Comorbidity Score (Elixhauser et al., 1998), was 3.69.[27] 6.3% of the teams in our sample (1,197 of 18,922) recorded a patient death, with a mean mortality rate of 3.58 per 1,000 patients. There is considerable variation in the frequency of deaths across wards (see Appendix Figure A6). There is also considerable within ward variation in deaths as shown in Table 3.

---

[25]Summary statistics for all staffing measures can be found in Table A1.

[26]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 staff was 89.1 (with a median of 92).

[27]We calculated the Elixhauser Cormorbidity Score for each patient using their recorded diagnoses, with coding from the International Classification of Diseases, tenth edition (ICD-10).

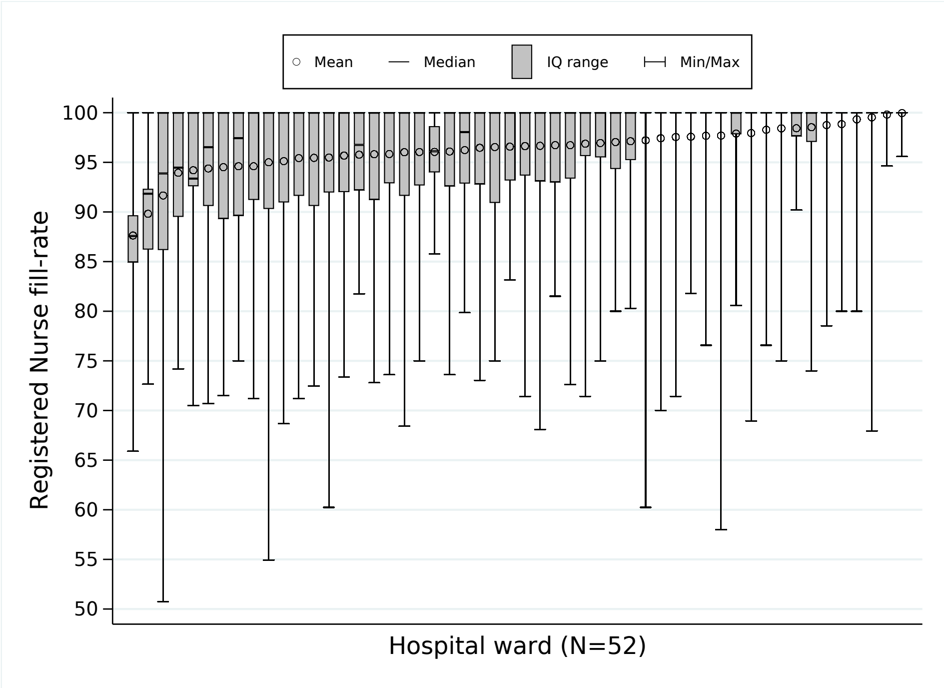
Table 1: Summary statistics: ward staffing levels and team composition

	All nursing staff		**Nursing Assistants**		**Registered Nurses**	
	Mean	SD	Mean	SD	Mean	SD
Planned (rostered) hours	194.8	89.8	53.4	30.7	141.4	91.4
Actual hours worked	186.7	88.8	50.2	29.1	136.6	90.6
Unfilled hours	8.0	10.1	3.2	6.1	4.8	7.6
Share of planned hours worked (%), *of which:*	95.7	5.4	94.8	10.9	96.4	5.8
Local/regular staff	67.9	15.7	63.9	26.4	70.5	16.9
Bank staff	20.6	12.7	28.8	23.8	16.3	12.7
Agency staff	7.2	8.2	2.0	8.4	9.5	11.0
Band 2 staff	23.9	14.6	78.0	25.5	–	–
Band 3 staff	3.7	5.4	15.7	23.6	–	–
Band 4 staff	0.2	1.0	0.9	3.8	–	–
Band 5 staff	47.1	11.5	–	–	68.4	15.8
Band 6 staff	17.0	12.5	–	–	22.8	13.3
Band 7 and 8 staff	3.8	4.5	–	–	5.2	5.8
Band 5 local staff	30.7	12.0	–	–	44.0	16.0
Band 5 bank staff	10.1	8.5	–	–	15.1	13.0
Band 5 agency staff	6.2	7.6	–	–	9.2	11.0
Band 6 local staff	15.6	10.6	–	–	21.2	11.9
Band 6 bank staff	1.1	3.3	–	–	1.2	3.7
Band 6 agency staff	0.3	2.0	–	–	0.3	2.0
Band 7 and 8 local staff	3.8	4.5	–	–	5.2	5.8
Band 7 and 8 bank staff	0.0	0.0	–	–	0.0	0.1
Band 7 and 8 agency staff	0.0	0.0	–	–	0.0	0.0
New joiners to the ward (%) †	12.2	9.4	16.4	19.7	10.7	10.6
Non-new joiners (%) †	83.5	10.7	78.4	21.9	85.6	11.9
Years of experience in the NHS Trust	6.3	2.0	5.6	3.2	6.7	2.5
Shifts in the ward in the past 90 days	30.3	5.2	27.8	9.7	31.5	5.9
Average shared work experience (shifts) with other team members (in any ward) in the past 90 days	14.5	5.0	13.5	6.3	14.9	5.1
Average shared work experience (shifts) with other RNs in the team (in any ward) in the past 90 days	–	–	–	–	15.2	5.7
Observations	18,922		18,922		18,922	
Hospitals	3		3		3	
Wards	52		52		52	

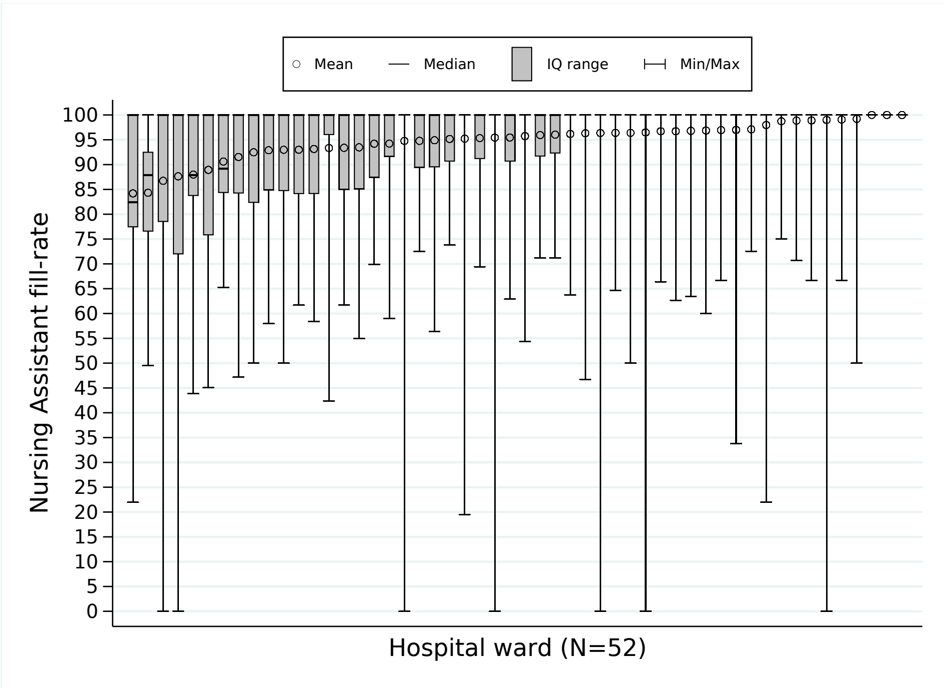
† New joiners are defined as those who are within 30 days of starting on the unit for the first time, having not done so in (at least) the past 90 days.
Note: Unit of analysis is the hospital unit-day. Nursing support staff are personnel without formal training or registration requirements, identified as those in NHS Agenda for Change pay bands 2–4), and typically employed as health care assistants or other support workers. 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, identified in our data as those in NHS Agenda for Change pay bands 5–8, with bands 5 and 8 representng the least and most senior nurses, respectively. Mean characteristics of team members are weighted by the number of hours worked.

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

(a) Registered Nurses



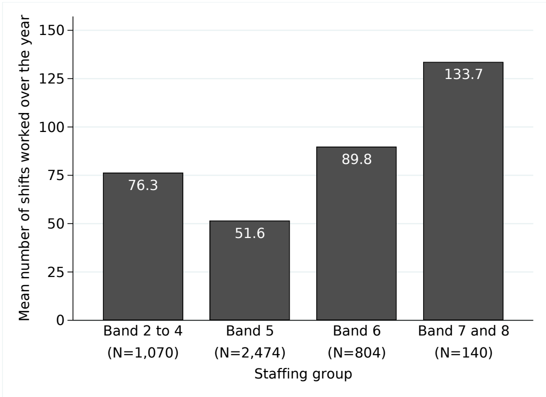
(b) Nursing Assistants



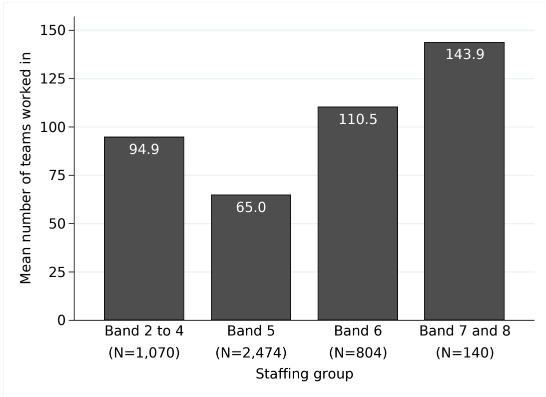
Note: Figures are calculated with respect to the 2017 calendar year (the period covered by our sample). The ordering of wards differs between panel a) and panel b).

Figure 2: Staff movements across teams, wards and hospitals

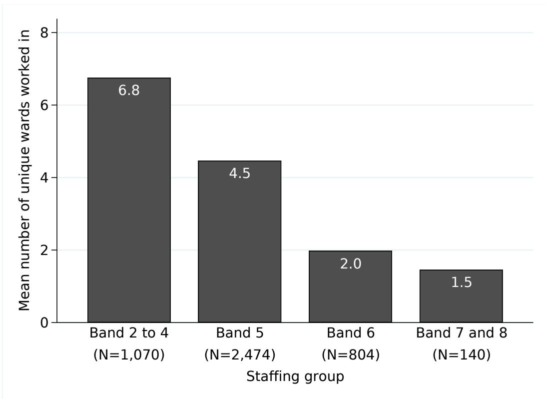
(a) Number of shifts (mean)



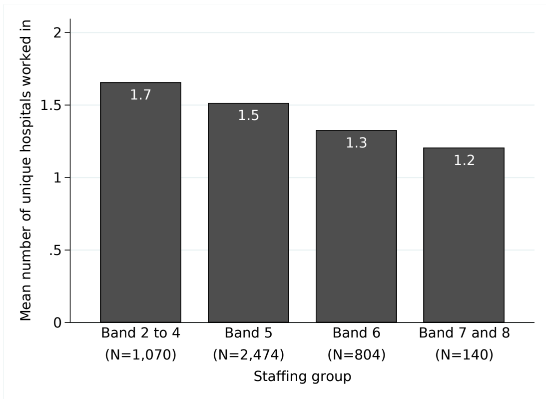
(b) Unique teams (mean)



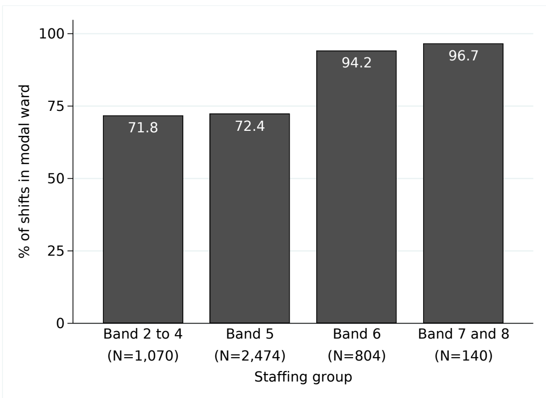
(c) Unique wards (mean)



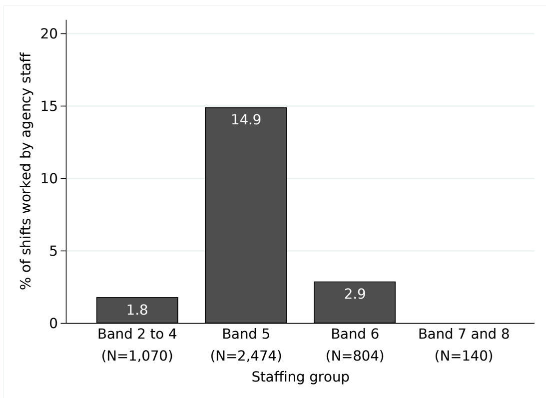
(d) Unique hospitals (mean)



(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).

Table 2: Summary statistics: patient outcomes and characteristics

	Mean	Std. Dev.
Number of patients treated in the unit	20.61	8.46
Age	64.18	9.86
Female (%)	0.46	0.18
Non-White (%) †	0.52	0.16
Elixhauser Comorbidity Index score	3.69	1.17
Average length of hospital stay (days) ‡	28.53	17.08
Patient death (binary)	0.063	0.243
Death rate per 1,000 patients	3.58	16.20
Observations	18,922	
Hospitals	3	
Wards	52	

Note: Unit of analysis is the hospital unit-day. Patient characteristics weighted by the hours spent in the unit on the day in question.
† Non-white includes those patients for whom ethnicity was recorded as unknown.
‡ Calculated as the average for the patients treated in a given unit on a given day, rather than the average for the universe of patients.

Table 3: 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 Empirical strategy

4.1 Baseline model

We begin by estimating the effect of a shortage of team members with less and more human capital. We estimate the following baseline fixed-effects logit model:

$$\begin{aligned} y_{it} &= \mathbb{1}(y^*_{it} > 0) \\ y^*_{it} &= \beta_1 FillRateNA_{it} + \beta_2 FillRateRN_{it} + \delta PatientChars_{it} + \lambda_t + \eta_i + \epsilon_{it} \end{aligned} \tag{6}$$

where y_{it} is a binary indicator of a patient death in ward i on day t (where the ward-day is the team). Quantity is measured by the fraction of planned NA and RN hours that were actually worked, denoted by $FillRateNA_{it}$ and $FillRateRN_{it}$, respectively (as formally defined in Section 2.4). $PatientChars_{it}$ captures time-varying characteristics of patients in ward i on day t. We include the number of patients treated in ward i on day t to control for variation in workload.

λ_t captures time effects (day of week; month; and a dummy variable indicating a public holiday). We include ward fixed effects (η_i) to capture time-invariant differences across hospital wards (such as their medical specialty, patient mix, hospital management quality, and physical layout). Identification comes from from within-ward variation in staffing levels and team composition. The extent of this within-ward variation is illustrated in Table 3. ϵ_{it} is an error term, assumed to be logistically distributed, with standard errors clustered at the ward level.

β_1 and β_2 are the coefficients of interest, and represent the impact of a change in the level of NA and RN staffing on the odds of a team experiencing a patient death, respectively. 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.[28]

4.2 Threats to identification

Identification in Equation (6) comes from the assumption that staffing is exogenous to outcomes. The institutional set-up makes such an assumption plausible. Nurse rosters are generated 3 months in advance to ensure that each ward contains the appropriate

[28]See Appendix Section A.1 for a full set of definitions and summary statistics.

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 consider these here.

4.2.1 Staffing levels and patient mix

Planned staffing levels are exogenous to eventual 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 adjust to changes in patient severity, such that there is a positive association between patient severity and staffing, this would bias our estimates upwards. Our model predicts a negative relationship between staffing and the risk of death; if staffing is endogenous to patient severity, 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 of all, 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 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 this key assumption. First, we examine the relationship between staffing levels on day t in ward i and observable patient severity and find no correlation (see Figure A3 and Table A2). We examine whether agency staffing – the margin that could be most easily used to respond to changes in severity – is a function of patient severity and find no evidence that agency staffing is higher when average patient severity is greater (see Table A3). If anything, greater patient severity – as measured by the Elixhauser Cormorbidity Index – is associated with severity is associated with *lower* levels of agency staffing (Figure A4). And, more generally, we are not able to predict the level of RN staffing using the observable characteristics of the patients under their care, consistent with random assignment (see 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 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 (see Table A4).

4.2.2 Staff shortages and ward performance

A second threat to our identification strategy is if staff shortages are endogenous to recent or expected ward performance. That is, staffing is lower because of 'no-shows' that are a response to recent patient outcomes. Appendix 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), with a small effect statistically significant 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 death two or three days previously. There are no statistically significant effects on NA staffing, nor on 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 (Appendix Table A6). For nurses we find no statistically significant relationship between the deaths and subsequent sickness in the previous three days. 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 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 suggests that staff shortages are not a response to recent or expected patient mortality outcomes.

4.2.3 Other staffing adjustments

A final concern is that because our data only includes information on nurse staffing, we are unable to capture any endogenous responses from other staffing 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. 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. Our inclusion of ward and month fixed effects will capture permanent differences in doctor quality across wards and any effects of seasonal doctor rotation, respectively.

4.3 The team as the unit of analysis

Our analysis is at the team level, motivated by our focus on the impact of team characteristics on outcomes. An alternative approach would be to estimate the model at the patient level and replace the dependent variable in equation (6) as the individual patient outcome and average patient characteristics with individual patient characteristics. However, the staffing variables would not vary at the patient level, so the only difference would be that we have an additional concern about unobserved patient severity. Averaging across all patients on the team (i.e. taking a team approach) reduces this concern. Analysis at the patient level would also raise the issue of harvesting: a patient may not die at time t, but die at $t + 1$, when in fact it was treatment at time t that caused their death.

Alternatively, we could use as the unit of analysis the patient hospital spell. One challenge with this approach is that patients with a longer hospital spell are both more likely to be exposed to various staffing events (for the simple fact that they are treated by a greater number of teams) and more likely to be severely ill and thus more likely to die. Length of stay is therefore correlated with our primary outcome of interest and with exposure to events like staff absences. Further, if we aggregated the staffing variables over the whole period of the patient's stay, for patients with longer stays we would expect such averages to revert to the mean, making identification difficult.

Finally, a relatively large proportion (more than one-third) of patient spells last a single day, and more than half are two days or shorter, precluding an event study design at the patient level due to a lack of ability to test for pre-trends.

5 Results

5.1 Overall staffing and general human capital

Table 4 reports our baseline results for quantity divided into higher and lower human capital.[29] 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 2.4), 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.

We find that an increase in NA staffing has no statistically significant impact on the odds of a patient death. In contrast, higher levels of RN 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.[30,31]

The size of the team therefore matters: teams with lower quantities of staffing, relative to target, have worse patient outcomes. But as important is the composition of the team. There are significant returns to qualifications: teams with higher levels of general human capital (as measured by a greater share of RNs rather than NAs in the team) are more productive.

The second column distinguishes between nurses of different seniority within the broad category of RNs. Those in higher pay bands have, on average, more experience and training, and thus can be considered to 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. These estimates are not statistically significantly different from one another, but the ordering and relative magnitudes of the point estimates are indicative. These results suggest that, for the

[29]These are also shown graphically in Appendix Figure A1.

[30]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 4, $0.9874^{8.6} \approx 0.9$, equivalent to a 10% reduction.

[31]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 seem broadly in line with these.

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 nurse: the average band 7–8 nurse is about 2.2 times as productive as a band 5 nurse. This is consistent with higher returns to greater general human capital.

The third column of Table 4 distinguishes between RNs of different contract types: local, bank and agency (see Section 2.3) . The results suggest that RNs 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. Here the coefficients on agency and non-agency staffing are very similar, suggesting that the general human capital of the team is more important than contract type.[32]

5.2 Firm-specific, ward-specific and team-specific human experience

5.2.1 Firm-specific human capital

Table 5 focuses on the impact of firm-specific (in our case, NHS hospital group-specific) experience, after controlling for the quantity and seniority mix of staffing.[33] All columns include the NA fill-rate, along with the RN fill-rate, split out by pay band. We then additionally examine the impact of higher levels of firm-specific experience.

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 group – is associated with better team performance.The estimate is statistically significant at the 10% level.[34] The second column separates returns from firm specific experience for RNs from that of NAs and shows that there are returns to firm-specific experience only for RNs. 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). A one standard deviation increase in

[32] Our results include separate ward and month fixed effects. We additionally allow for the possibility of ward-specific time trends by including ward-month fixed effects (Appendix Table A8) and ward-quarter fixed effects (Appendix Table A9). Our results are robust to these alternative specifications and the magnitude of our main estimates are unchanged.

[33] These are also shown graphically in Appendix Figure A2.

[34] Adding firm-specific experience does not change the results for quantity of general human capital – the results for general human capital are similar to those in the second column of Table 4.

Table 4: 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 (0.004) [0.360]	1.003 (0.004) [0.356]	1.003 (0.004) [0.348]	1.003 (0.004) [0.348]
Share (ppt) of planned Registered Nurse (RN) hours filled	0.987** (0.006) [0.038]			
Share (ppt) of planned Registered Nurse (RN) hours filled by:				
Band 5 nurses		0.989* (0.006) [0.087]		
Band 6 nurses		0.984** (0.007) [0.020]		0.984** (0.007) [0.020]
Band 7 and 8 nurses		0.976** (0.011) [0.025]		0.976** (0.011) [0.026]
Local (regular) nurses			0.987** (0.006) [0.037]	
Bank nurses			0.986** (0.007) [0.038]	
Agency nurses			0.991 (0.006) [0.140]	
Band 5 Local and Bank nurses				0.989* (0.006) [0.084]
Band 5 Agency nurses				0.991 (0.006) [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. 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.

average firm-specific experience (2.5 years, as per Table 1) reduces the odds of a patient death by 17.0%, after controlling for the seniority mix. In contrast, a greater level of firm-specific experience among the NAs has no statistically significant impact. The third column of Table 5 provides indicative evidence that the returns to firm-specific experience increase more or less monotonically with experience, though the point estimates are not statistically significantly different from each other.

These results suggest that teams with 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 RNs but not NAs suggests that that general and firm-specific human capital are complementary.

5.2.2 Ward-specific human capital

We now turn to an examination of the potential returns to ward-specific human capital, analogous to plant-specific huan capital within a firm, which in our context (as discussed in Section 2) is a combination of location-specific and task-specific capital.

The first column of Table 6 examines the impact of recent work experience in the ward among NAs and RNs, conditional on the size and general human capital (seniority mix) 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 took place in the ward and similarly finds no significant effects.

In the third and fourth column, we employ our definition of a 'new joiner': an individual who is within their first 30 days on the ward, having not worked there in at least the previous 90 days.[35] We examine whether teams containing a new joiner perform differently to those who do not. There is 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.

5.2.3 Team-specific experience

In Table 7 we examine the effect of familiarity between team members. As above, all regressions control for the size and general human capital of the team, patient character-

[35] See Section 2.4.2 and the Appendix for a full set of definitions for all variables used in the analysis.

Table 5: 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 (0.003) [0.386]	1.003 (0.003) [0.331]	1.003 (0.003) [0.365]
Share (ppt) of planned Registered Nurse (RN) hours filled by:			
Band 5 nurses	0.989* (0.006) [0.082]	0.989* (0.006) [0.062]	0.988* (0.006) [0.067]
Band 6 nurses	0.986** (0.007) [0.037]	0.986** (0.007) [0.036]	0.987** (0.007) [0.049]
Band 7 and 8 nurses	0.979* (0.011) [0.066]	0.980* (0.011) [0.078]	0.980* (0.011) [0.088]
Average experience in the Trust (years)	0.939* (0.034) [0.079]		
Average experience in the Trust for NAs (years)		1.017 (0.015) [0.246]	
Average experience in the Trust for RNs (years)		0.928*** (0.021) [0.001]	
Percent of Registered Nurses (RNs) with: †			
1-3 years Trust experience			0.991** (0.004) [0.047]
3-7 years Trust experience			0.990** (0.005) [0.034]
7-11 years Trust experience			0.987*** (0.004) [0.003]
11-15 years Trust experience			0.988** (0.005) [0.016]
15+ years Trust experience			0.985*** (0.005) [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. See notes to Table 4 for a full list of controls.

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

Table 6: 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 (0.005) [0.865]	1.000 (0.002) [0.874]		
Average shifts in the ward in past 90 days (RNs)	0.998 (0.008) [0.823]			
Average % of hours in the ward in past 90 days (NAs)		1.000 (0.002) [0.874]		
Average % of hours in the ward in past 90 days (RNs)		0.999 (0.004) [0.844]		
NA new joiner in the team (binary) †			0.879 (0.072) [0.116]	0.880 (0.072) [0.118]
RN new joiner in the team (binary) †			1.131 (0.121) [0.251]	
Band 5 RN new joiner in the team (binary) †				1.195* (0.115) [0.065]
Band 6 RN new joiner in the team (binary) †				1.021 (0.111) [0.849]
Band 7-8 RN new joiner in the team (binary) †				0.793 (0.222) [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.

istics, and time trends. The results in the first column indicate that the average amount of shared experience in the team does not have a statistically significant effect on team performance, for neither NAs nor RNs. The second column examines whether the average amount of shared experience between RNs affects team performance. Again we find 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.

5.2.4 Summary of results

The results so far have demonstrated the importance of both team size and composition in this setting. For the average team, an extra RN doing a 12-hour shift would reduce the odds of a patient death by around 10%. Within the broad category of RNs, the addition of a senior nurse has around twice as much impact as adding a newly qualified nurse. There are statistically significant returns to firm-specific experience: increasing the average tenure in the hospital group among RNs by one year reduces the odds of a patient death by 7.2%, around two-thirds the impact of adding an additional nurse. We find little-to-no evidence of any returns to recent ward-specific or team-specific experience: teams composed of nurses with more (recent) experience working in the ward, or with their fellow team members, do not perform systematically differently.[36] To further examine, this Appendix Table A12 reports the coefficients when all three measures of specific human capital (firm-, ward- and team-specific) are included alongside the size and general human capital of the team. The results reaffirm the importance of firm-specific experience among registered nurses and again indicate no (statistically significant) returns to ward- or team-specific experience.

5.3 Further analyses: team spillovers, unexpected absences and heterogeneity

5.3.1 Team spillovers

The analysis thus far has focused on the contemporaneous impacts of team composition on the outcomes of patients under the care of the team. This would not capture any

[36] The limited returns to ward-specific experience that we find contrasts with Bartel et al. (2014) who found significant returns to unit-specific experience among nurses in VA hospitals in the USA.

Table 7: 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 (0.010) [0.725]			
Average number of shifts with other team members in past 90 days (RNs)	0.993 (0.012) [0.552]			
Average number of shifts with other RNs in past 90 days (RNs)		0.998 (0.009) [0.791]		
Average share of team members worked with before (NAs)			1.005 (0.004) [0.155]	
Average share of team members worked with before (RNs)			0.994 (0.005) [0.236]	
Average share of team members worked with 10+ times before (NAs)				1.001 (0.002) [0.547]
Average share of team members worked with 10+ times before (RNs)				0.996 (0.003) [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.

spillovers onto other teams. Appendix Table A4 shows that short staffing is not associated with an increase in patient outflow. In other words, we find no evidence that teams respond to being short-staffed by sending patients to other parts of the hospital. It could be, though, 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 Appendix Table A11. 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.[37] We find no evidence of team size or composition on subsequent deaths among patients treated by the team, indicating that there are limited lagged effects on patients and/or spillovers onto other teams.

5.3.2 Unexpected absences

Variation in staffing levels and team composition can be caused by staff absences. Some forms of absences are likely to result in a greater degree of disruption to a team than others. When absences are known about in advance (such as a pre-agreed period of holiday or training leave), rosters can be re-arranged in advance with little disruption. Unexpected absences, such as those due to short-term sick leave, may be harder to adjust to, as managers may struggle more to find an appropriate replacement at short notice. Our data contains some limited information on staff absences. In Appendix Table A10, we examine the impact of staff absences on the productivity of the team. We find results qualitatively consistent with our main results: an unexpected absence of a senior nurse (in pay band 7 or 8) is associated with impaired team performance and the effect size is large.

5.3.3 Heterogeneity by patient severity

Our data contains no information on cause of death or nurse-patient interactions, limiting our ability to identify mechanisms through which team size and composition affect patient mortality. It may be that certain types of patient suffer when teams are short-staffed, which would indicate where resource-constrained teams focus their efforts. We examine this in Appendix Tables A13 and A14, by separately examining the impact of nurse staffing on deaths of low- and high-severity patients.[38] The results in Table

[37]We have also conducted the same analysis with longer time periods and find very similar results.

[38]The median Elixhauser Cormorbidity Index score among patients who die in hospital in our sample was 5. We define a low-severity patient as one with an Elixhauser score of less than 5, and a high-severity patient as one with an Elixhauser score of 5 or above. A death of a low-severity and high-severity patient

A13 suggest that the impacts of shortages of RNs (and in particular senior RNs) fall almost entirely on low-severity patients; the results in Table A14 similarly demonstrate that the impacts of lower firm-specific experience fall predominantly upon low-severity patients. This is consistent with the idea that short-staffed teams, and teams lacking in seniority and firm-specific experience, focus on providing care to the most (observably) sick patients, and that care for other (less observably sick) patients is rationed, with those patients more likely to die as a result. It could also be that senior and more experienced nurses are better at identifying patients with a greater mortality risk not predicted by readily observable characteristics (such as age and recorded diagnoses). Additional analysis (not reported here) indicates that this result is not driven by or confined to a particular set of wards or medical specialties (such as intensive care or geriatric wards); it rather appears to be more general.[39]

This analysis has important 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. This implies a greater overall cost to social welfare.

6 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 among patients under their care.

We use qualifications and rank 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 (NA) has no impact on the likelihood of a patient death, but a greater number of hours worked by Registered Nurses (RNs) is associated with lower inpatient mortality, with the magnitude of these effects increasing with qualifications and experi-

is experienced by 3.6% and 3.0% of teams in our sample, respectively (Appendix Table A1).

[39]These results are available from the authors upon request.

ence. The most senior nurses, who are responsible for managing their ward, are around 2.2 times as productive as their newly qualified counterparts. Notably, these productivity differences are somewhat wider than the pay differences between the most junior and more senior nurses. Our results are almost entirely driven by mortality impacts on relatively low-severity, rather than high-severity, patients. This is consistent with the idea that short-staffed teams or teams lacking in senior staff focus effort on patients who are more severely ill.

We also find evidence that a team performs better when its constituent members have more experience in working for the firm (hospital), though there is much less 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.

A key conclusion of this analysis is that nurse staffing shortages, endemic to the English NHS, have adverse consequences for patient care but that efforts to substitute less skilled individuals for more highly individuals are unlikely to boost team performance.[40] Given the returns to hospital experience that we find, one alternative strategy would be to increase efforts to retain more skilled staff.

[40] This also relates to recent findings for the substitution of nurse practitioners for doctors in the VA system in the US (Chan and Chen, 2022).

References

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.

Autor, D. H. and S. N. Houseman (2010, July). Do Temporary-Help Jobs Improve Labor Market Outcomes for Low-Skilled Workers? Evidence from "Work First". American Economic Journal: Applied Economics 2(3), 96–128.

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.

Bryson, A. (2013, April). Do temporary agency workers affect workplace performance? Journal of Productivity Analysis 39(2), 131–138.

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.

Chan, D. C. and Y. Chen (2022). The Productivity of Professions: Evidence from the Emergency Department. Unpublished manuscript.

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.

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. pp. 45.

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.

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.

Jahn, E. J. and D. Pozzoli (2013, October). The pay gap of temporary agency workers — Does the temp sector experience pay off? Labour Economics 24, 48–57.

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.

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.

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.

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 and definitions

Table A1: Summary statistics: all variables used in 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

† Non-whote includes patients with ethnicity recorded as .

‡ 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.

Staffing quantity and general human capital definitions

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

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

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

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

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

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

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

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

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

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

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

Firm-specific human capital

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

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

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

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. 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.

Ward-specific human capital

$$\text{Individual } i \text{ shifts in ward } k \text{ in the past 90 days} = Y_{ihkt} = \sum_{\tau=t-90}^{t-1} \mathbb{1}(i \in T_{\tau k}) \forall k \in K(h)$$

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

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

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

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

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

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

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

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

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

Team-specific human capital

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

Individual i's average shared work experience with team $= Z_{-ihkt} = \frac{1}{|T_{kt}|-1} \cdot \sum_{j \in T_{kt}, j \neq i} E(i,j,t)$

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

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

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

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

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

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

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

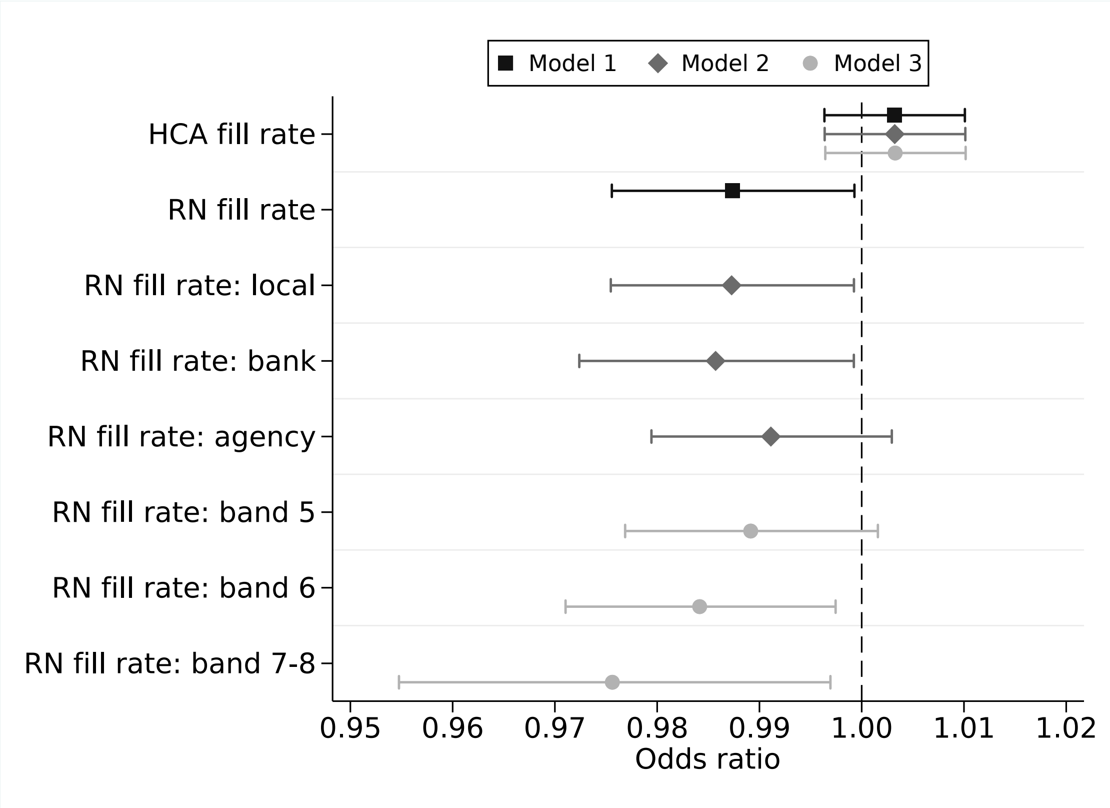
Mean % of team members worked with in past 90 days $= \frac{1}{|T_{kt}|} \sum_{i \in T_{kt}} WorkedWith_{-ihkt}$

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

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. The mean % of team members worked with ever or 10+ times is defined separately for NAs and RNs.

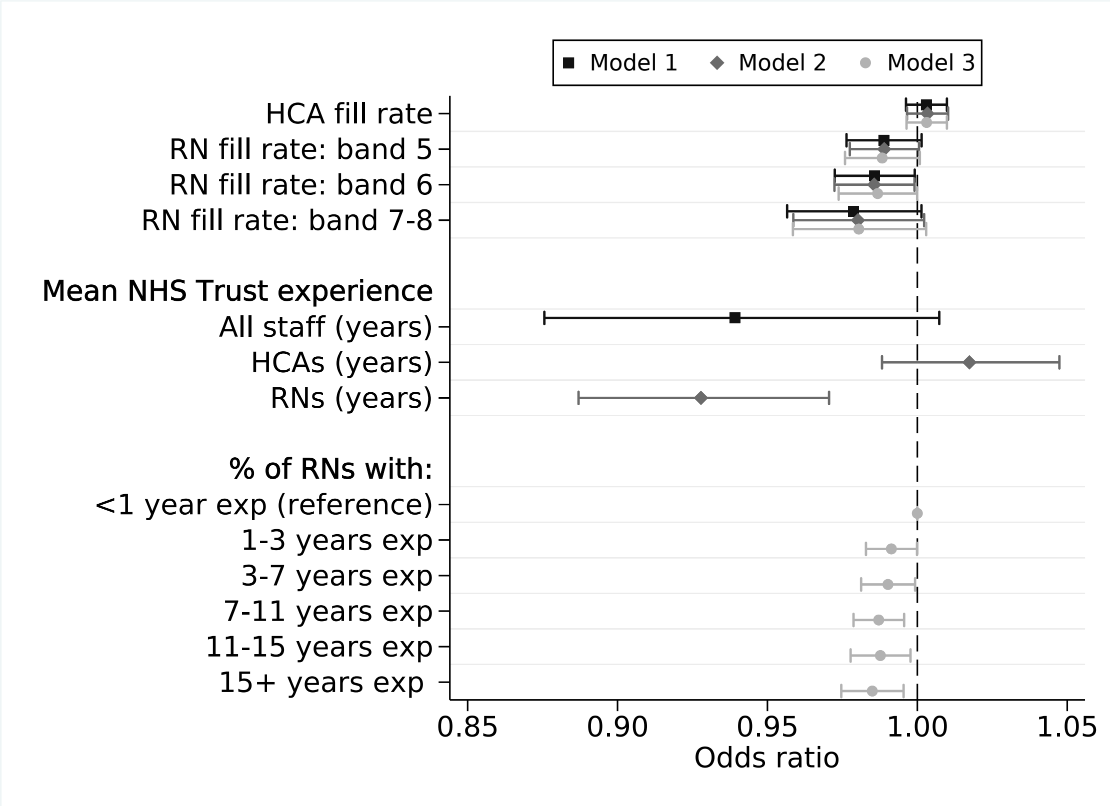
A.2 Main results displayed as coefficient plots

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.

A.3 Additional results and figures

Table A2: Relationship between patient severity and team composition

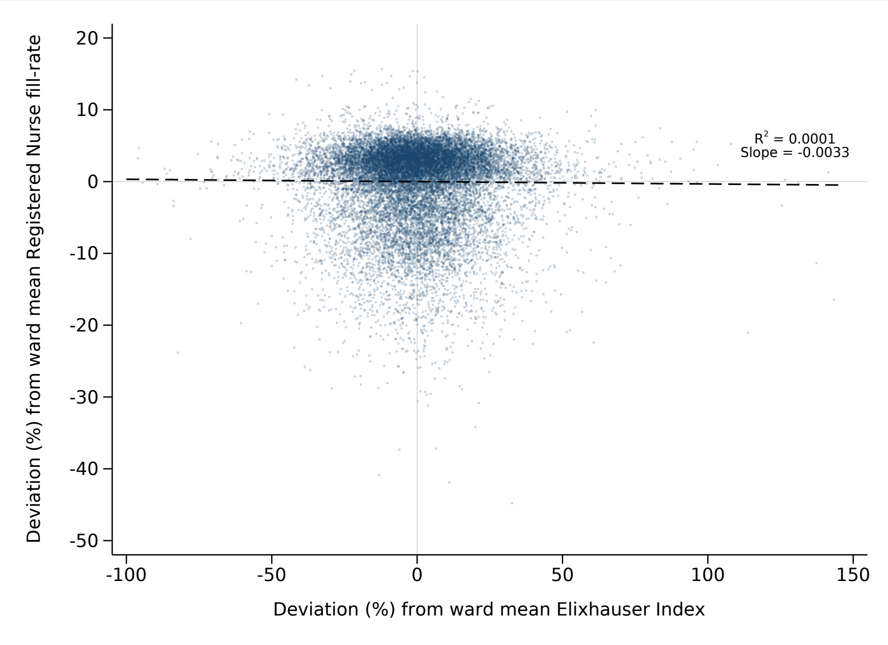
	Outcome: mean patient Elixhauser Cormorbidity Index			
	(1)	(2)	(3)	(4)
Share (ppt) of planned Nursing Assistant (NA) hours filled	-0.001 (0.001) [0.270]	-0.001 (0.001) [0.275]	-0.001 (0.001) [0.278]	-0.001 (0.001) [0.277]
Share (ppt) of planned Registered Nurse (RN) hours filled	-0.000 (0.001) [0.784]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

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

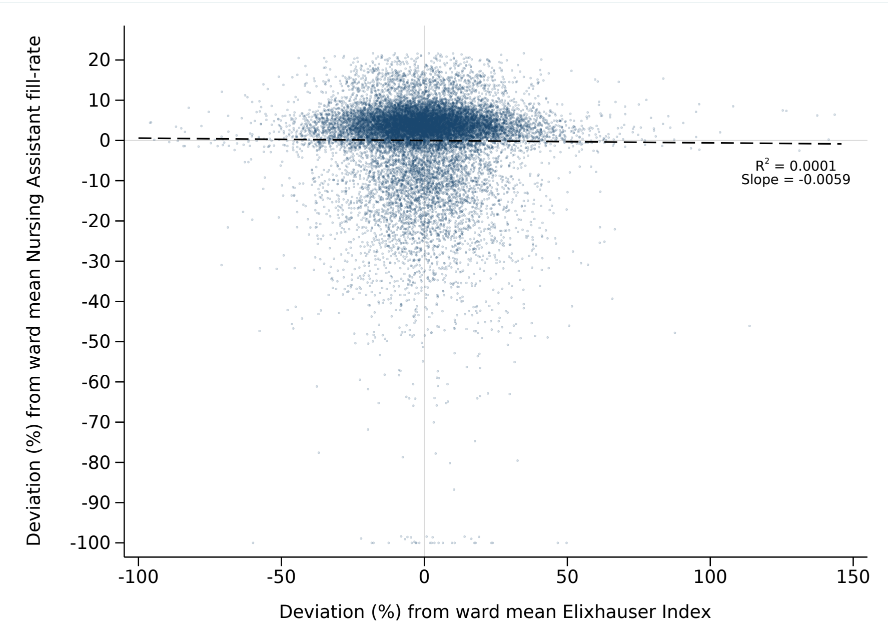
Coefficients shown are from a fixed-effects regression, with hospital ward fixed effects. Robust standard errors, clustered at the hospital unit level, are shown in parentheses; p-values are shown in square brackets. The unit of analysis is the hospital unit-day. Time controls include a dummy for each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

Figure A3: Relationship between staffing levels and observable patient comorbidity, as measured by Elixhauser Cormorbidity Index

(a) Registered Nurses

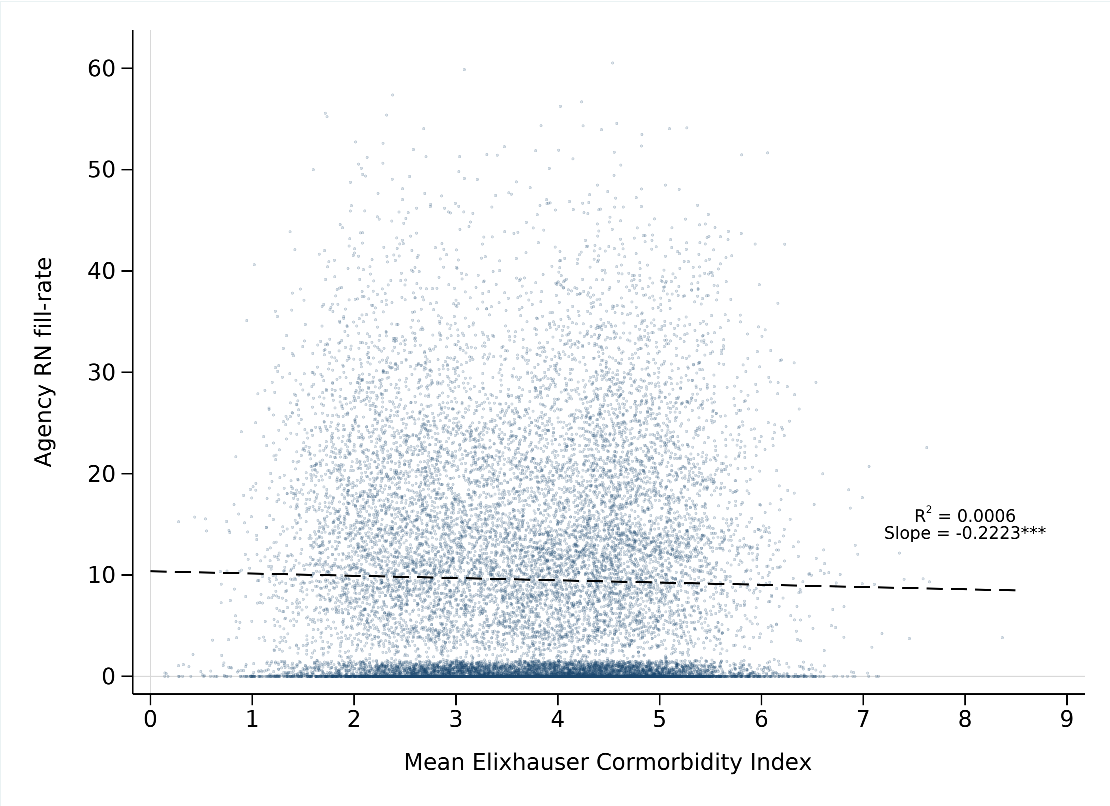


(b) Nursing Assistants



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 A4: 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

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

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 (0.005) [0.827]	0.001 (0.005) [0.832]	0.001 (0.005) [0.800]	0.001 (0.005) [0.800]
Share (ppt) of planned Registered Nurse (RN) hours filled	0.005 (0.009) [0.600]			
Share (ppt) of planned Registered Nurse (RN) hours filled by:				
Band 5 nurses		0.005 (0.010) [0.609]		
Band 6 nurses		-0.003 (0.011) [0.781]		-0.003 (0.011) [0.769]
Band 7 and 8 nurses		0.023 (0.025) [0.381]		0.022 (0.025) [0.386]
Local (regular) nurses			0.005 (0.009) [0.595]	
Bank nurses			0.007 (0.011) [0.538]	
Agency nurses			-0.001 (0.010) [0.937]	
Band 5 Local and Bank nurses				0.006 (0.010) [0.563]
Band 5 Agency nurses				-0.000 (0.010) [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 (0.245) [0.827]	-0.232* (0.120) [0.058]	0.094 (0.273) [0.731]
Death in the ward on day T-2	-0.029 (0.394) [0.941]	-0.072 (0.154) [0.640]	-0.284 (0.343) [0.411]
Death in the ward on day T-3	-0.119 (0.217) [0.585]	0.239 (0.150) [0.118]	0.195 (0.261) [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 (0.262) [0.381]	-0.095 (0.095) [0.322]	0.225 (0.262) [0.395]
Death in the ward on day T-2	0.268 (0.269) [0.323]	-0.057 (0.148) [0.703]	-0.018 (0.230) [0.938]
Death in the ward on day T-3	0.016 (0.190) [0.934]	0.028 (0.135) [0.834]	0.186 (0.246) [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 (0.082) [0.171]	0.959 (0.070) [0.570]
Death in the ward on day T-2	1.364*** (0.140) [0.002]	1.025 (0.076) [0.734]
Death in the ward on day T-3	0.925 (0.152) [0.635]	1.057 (0.122) [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 (0.157) [0.986]	1.004 (0.150) [0.980]	1.000 (0.149) [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 (0.004) [0.533]	1.002 (0.004) [0.532]	1.002 (0.004) [0.525]	1.002 (0.004) [0.525]
Share (ppt) of planned Registered Nurse (RN) hours filled	0.988* (0.007) [0.065]			
Share (ppt) of planned Registered Nurse (RN) hours filled by:				
Band 5 nurses		0.990 (0.007) [0.117]		
Band 6 nurses		0.984** (0.008) [0.040]		0.984** (0.007) [0.038]
Band 7 and 8 nurses		0.975** (0.011) [0.022]		0.975** (0.011) [0.021]
Local (regular) nurses			0.987* (0.007) [0.053]	
Bank nurses			0.987* (0.007) [0.063]	
Agency nurses			0.993 (0.007) [0.304]	
Band 5 Local and Bank nurses				0.989 (0.007) [0.100]
Band 5 Agency nurses				0.992 (0.007) [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 (0.004) [0.454]	1.003 (0.004) [0.450]	1.003 (0.004) [0.445]	1.003 (0.004) [0.444]
Share (ppt) of planned Registered Nurse (RN) hours filled	0.987** (0.006) [0.041]			
Share (ppt) of planned Registered Nurse (RN) hours filled by:				
Band 5 nurses		0.988* (0.007) [0.080]		
Band 6 nurses		0.984** (0.007) [0.029]		0.984** (0.007) [0.029]
Band 7 and 8 nurses		0.975** (0.011) [0.019]		0.975** (0.011) [0.019]
Local (regular) nurses			0.986** (0.006) [0.034]	
Bank nurses			0.986** (0.007) [0.041]	
Agency nurses			0.991 (0.007) [0.210]	
Band 5 Local and Bank nurses				0.988* (0.007) [0.069]
Band 5 Agency nurses				0.991 (0.007) [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 staff absences on patient mortality

	Outcome: patient death			
	(1)	(2)	(3)	(4)
Share (ppt) of planned Nursing Assistant (NA) hours filled	1.003 (0.004) [0.345]	1.003 (0.004) [0.346]	1.003 (0.004) [0.349]	1.003 (0.003) [0.357]
Share (ppt) of planned Registered Nurse (RN) hours filled by:				
Band 5 nurses	0.989* (0.006) [0.081]	0.989* (0.006) (0.006)	0.989* (0.006) [0.073]	0.990* (0.006) [0.099]
Band 6 nurses	0.984** (0.007) [0.017]	0.984** (0.007) [0.021]	0.984** (0.007) [0.014]	0.978** (0.011) [0.040]
Band 7 and 8 nurses	0.975** (0.011) [0.024]	0.977* (0.012) [0.056]	0.975** (0.011) [0.025]	0.978** (0.011) [0.040]
Absence for a reason other than sickness, lasting $\leq$ 7 days:				
RN of any band	1.041 (0.073) [0.565]			
Band 7–8 nurse		1.030 (0.085) [0.717]		
Absence due to sickness, lasting $\leq$ 7 days:				
RN of any band			0.966 (0.076) [0.656]	
Band 7–8 nurse				1.632*** (0.239) [<0.001]
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. See notes to Table 4 for a full list of controls.

Table A11: 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 (0.003) [0.638]	0.998 (0.002) [0.496]	0.998 (0.002) [0.282]
Share (ppt) of planned Registered Nurse (RN) hours filled by:			
Band 5 nurses	0.994 (0.005) [0.208]	0.998 (0.005) [0.769]	1.000 (0.005) [0.966]
Band 6 nurses	0.998 (0.006) [0.654]	0.999 (0.006) [0.872]	1.001 (0.006) [0.870]
Band 7 and 8 nurses	0.989 (0.010) [0.234]	0.992 (0.008) [0.351]	0.999 (0.008) [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 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: 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 (0.005) [0.544]	1.005 (0.004) [0.171]
Share (ppt) of planned Registered Nurse (RN) hours filled by:		
Band 5 nurses	0.984* (0.009) [0.063]	0.998 (0.011) [0.851]
Band 6 nurses	0.980** (0.009) [0.028]	0.991 (0.010) [0.384]
Band 7 and 8 nurses	0.964** (0.015) [0.018]	0.993 (0.017) [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. 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 A14: 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 (0.006) [0.722]	1.007 (0.005) [0.160]
Share (ppt) of planned Registered Nurse (RN) hours filled by:		
Band 5 nurses	0.984* (0.008) [0.060]	0.996 (0.011) [0.731]
Band 6 nurses	0.984* (0.009) [0.071]	0.990 (0.011) [0.351]
Band 7 and 8 nurses	0.977 (0.016) [0.175]	0.987 (0.016) [0.443]
Average experience in the Trust for NAs (years)	1.008 (0.021) [0.683]	1.018 (0.017) [0.289]
Average experience in the Trust for RNs (years)	0.909*** (0.028) [0.002]	0.952 (0.033) [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 A5: Distribution of time of recorded patient death

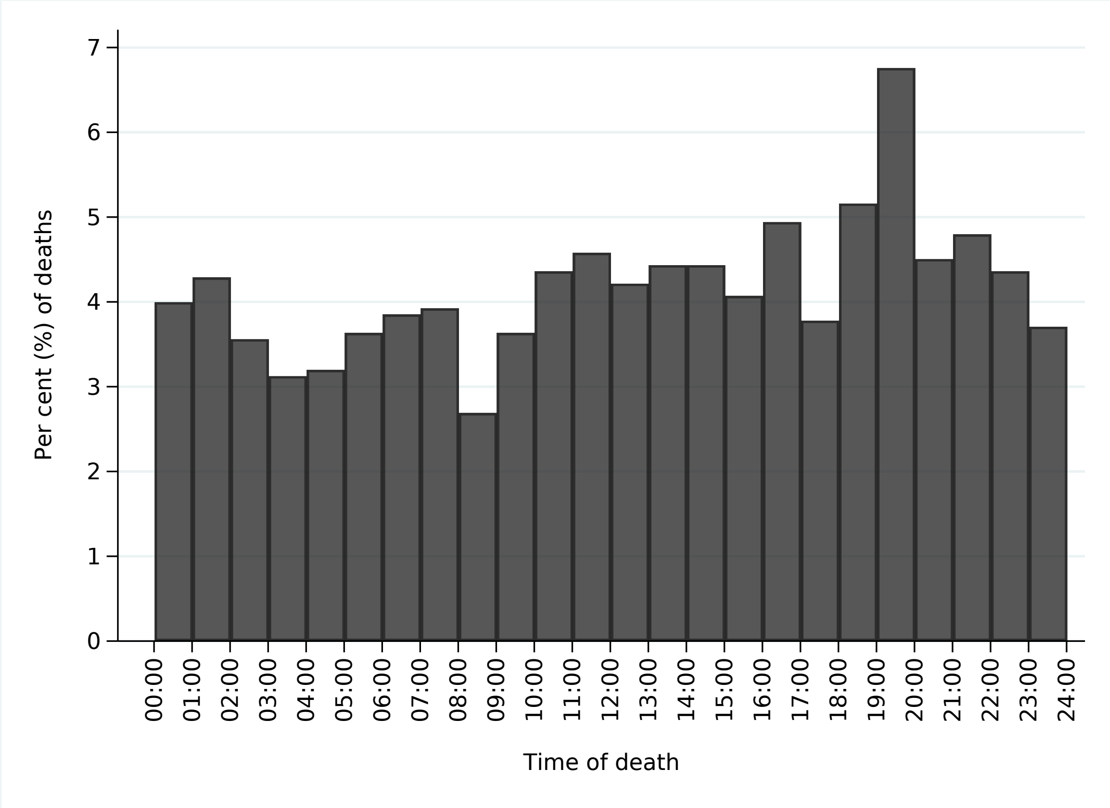


Figure A6: Frequency of patient deaths across wards