HARVARD | BUSINESS | SCHOOL

The Impact of CEOs in the Public Sector: Evidence from the English NHS

Katharina Janke
Carol Propper
Raffaella Sadun

Working Paper 18-075

The Impact of CEOs in the Public Sector: Evidence from the English NHS

Katharina Janke
Lancaster University

Carol Propper
Imperial College London

Raffaella Sadun
Harvard Business School

Working Paper 18-075

Funding for this research was provided in part by Harvard Business School.

The Impact of CEOs in the Public Sector: Evidence from the English NHS

Katharina Janke, Carol Propper and Raffaella Sadun*

September 15, 2020

Abstract

Governments worldwide have sought to reform the delivery of public services by mimicking private sector governance models that grant CEOs greater autonomy and give them responsibility for meeting key government targets. We examine the effectiveness of this approach in the context of English public hospitals, complex organizations with multi-million turnover. We find little evidence of CEOs' impact on hospital production, though estimated pay differentials suggest that the CEOs are perceived to be differentiated by the market. These findings are not driven by endogenous sorting of CEOs. The results question the effectiveness of top-down managerial approaches to improve public sector performance. (JEL H51, I11, L32, M12)

*Janke: Lancaster University, Furness Building, Lancaster, LA1 4YG, UK (email: k.jankemarie@lancaster.ac.uk). Propper: Imperial College Business School, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK, and CEPR (email: C.Propper@imperial.ac.uk). Sadun: Harvard Business School, Soldiers Field, Boston, MA, 02163, CEPR and NBER (email: rsadun@hbs.edu). We thank for their comments seminar participants at Birkbeck, Bocconi, Brisbane, Boston University, Dartmouth, CEPR, Harvard, LBS, LSE, Leuven, MIT, Monash, Munich, NBER Summer Institute, Oxford, Toronto.

1 Introduction

From the 1980s onward, governments worldwide sought to reform the delivery of public services by moving away from traditional centralised bureaucracies. Key features of these reforms were changes towards more specialized and autonomous organizations, coordinated by means of market mechanisms and contractual relationships rather than hierarchies of authority. This model, which later became known as the New Public Management (NPM) (Pollitt & Bouckaert 2000), puts much more emphasis on the role of senior managers, giving them greater autonomy to run their organizations while holding them accountable through manager-specific compensation policies, performance-related pay, and tighter monitoring and dismissal for failure to meet explicit performance targets (Besley & Ghatak 2003, Le Grand 2003, Dixit 2002).

A central plank of the success of this model is that senior managers are able to have an impact on the public service delivery organisations they run. The impact of CEOs for private sector organisations has been demonstrated in a number of influential studies, beginning with Bertrand & Schoar (2003).[1] The evidence of the role of CEOs for public sector organizations, however, has so far been limited to the study of relatively small organizations such as schools and development projects, where top managers may have a greater chance of having an impact. In contrast, the effect of top managers on large and complex public sector organizations has hardly even been examined. Can CEOs make a difference in this context? If the answer is no, it calls into question the general applicability of this key component of the New Public Management model.

To address this question, we examine the impact of CEOs of very large and complex public sector organizations that are perhaps the canonical example of the application of the NPM model: hospitals in the English National Health Service (NHS). NHS hospitals are large, with on average 4,500 employees, treat over 75,000 patients per year and have multi-million pounds turnover. From the late 1980s onwards, the UK government embarked on a large reform program that replaced a traditional administrative approach to hospital management with a highly decentralized managerial model, in which CEOs were given responsibility for the management and performance of individual public hospitals, and individual hospital boards could select and reward individual CEOs in an autonomous fashion. While hospitals remained under public ownership, central budgets were replaced by local contracts for service delivery, which were won in competition with other NHS hospitals.[2] CEOs were held accountable through clinical and financial targets set by

[1] Recent examples include Bamber et al. (2010), Dejong & Ling (2013) and Bennedsen et al. (2006).

[2] There was also competition with a small private sector for non-urgent (elective) services but this competition was limited during the period we study.

the central government, against which the performance of the CEOs was to be regularly assessed. Failure to meet these targets could lead to public 'naming and shaming' and dismissal of the CEO. These changes were accompanied by frequent movements of the same CEOs across different but comparable NHS hospitals, providing an ideal setting to examine the role of managerial discretion for hospital performance in isolation from other persistent differences in hospital characteristics.

We examine whether CEOs were able to affect the performance of their hospitals using a wide variety of outcomes. We start with hospital production measures that have been used as targets by the government regulator for NHS CEOs at various points during (and for some measures before) the 14-year period we examine. These measures cover both clinical and non-clinical aspects of hospital production. We also examine a variety of other metrics considered to be key for the efficiency and quality of care provided by hospitals, such as staff satisfaction, which has been shown to be associated with clinical quality of care as well as recruitment and retention.[3] Additionally, the NHS requires hospitals to publish the pay awarded to their top managers, allowing us to examine whether there are differences in CEO compensation, which we would expect to see if hospital boards perceived CEOs to be differentiated in terms of their managerial quality.

To examine whether CEOs can affect production we use two approaches: a parametric approach pioneered by Bertrand & Schoar (2003) and a non-parametric approach. The first approach assesses whether deviations in a hospital production measure from the expected level, determined by a regression of the measure on time-varying firm characteristics and firm fixed effects, can be replicated by the same CEO in another hospital. The second compares changes in targeted performance measures after a CEO turnover event to changes experienced by matched hospitals without such an event.

Using both approaches, we find little consistent evidence that CEOs are able to generate persistent performance differentials with respect to the key targets that governments have set for them. This result holds for a larger set of (targeted and non-targeted) input, throughput and output measures. On the other hand, we find that the market perceives hospital CEOs as differentiated in that there are considerable and persistent (i.e., independent of the hospitals they work for) differences across managers. Moving from the 25th percentile to the 75th percentile of the distribution of CEO fixed effects in pay is associated with a 14% increase in pay relative to mean CEO pay. While not large in comparison to private sector CEO differences in rewards (many of which come from stock options and not pay) these persistent differences are substantial in the public sector context.

[3]Staff recruitment and retention are central in hospital settings as labour makes up over 70% of inputs into production.

We examine a number of possible reasons for this apparent lack of CEO influence over hospital production. First, we examine whether the lack of persistence in the CEO fixed effects is to due the endogenous assignment of "good" CEOs to poorly performing hospitals, or to hospitals that have structural features which may negatively affect achievement of good performance. If this were the case, differences in pay across CEOs would not necessarily be mirrored by differences in hospital production, since the best CEOs would be systematically assigned to harder-to-change organizations. We also study whether better performing CEOs are more likely to exit the NHS, perhaps because they can aspire to higher pay relative to public sector employment. Second, we examine whether there is evidence of CEO-hospital match effects, i.e. whether specific types of CEOs may perform better when in a specific type of hospital. Such match effects would imply that CEOs matter, but only when the fit between the manager and the hospital is appropriate, thus explaining the lack of general effects of CEOs. Third, we examine whether the lack of persistent CEO effects may be driven by the fact that mover CEOs, who are key to our identification strategy, tend to have short tenure. Tenure effects would imply that CEOs may in fact differ in terms of their potential effect on hospital performance, but these effects fail to materialize in the estimates given the short time horizons we observe the managers for in each hospital (on average 3.8 years).

We find little evidence of endogenous assignment. We do, however, find some evidence of match effects and tenure effects, which suggest that the government and hospital boards may overestimate the ability of CEOs to bring about change over short time periods and/or regardless of the circumstances they face.

Our results indicate that a key aspect of the NPM model does not appear to deliver, in that CEOs of large public hospitals do not necessarily impact key aspects of hospital production. This result stands in stark contrast with earlier findings in private sector and smaller public sector organizations. Various structural factors may account for the lack of a "CEO effect", including the fact that NHS CEOs, whilst well paid relative to other public sector CEOs, may not be paid enough relative to the private sector to attract capable managers. However, the lack of a CEO effect may also be due, more broadly, to the complexity of hospital production, which transcends the constraints imposed by public ownership. More generally, our results cast doubt on the effectiveness of a "turnaround CEO" approach – the model in which top managers frequently rotate to induce meaningful changes in performance – for large public sector organizations.

Our study complements prior research on the impact of CEOs in the public sector. This question has to date been investigated for either relatively small organizations or ones in which the level of task complexity is low. Several papers investigate the impact of principals on student performance. Böhlmark et al. (2016) present evidence of principal fixed effects in students' outcomes. Other papers include Branch et al. (2012), Coelli & Green (2012) and Grissom et al. (2015). Lavy & Boiko (2017) find that superintendents affect student performance. Bloom, Lemos, Sadun & Van Reenen (2015) find managerial practices adopted by school principals are correlated with school performance. Fenizia (2019) finds managers affect the productivity of workers in a government agency administering social security benefits in Italy. Examining World Bank employees, Limodio (2018) documents negative matching between high performing managers and low performing countries. Rasul & Rogger (2018) examine management practices in the Nigerian Civil Service and find they affect the behaviour of government bureaucrats.

There are, of course, many other issues that arise in the NPM model that have been extensively studied by economists (and others). Prominent amongst these are the limits to the use of targets to influence behaviour in public service delivery. Two key challenges identified by the literature are a) the presence of multiple and hard to measure outputs; and b) the potential conflicts associated with the use of financial incentives when employees are characterized by high public service motivation. Early studies of the use of performance targets in the public sector include Prendergast (2001, 2002) and Heckman et al. (2002). Dixit (2002) provides support for the use of low-powered financial incentives for bureaucrats, while Besley & Ghatak (2003) suggest that matching between mission orientated firms and employees replaces the need for financial incentives all together. There is a large, and growing, empirical literature examining the impact of mission matching, particularly in the provision of health and education programs in developing countries, where many field trials have been undertaken. A recent example is Ashraf et al. (2014). In the UK context, empirical studies of the impact of financial incentives on the performance of public service providers include Gravelle et al. (2010) and Propper et al. (2002) for family doctors and Burgess et al. (2017) for UK job placement agencies. Propper et al. (2010) examine the impact of performance targets on NHS hospitals.

The paper is organized as follows. Section 2 describes the components of the NPM and the evolution of the market for CEOs in the NHS, including the accompanying changes in the level and variance of CEO pay. The section also presents results from pay regressions using the Abowd et al. (1999) methodology. Section 3 describes the data. Section 4 provides details on our two econometric approaches. Section 5 present the results while Section 6 examines possible reasons for the lack of persistent CEO effects. Section 7 concludes.

2 Institutional Background

2.1 Key elements of the New Public Management in the English NHS

The NPM model implemented in the English NHS involved giving hospitals autonomy and relying on market mechanism while the central government would set targets against which hospital CEOs were to be assessed. We explain some key institutional features introduced by the reforms below.

Hospital Autonomy and Market Mechanisms From the early 1990s, English public hospitals started operating as free-standing organizations known as NHS Hospital Trusts, earning their revenue from contracts won in competition with other public hospitals. From the early 2000s, the government sought to further stimulate competition by placing contracts with a small number of private hospitals, known as Independent Sector Treatment Centres (ISTCs), that provided a selected set of planned operations and diagnostic tests. This policy was later expanded to include any private provider for all elective treatments.[4] The overall policy goal was for English NHS hospitals to operate subject to market forces rather than central guidance. Within this general policy framework, which applied to *all* hospitals, if a hospital achieved certain targets (relating primarily to financial performance and access) they were formally given a higher level of autonomy, known as Foundation Trust (FT) status. The aim was that all Trusts would get FT status by 2008, though in practice this was not achieved.[5]

Corporate Governance Structures The changes to hospital autonomy were supported by significant reforms to the management of hospitals, which gradually replaced a bureaucratic consensus management system with a general manager who had overall responsibility for service performance and management (Baggott 1994).[6] During the wave of mid-1990s market reforms, these general managers were renamed Chief Executives,

[4]ISTCs were privately owned and specialised in the provision of a limited set of planned treatments (e.g. joint replacements). They initially received favourable five-year contracts with revenue that did not vary with the number of patients treated (Naylor & Gregory 2009) but later all providers, public or private, were paid according to the same DRG-type tariffs.

[5]By the end of our sample period, 62% of hospitals in our sample had FT status. We control for FT status in all regressions in the paper.

[6]The initial push in favor of general managers and private sector managerial practices more broadly for the NHS came from the Griffiths' report of 1983 (Lord Griffiths was at the time the Deputy Chairman of the supermarket chain J. Sainsbury plc, and was tasked by Margaret Thatcher to study the management of the NHS).

and hospitals started being subject to corporate governance standards similar to the ones brought into private sector firms in the UK in 1992 (Cadbury 1992). The role of hospital boards was also strengthened, and they became responsible for managing the day-to-day operation of a hospital. Trust boards had to include the Chief Executive, a Finance Director, a Nursing Director and a Medical Director, but could have more positions.[7] These executive positions were matched by their non-executive director counterparts, who were hired with the expectation that they would need to dedicate at least three days a month to the hospital.[8] They were also, in contrast to hospital board members in the USA, remunerated. Boards generally met monthly (Jha & Epstein 2013) and were hands-on in terms of monitoring of the financial performance and the quality of care provided by the hospital.[9]

CEO Responsibilities In addition to managing day-to-day hospital operations, Chief Executives and their boards were responsible for delivering government policy, which was embodied both in targets and in guidance. Performance against targets was subject to close scrutiny by central government. During most of the period we study, targets were predominantly focused on financial performance and reducing waiting times.[10] From 2001 onwards, the central government regulator published hospital ratings, which were based on detailed quantitative data on both financial and process metrics. From 2011 the targets started including clinical quality metrics.[11] Missing key performance targets could place a CEO under threat of dismissal. Ballantine et al. (2008) document a strong association between a limited number of hospital performance measures and CEO turnover between 1998 and 2005. In sum, NHS CEOs became responsible for both meeting government targets and day-to-day operations of large and complex organizations operating in a potentially competitive market.[12]

[7]Good practice for NHS boards is set out in `https://www.leadershipacademy.nhs.uk/wp-content/uploads/2013/06/NHSLeadership-HealthyNHSBoard-2013.pdf`. A statutory instrument (`http://www.legislation.gov.uk`) sets out the board voting members.

[8]The minimum number of non-executive directors was a Chair and three others.

[9]Jha & Epstein (2013) found that approximately 40% of Boards had received formal training in quality management and that they frequently reviewed and monitored quality of care issues. 98% of Boards reported that quality of care was on the agenda at every board meeting, 77% reported to actively use patient safety data to provide staff feedback.

[10]For example, achieving FT status was conditional on meeting these targets. Propper et al. (2010) provide details on waiting times targets and their impact on performance.

[11]The focus on clinical quality was the result of an extensive investigation into systemic failure at a single hospital, Mid-Staffordshire. The final recommendations were published in 2013 in `https://www.gov.uk/government/publications/report-of-the-mid-staffordshire-nhs-foundation-trust-public-inquiry`.

[12]Bloom, Propper, Seiler & van Reenen (2015) show that NHS hospital management quality responds positively to greater product market competition.

CEO Selection Boards had guidance from the central government regulator on making senior appointments, including the CEO. CEOs were hired in a manner similar to those of private sector firms. The Chair of the Board and the appointment committee would generally use private sector headhunters for the selection and hiring of the CEO, and they would also either consult with, or include in the decision making process, a representative from the national government organization responsible for overseeing the NHS.[13] CEOs had career paths that involved movement between different parts of the NHS. They were predominantly individuals who had entered the NHS relatively early in their career (either as managers or as clinicians), and who were typically promoted by moving between organizations in the NHS. Thus, a typical CEO would have considerable experience of working across a number of NHS organizations. However, individuals who had private sector experience (either as private consultants in the health sector or in running private sector organizations) were also sometimes appointed to CEO positions. There was also movement of CEOs to the private sector, often to posts within the wider healthcare sector.

CEO Remuneration CEO remuneration was set by the Board. From 2003 hospitals that had achieved FT status were free to set CEO and other executive and non-executive director pay, decided upon by the remuneration committee as in any private company.[14] The remuneration committee could also decide whether to link CEO (and other director) remuneration to corporate and individual performance. Performance, particularly against government targets, could affect CEO pay, job tenure and future rewards. Poorly performing CEOs could be dismissed and well performing CEOs rewarded by appointment to a more prestigious NHS (or private sector) organization. In addition, good performance could also be recognized by the award of a national public honour granted by the Head of State. In contrast, the pay of clinical staff (including physicians) and lower level managerial staff in all NHS hospitals was (and is) set at national level (with some regional uplifts) by a public sector pay review body and was therefore essentially the same across all hospitals.

[13]Non-FT hospitals had to include a representative of the central government regulator.

[14]The remuneration committee is composed of at least three independent non-executive directors. It decides on pay of all executive directors and is to position its NHS FT relative to other NHS FTs and comparable organizations (Monitor 2014). Boards of non-FTs were more constrained in their decisions on pay of both executive and non-executive directors and had to follow regulator guidance. For CEOs see `https://improvement.nhs.uk/resources/supporting-providers-executive-hr-issues/` and for Chairs and non-executive directors see `https://improvement.nhs.uk/resources/terms-and-conditions-nhs-trust-chairs-and-non-executive-directors`. Executive and non-executive directors of FT hospitals are more highly paid than directors of non-FT hospitals.

2.2 New Public Management and NHS CEO Pay

The increased emphasis on the role and responsibilities of senior managers brought by the NPM reforms were accompanied by significant changes in CEO remuneration, both in terms of absolute levels and variance across CEOs. We examine these effects in detail using novel data we compiled from the NHS Boardroom Pay Reports published by IDS Incomes Data Services for the financial years 2000/01 to 2010/11, which we extended by hand-collecting data from hospitals' annual reports for 2011/12 to 2013/14. These reports provide data on salary, taxable benefits and total remuneration of executive directors for nearly all NHS hospitals.

Changes in Levels and Variance of CEO Pay The NPM reforms were accompanied by significant growth in CEO pay, both relative to the level at the beginning of the 2000s and relative to the level of pay for clinical staff and middle managers. This is shown in Figure 1, which plots the level of CEO pay over our sample period of 2000 to 2013, together with the mean pay of nurses, consultants (senior physicians) and middle managers. Over this period CEO pay was, not surprisingly, higher than pay of other NHS employees, but also increased faster. The increase in pay was also accompanied by an increase in the variance in CEO pay. The difference between the 10th and the 90th percentile (the shaded areas in the Figure) increased from £40,000 in 2000 to £65,000 in 2013 and at the top of the distribution CEO pay increased from £120,000 in 2000 to £175,000 in 2013.[15]

While this growth in pay did not compensate for differences in pay relative to CEOs in the UK corporate sector and hospital executives in the US,[16] these trends put the remuneration packages of NHS CEOs at the high end of the compensation distribution of the UK public sector and of UK public service providers more generally.[17] To show this we present pay data from the Quarterly Labour Force Survey, the largest household study in the UK, from April 2000 to March 2017.[18] The survey includes respondents' gross weekly pay and industry classification (SIC), occupation classification and whether

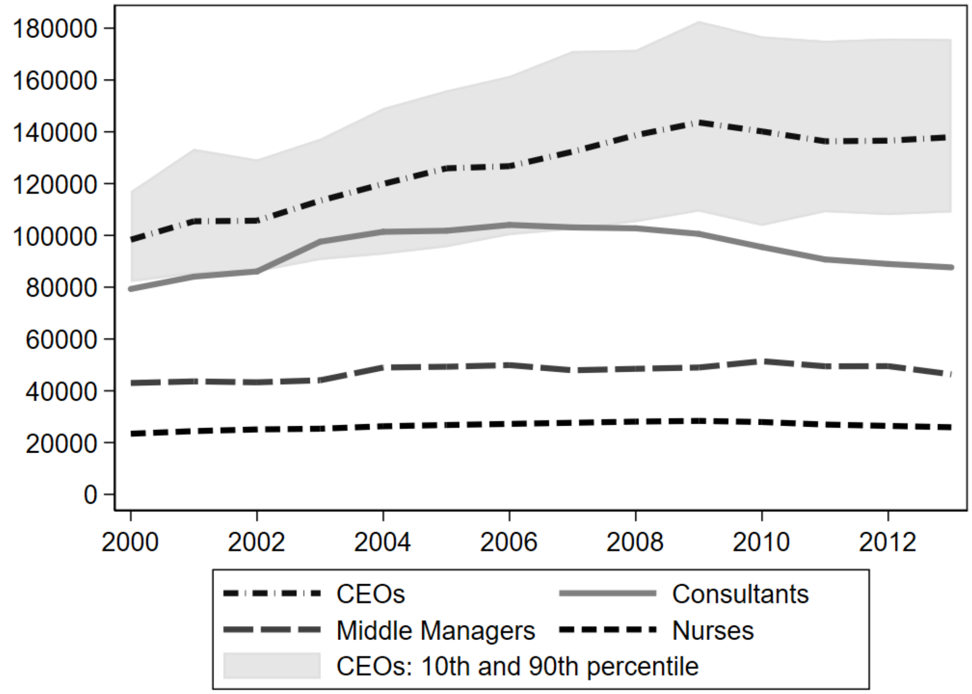
[15] NHS CEOs also receive generous pension benefits which are excluded from this analysis.

[16] Bell & Van Reenen (2016) report mean total compensation of CEOs of the top 300 UK primary-listed companies increased from £900,000 in 1999 to £1,900,000 in 2014. These remuneration packages are larger by an order of magnitude. Joynt et al. (2014) report that mean compensation of CEOs of US non-profit hospitals was $596,000 (approximately £400,000) in 2009. The majority of CEOs in their sample served at hospitals with fewer than 300 beds, well below even the 25th percentile of 446 in our sample. Similarly, figures Joynt et al. (2014) report for the highest decile of the compensation distribution, which has the largest mean number of beds (310), show mean compensation of $2,100,000 (approximately £1,400,000).

[17] The Prime Minister's salary is around £145,000. The higher pay of NHS CEOs has traditionally attracted considerable negative attention from the popular media. Articles about "NHS fat cats" receiving "six-figure salaries" or "earning more than the Prime Minister" are common (Ham et al. 2011).

[18] The Quarterly Labour Force Survey provides the official measures of employment and unemployment.

Figure 1: Annual means of pay for NHS staff by job type



Note: Adjusted for inflation using Consumer Price Index, base year = 2000

they work in the public or private sector. We focus on respondents whose occupation classification is "Directors and Chief Executives of Major Organisations".[19]

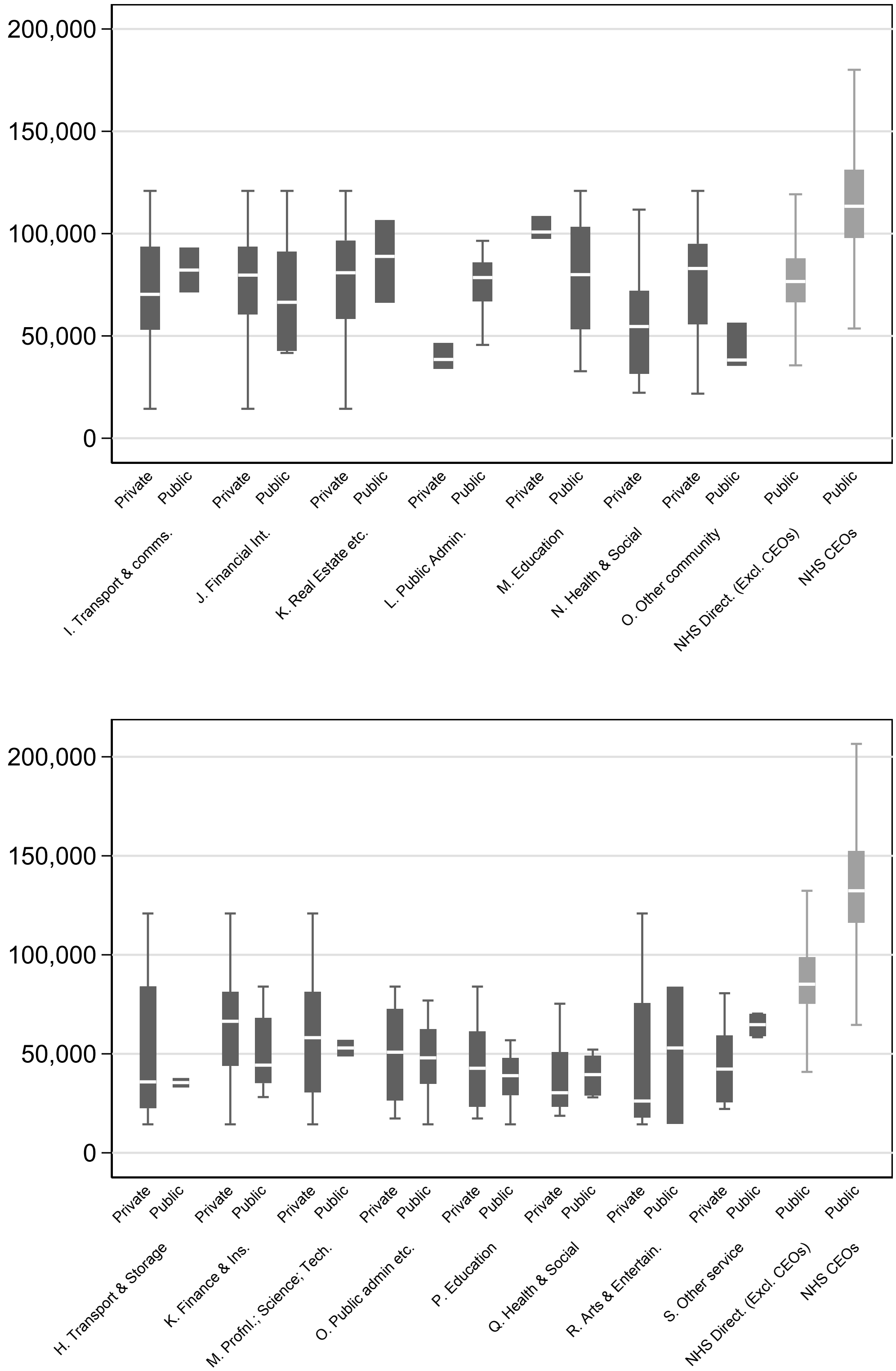
Figure 2 presents pay split by industry and public and private sector. As the industry classification was changed substantially in 2009, we present separate graphs for 2000-2008 and 2009-2017.[20] In both periods NHS hospital CEOs and non-CEO directors were well paid relative to top managers at a wide range of organizations in both the public and private sector. On average NHS CEOs received the largest pay packages. A comparison between the upper and lower graph suggests the gap between the pay of NHS directors and that of directors in other organizations grew rather than diminished over the period.[21] Thus, while NHS hospitals were unable to provide pay comparable to that offered in similar large and complex private sector companies, over the time period we consider NHS CEOs were among the most highly rewarded executives in public sector organizations.

[19] We convert weekly gross pay to annual gross pay and adjust for inflation using the consumer price index (base year = 2000). To deal with outliers and limited cell sizes for some industry-sector combinations, we windsorize the pay data at the 5% level, with the top 5% of data replaced with the 95th percentile and the bottom 5% replaced with the 5th percentile.

[20] Details of the industries in Figure 2 and changes in the classification are in Web Appendix W-1.

[21] These comparisons do not take into account pension entitlements which are also more generous in the NHS than in other public and private sector organizations.

Figure 2: Annual gross pay for "Directors and Chief Executives of Major Organisations" and basic pay for NHS CEOs and non-CEO directors in 2000-2008 (top) and 2009-2017 (bottom)



Notes: All pay values adjusted for inflation using Consumer Price Index (base year = 2000). Non-NHS pay data from Quarterly Labour Force Survey, winsorized at the 5% level.

CEO Fixed Effects in Pay While we see an increase in the variance of CEO pay over time, indicating that some CEOs may be perceived as being increasingly different to others, these statistics are not adjusted for factors that may affect individual pay, such as age or tenure. The distribution of such factors may be changing over the 14-year period in Figure 1, reflecting changes in the stock of CEOs. Nor are the statistics in Figure 1 adjusted for observable features of the hospitals. Again, the distribution of hospital characteristics may be changing over the sample period. To extract permanent differences in pay between CEOs – which allow us to see if the market perceives CEOs to be differentiated by (unobservable to us) factors such as managerial ability or leadership effectiveness – we use the Abowd et al. (1999) methodology plus post-estimation shrinkage to estimate CEO fixed effects in pay. Web Appendix W-2 discusses the data, the methods – including the requirements for identification of the CEO fixed effects in pay – and presents detailed results.[22]

Figure 3 presents the distribution of the shrunk effect estimates for a) all executive directors, b) the subset of CEOs and c) the subset of CEOs who we observe in two different CEO positions.[23] The figure shows that CEOs have higher fixed effects in pay than all other executive directors, and those CEOs that move between hospitals have slightly higher pay effects than those who do not, though the differences between these two sets of CEOs are not large. The interquartile range of the effect estimates for the 95 mover CEOs is £11,400, about one-tenth of the sample mean for CEO pay of £126,230. Thus, even after controlling for observable CEO and hospital characteristics there is dispersion in CEO pay, suggesting that employers perceive CEOs to be differentiated and pay them accordingly.[24]

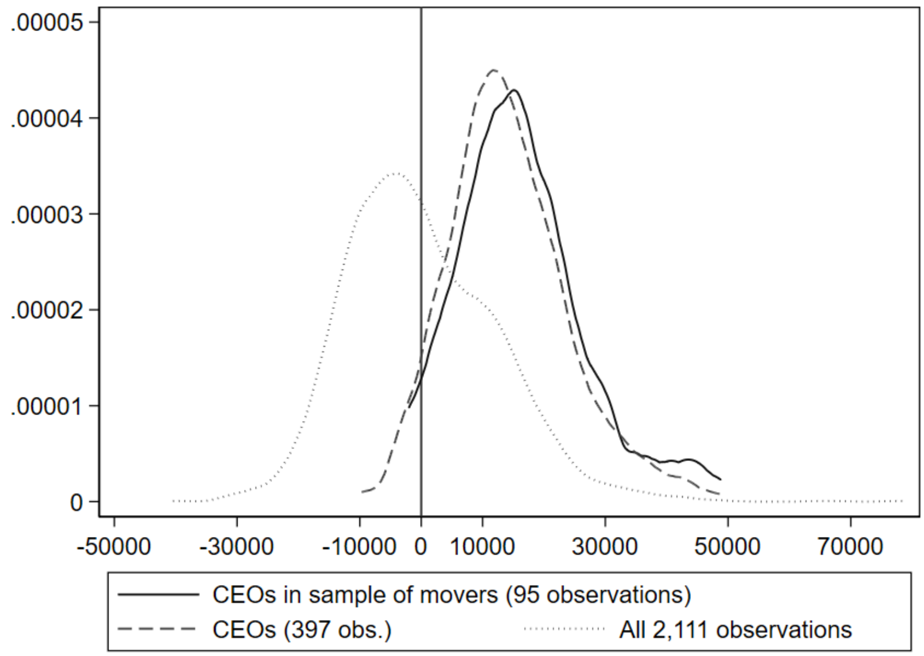
Summary In sum, the NPM reforms resulted in a significant shift in the role of hospital CEOs and the way in which they were selected and remunerated. These changes were predicated on the belief that individual top managers–when adequately selected, compensated and monitored–could deliver improvements in efficiency and quality. We investigate the extent to which this belief is supported by evidence in the rest of the paper.

[22]In brief, worker mobility is required to estimate separate hospital (firm) and CEO (worker) fixed effects. We use pay data for all executive directors and not just for CEOs to maximise the size of the set of hospitals connected by worker mobility.

[23]The effect estimates are normalised by transforming them into deviations from the mean of all the effect estimates.

[24]In contrast to analyses of pay fixed effects for CEOs in the private sector, our sample is based on a very homogeneous set of establishments – acute care hospitals – especially after controlling for hospital fixed effects and time-varying hospital characteristics such as size and case mix.

Figure 3: Kernel density plots of shrunk director effects in total pay



Notes: Kernel density plot of deviations of estimated director effects in total pay, CPI adjusted, from mean. Shrunk estimates obtained using empirical Bayes (Chandra et al. (2016)).

3 Data

Our analysis is based on information derived from various administrative data sources, which we have brought together for the first time. Variable definitions and details of the sources are in Appendix: A. In this section, we present summary statistics on our hospital production measures and detailed information on the movement of CEOs across hospitals, which we rely on to estimate the impact of CEOs on hospital production.

3.1 Hospital Production Measures

We have collated a rich set of production measures at the hospital level for the financial years 2000/01 to 2013/14.[25] The NHS made these data publicly available as part of the more general NPM reforms. Thanks to this policy of transparency, we can access a wide range of data on hospital production, including measures of inputs, throughputs (e.g. access to care metrics such as waiting times, which are important in a system where care is rationed), outputs (financial performance and measures of the quality of clinical care)

[25] Technically, these data are available at NHS hospital trust level. For readability we refer to NHS hospital trusts as hospitals. The financial year runs from 1 April to 31 March.

and staff job satisfaction. While CEO behaviour may affect all of these measures, the NPM model followed by successive governments has meant that a subset of these have been high profile targets for CEOs and hospitals in the NHS. These have not necessarily stayed the same over our sample period, as governments have changed their focus once targets for one aspect of hospital production were judged to have become less salient to voters. In the period we study, the most high profile targets were waiting times for elective care (for which there have been a series of targets in operation since 2000)[26] and financial performance (the operating surplus)[27] of hospitals. Waiting times have been a particularly important political target as the NHS rations excess demand by means of waiting lists and these are seen by the public as a measure of NHS failure when they reach long levels (Propper et al. 2010). Financial performance is also key, as all care is funded from the public purse. Both these measure were used by the NHS regulator to assess whether a hospital qualified for Foundation Trust status, which give the Board and CEO greater autonomy in terms of making large capital investments. As noted above, CEOs could be, and were, dismissed for not achieving these targets (Ballantine et al. 2008).

Other important targets reflected successive governments' general concerns over NHS expenditure and 'value for money' and were important issues for the NHS over a number of years. These targets included increasing the number of operations carried out as day cases (i.e. without overnight stay) and decreasing the average length of stay.[28] There was less focus on clinical outcomes for much of the period we study, but reducing hospital acquired infections (meticillin-resistant Staphylococcus aureus [MRSA] rates) was an important governmental concern during the period, as a number of reports drew attention to rising levels in the early 1990s.[29]

In our analyses we focus on these high profile target variables. In an extension we repeat our analyses for the non-targeted hospital production measures. Additionally, we

[26] In 2000, a policy document set out a target of a maximum waiting time for inpatient treatment of 6 months by 2005 (Department of Health 2000). In 2004, after progress had been made in reducing waiting times, a further target was set of reducing the waiting time from GP referral to hospital treatment to 18 weeks by 2008 (Department of Health 2004).

[27] Throughout our study period, hospitals were required to "ensure that its revenue is not less than sufficient, taking one financial year with another, to meet outgoings properly chargeable to revenue account" (National Health Service and Community Care Act 1990 and National Health Service Act 2006). This requirement is known as the breakeven duty. It is commonly interpreted to mean that over a three-year period hospitals' income must match their expenditure (National Audit Office 2004).

[28] In 2000, a policy document set out a target of 75% of elective surgery to be performed as day cases (Department of Health 2000). A key aim of increasing the number of day cases was to reduce length of stay to free up capacity to reduce waiting times.

[29] In 2004 the Department of Health introduced a target to reduce MRSA bloodstream infections across all NHS acute hospitals by 50% by 2008 (National Audit Office 2009). A report by the Health Foundation provides an instructive discussion of how this target was achieved (The Health Foundation 2015).

present results using a “stacked” approach, with our hospital production measures grouped into input measures, throughput measures and clinical performance measures.

Table 1 presents descriptive statistics for the hospital production measures with the targeted measures that we focus on in bold. For each variable, we show the overall mean and standard deviation as well as the mean at the beginning, in the middle and at the end of our sample period. The number of observations for the hospital production measures is determined by their availability and is reflected in the observations used in our estimations.

Table 1: Descriptive statistics for hospital production measures

				Mean of variable in		
	Obs.	Mean	St. Dev.	2000	2006	2013
Input measures:						
Doctors + nurses/beds	2,382	2.27	0.78	1.70	2.24	2.98
Senior doctors/staff (%)	2,396	8.57	2.64	6.24	7.89	10.6
Nurses/staff (%)	2,396	32.2	3.82	33.7	32.5	31.1
Technology index	2,399	0.38	0.24	0.29	0.39	0.43
Throughput measures:						
Admissions (count)	2,392	74,488	42,778	54,000	74,229	92,422
Waiting time, mean (days)	2,356	70.5	30	93.5	73.9	48.9
Day cases (%)	2,383	31.3	8.7	29.5	30.0	34.9
Length of stay, mean (days)	2,386	5.23	2.87	7.29	4.80	4.33
Cancelled operations (count)	2,332	373	290	401	301	404
Clinical performance measures:						
AMI deaths (%)	1,757	7.25	2.87	9.18	6.75	5.44 (2012)
Stroke deaths (%)	1,965	22.7	5.29	27.1	23.0	17.5 (2012)
FPF deaths (%)	1,920	8.94	2.58	9.16	9.20	7.21 (2012)
Readmissions (%)	2,070	9.80	1.66	8.34	10.2	11.2 (2011)
MRSA rate (per 10,000 bed days)	2,055	10.2	8.36	15.7 (2001)	16.6	2.4
Surplus (£000)	2,396	-1,965	15,101	259	-796	-4,975
Staff job satisfaction (1=dissatisfied, 5=satisfied)	1,838	3.47	0.10	3.47 (2003)	3.39	3.61

AMI deaths are deaths within 30 days of emergency admission for acute myocardial infarction. Stroke deaths are deaths within 30 days of emergency admission for stroke. FPF deaths are deaths within 30 days of emergency admission for fractured proximal femur. Means of total pay in 2000 are based on only 1 observation each for Finance Director, Chief Operating Officer, Nursing Director, HR Director and Other. Definitions and sources of all variables in Table A-1 in Appendix: A.

On average, the hospitals in the English NHS are large. The sample average is 75,000 admissions per year, and these have steadily grown due to increased demand and consolidation in the sector (see Figure 4b below). The skill mix also changed. The ratio of the most skilled staff to number of beds–a measure of the labour-to-capital ratio–and the ratio of senior doctors to staff–a measure of the labour skills ratio–grew by around 70% over our sample period. Some clinical outcomes improved: for example, acute myocardial infarction (AMI) mortality rates declined from 9.18% in 2000 to 5.44% in 2012 but other clinical outcomes display more stagnant patterns. Deaths after admission for fractured proximal femur (FPF) did not reduce much and readmissions increased. In line with the literature, there is also evidence of large variations in performance across hospitals, for inputs, throughputs and outcomes.[30]

In terms of performance with respect to key targets, financial performance plummeted over our sample period, moving from an average surplus of £259,000 in 2000 to an average deficit of £4,975,000 in 2013. There were sector wide falls in waiting times (the time between decision to admit and actual admission) from 93.5 days in 2000 to 48.9 days in 2013. The number of day cases rose from just under 30% to just under 35% of admissions, length of stay declined from 7.29 days to 4.33 days, and the meticillin-resistant staphylococcus aureus (MRSA) rates dropped from 15.7% in 2001 to 2.4% in 2013.

Table W-1 in Web Appendix W-2 provides descriptive statistics on time-varying hospital characteristics that we use as controls in the hospital production function. They include governance measures, capital measures and case-mix measures.

3.2 Identifying Movements of CEOs across Hospitals

From the CEO pay data that we use to estimate CEO fixed effects in pay (see Section 2.2 and Web Appendix W-2) we identify movements of CEOs across hospitals, which are necessary to estimate CEO fixed effects in performance.[31] NHS hospitals are characterized by very high CEO turnover rates. Figure 4a shows that every year between 12 to 25% of hospitals in our sample have a turnover event.[32] Figure 4b shows the decrease in the number of hospitals on which the turnover rates in Figure 4a are based. The high turnover is partly due to consolidation which occurred between 1997 and 2003.[33]

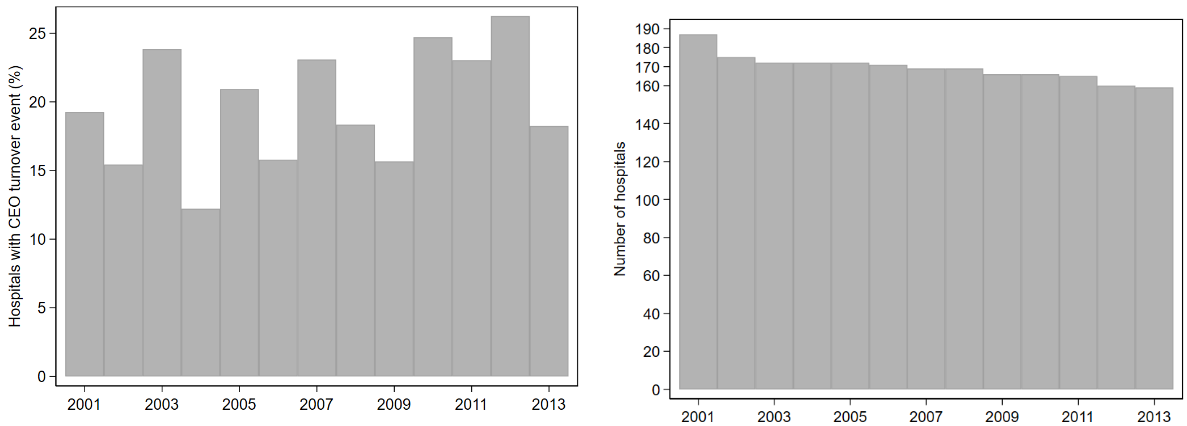
[30] Chandra et al. (2016) show large and persistent performance differentials across U.S. hospitals.

[31] We complemented the published data with extensive manual searches to check the personal identifiers for all executive directors in the data.

[32] As our data start in 2000 we report turnover events only from 2001 onward. Some hospitals experienced more than one CEO turnover event in a financial year. The turnover events in Figure 4a include all CEO turnover events in our sample, i.e. not only the turnover events for the CEOs observed in at least two hospitals.

[33] Over half the NHS acute hospitals in 1997 had been involved in some kind of merger or reconfiguration with other NHS hospitals by the end of 2003 (Gaynor et al. 2012). All consolidation was within the NHS.

Figure 4: CEO turnover and hospital consolidation



(a) Annual proportion of hospitals with CEO turnover event

(b) Annual number of hospitals

Figure 5a shows sample entry and exit of hospitals between 2000 and 2013. At the beginning of our sample period in 2000 and 2001 over 10% of hospitals exit the sample. Hospitals only enter the sample because of mergers.[34] Figure 5b shows sample entry and exit of CEOs over the same period. Hospital consolidations mechanically led to increased CEO turnover, since at the very least only one of the CEOs of the formerly separate hospitals continued in post, and frequently a new CEO was appointed to lead the consolidated hospital. However, CEO turnover appears to have increased even in absence of merger events. CEO sample entry and exit is on average considerably higher than hospital sample entry and exit, at around 14% for the whole period. CEO entry and exits in our sample are highest during the period of consolidation in the early 2000s and then fall and remain relatively stable after 2004, but are still both over 10% at the end of the period.[35]

Figure 6a shows CEO turnover per hospital for the subset of hospitals observed for at least 11 years.[36] These hospitals have on average 3.5 CEOs during the sample period. Only a minority of hospitals have the same CEO throughout and the majority have two to five CEOs over the sample period of 11 to 14 years. Hospitals with more CEOs over the sample period tend to be in certain regions of England but few other time-invariant

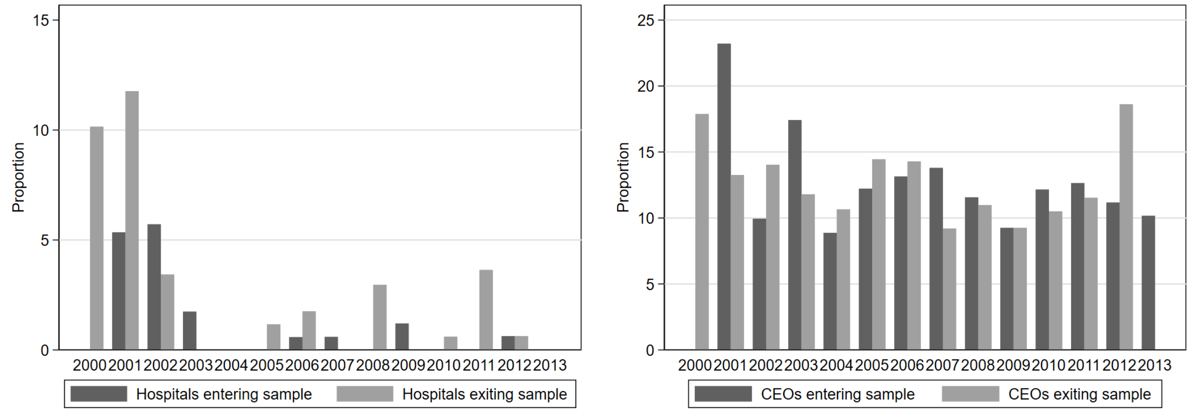
Consolidation meant that NHS hospitals grew in size, providing services from a number of sites in the same local area. There are no NHS hospital chains.

[34] Following a merger, the merged entity was generally given a new NHS code. We treat each new code as a separate hospital in our analysis. Mergers in which a much larger hospital absorbed a smaller hospital and kept its name and NHS code are captured by the acquisition dummy variable.

[35] The rise in exits in 2012 reflects the uptick in hospital consolidation in 2011.

[36] As our data set excludes CEOs that served for a part year only, the number of CEOs per hospital is a lower bound.

Figure 5: Sample entry and exit of hospitals and CEOs



(a) Proportion of hospitals entering and exiting sample in each year

(b) Proportion of CEOs entering and exiting sample in each year

characteristics are associated with the number of CEOs a hospital has.[37]

Figure 6b is critical for our analysis, in that it shows that a sizeable number of CEOs are observed as CEO in two or more NHS hospitals. This mover CEO sample is necessary to estimate CEO fixed effects separately from fixed or time-varying hospital characteristics. To examine whether mover CEOs are different from non-movers, we regress fixed characteristics of the CEO against a dummy variable indicating whether a CEO is one of the 95 CEOs included in our two-step approach estimation sample (which requires that they are a mover and that they are in each hospital for at least two years). The characteristics examined are gender, whether the CEO has a clinical qualification and whether they have a postgraduate management qualification. Mover CEOs are slightly more likely to have a postgraduate management qualification but do not differ in terms of gender or clinical background from non-movers.[38] Tenure for mover CEOs is short, with a median of 4 and a mean of 4.5 years, though longer than the average tenure of 3.8 years for the complete set of CEOs.

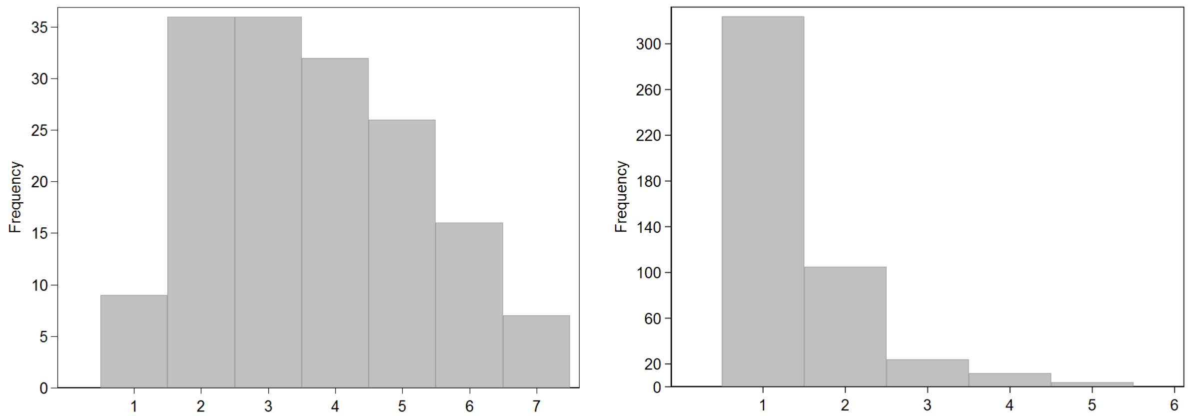
4 Methods

The most commonly used method to estimate the impact of CEOs on organizational performance is the fixed effects approach pioneered by Bertrand & Schoar (2003). The essence of this approach is to examine mover CEOs and estimate CEO fixed effects in a regression of firm-level outputs on time-varying firm characteristics and firm fixed effects.

[37] Details are in Table W-7 in Web Appendix W-3.

[38] For details see Table W-8 in Web Appendix W-3.

Figure 6: Number of CEOs per hospital and number of hospital spells per CEO



(a) CEOs per hospital (N = 162) (b) Hospital spells per CEO (N = 469)

Notes: CEOs per hospital is the number of CEOs observed per hospital for hospitals observed for at least 11 years. Hospital spells per CEO is the number of CEO spells at different hospitals for executive directors that are observed in a CEO position at least once.

If the set of CEO fixed effects is statistically significantly different from zero according to an F-test, then it is inferred that there are CEO fixed effects.[39] However, a possible shortcoming of the Bertrand and Schoar (2003) approach is that a large residual in one hospital might result in a mean residual that is statistically significantly different from zero as a consequence of a period-hospital-specific effect, rather than a persistent CEO effect (see also Fee et al. (2013)). To test whether this was an issue in our data, we implemented the Bertrand & Schoar (2003) fixed effects method and then undertook a placebo test in which we randomly assigned CEOs to hospitals and tested the joint significance of the fixed effects for each of the 100 replications of the random assignment. Our analyses, (see Web Appendix W-4) showed that while the CEO fixed effects for the real CEO-hospital matches were statistically significant, the means of the F-statistics across the 100 placebo samples were almost identical to those for the real data. We therefore did not use this approach, but resorted to two complementary methods to identify the CEO fixed effects, which we describe below.

[39] The left-hand side variable in Bertrand & Schoar (2003) is a firm level production measure in year t rather than the pay of individual i at the firm in year t. The estimation challenges thus differ from those discussed in the literature which focuses on worker and firm level effects using the Abowd et al. (1999) methodology. For example, while in Abowd et al. (1999) only the firm fixed effects of firms connected by worker mobility are identified, in Bertrand & Schoar (2003) all firm fixed effects are identified. On the other hand, identification of CEO fixed effects (i.e. worker fixed effects) is not straightforward in Bertrand & Schoar (2003). Bertrand & Schoar (2003) deal with this issue by estimating CEO fixed effects only for CEOs observed in two firms for at least three years each.

4.1 Parametric Approach: Two-step Procedure

The first approach is a parametric two-step procedure proposed by Bertrand & Schoar (2003). The first step involves estimating regressions of the following form:

$$y_{jt} = \boldsymbol{X'}_{jt}\boldsymbol{\beta} + \lambda_t + \psi_j + \varepsilon_{jt} \tag{1}$$

The dependent variable, y_{jt}, is one of the set of target measures discussed in Section 3 above for hospital j in financial year t. $\boldsymbol{X}_{jt}$ is a vector of the following time-varying observable hospital characteristics: foundation trust status, year of merger, years since merger, beds, technology index and case mix measures (patients aged 0 to 14, patients aged 60 to 74, patients aged 75+, male patients). A full set of financial year effects, λ_t, provides non-parametric control for trends in the hospital production measure that are national in scope while a full set of hospital effects, ψ_j, controls for non-time varying unobserved differences between hospitals.

For each CEO that we observe in two hospitals for at least two years duration in each, two residual means are calculated from Equation 1.[40] These are the mean of the residuals for the financial years $t_1^{i,A}$ to $t_n^{i,A}$ when CEO i is observed in hospital A and the mean of the residuals for the financial years $t_1^{i,B}$ to $t_n^{i,B}$ when CEO i is observed in hospital B. In the second step, the mean for CEO i's spell in hospital B is regressed on the mean for CEO i's spell in hospital A:

$$\frac{1}{n^{i,B}} \sum_{t=t_1^{i,B}}^{t_n^{i,B}} e_{Bt} = \delta_1 + \delta_2 \frac{1}{n^{i,A}} \sum_{t=t_1^{i,A}}^{t_n^{i,A}} e_{At} + \varepsilon_i \tag{2}$$

The coefficient of interest is δ_2. A positive value indicates that individual CEOs' deviations from the expected level of the dependent variable y_{jt} are similar across two different hospitals, which would be supportive of a persistent CEO effect.[41]

[40]For CEOs observed in three or four hospitals for at least two years in each, we use only the two most recent spells to be comparable with other CEOs that we only observe in two hospitals. If CEOs are observed in a hospital for two years but served for only part of each of these two years, we omit these observations. We omit CEO spells if the CEO was observed in a hospital for the exact same time period as we observe the hospital for since the hospital effect cannot be distinguished from the impact of the CEO.

[41]To test the validity of the two-step procedure, we carried out a placebo experiment similar to the one described above for the Bertrand & Schoar (2003) fixed effects methods. In particular, we repeatedly randomly assigned CEOs to hospitals and re-estimated Equations 1 and 2. The results indicate that the two-step procedure is valid. More details are in Web Appendix W-4.

Testing for Pre-assignment Trends We run the placebo regression proposed by Bertrand & Schoar (2003) to test for pre-assignment trends. Instead of using the mean of the residuals at hospital B during the time the CEO was observed there, $\frac{1}{n^{i,B}} \sum_{t=t_1^{i,B}}^{t_n^{i,B}} e_{Bt}$, we use the mean of the the residuals e_{jt} at hospital B during the three financial years before the CEO arrived at the hospital B.

The idea is that a positive δ_2 in Equation 2 might wrongly suggest that individual CEOs have an impact on hospital production. Instead, hospital boards might recruit CEOs that have experience of an environment similar to the one the hospital is currently operating in. For example, a CEO who has overseen a shift to more day case procedures at hospital A might be recruited to oversee a similar shift to more day case procedures at hospital B. In this case, deviations from the expected proportion of day case procedures at hospital B might precede the new CEO's arrival. A positive association between CEO i's deviations from the expected proportion of day cases at hospital A and hospital B's deviation from the expected proportion during the three years before CEO i arrived there is suggestive of selection of the CEO rather than of the CEO imposing their style.

4.2 Non-Parametric Approach

The two-way procedure relies heavily on the statistical model of hospital production in Equation 1. It also relies on CEOs having an impact on the same targeted measure across two hospitals. Therefore, we use a non-parametric approach to sidestep both problems. Our approach resembles a difference-in-difference estimator combined with matching. We compare the changes in hospital production following a CEO turnover event to changes at matched hospitals without a CEO turnover event. If there is any impact of CEOs on targeted production measures, we expect to see different changes after a CEO turnover event compared to changes at otherwise similar hospitals with no CEO turnover event.

This analysis uses all CEO turnover events in our sample. We identify hospitals that had a CEO turnover event that resulted in subsequent stable leadership of at least two years. From this set of observations we select those CEO turnover events that were preceded by two years of stable leadership. The treated observations are hospitals with a CEO turnover event in t and the new CEO staying on in $t+1$ and no CEO turnover in $t-1$ and $t-2$. This selection criterion excludes those NHS hospitals characterized by frequent CEO turnovers within a short time period–most likely hospitals in a crisis–for which it is hard to find a suitable control group. We match each treated observation to a control hospital with no CEO turnover event over the period $t-2$ to $t+1$.

We match, with replacement, treated hospitals to control hospitals exactly on year, teaching status, specialist status and foundation trust status in $t-1$. Matching exactly on year implies that we compare, for example, the difference in waiting times between 2006 and 2008 for a hospital with a CEO turnover event in 2007 to the difference in waiting times between 2006 and 2008 for a hospital with no CEO turnover event in 2007. Thus our results will not be confounded by period effects (Gaynor et al. 2012).[42] This tends to result in more than one match for each treated hospital. Therefore we use nearest neighbor matching on beds in $t-1$ to choose one control hospital from among the exactly matched hospitals. Where nearest neighbor matching on beds results in ties, we choose from among the (usually two) hospitals with the same absolute difference in number of beds the nearest neighbor in terms of the technology index in $t-1$.

We then compare the difference in the targeted production measures between the year before the CEO turnover event and the end of the two-year period, i.e. between $t-1$ and $t+1$, to the equivalent difference in the matched hospital. We report the mean of the change, $\frac{1}{n}\sum_{j=1}^{n}\left(y_{j(t+1)}-y_{j(t-1)}\right)$, and its standard error for both the treated and control hospitals. We present the difference between the two means as well as the standard error and p-value from a two-sample t-test with equal variance. To enable comparison of effect sizes between the different measures, we standardise all outcome variables to have a mean of zero and a standard deviation of one.

5 Results

5.1 Parametric estimates

Main results Table 2, Column 1, presents our estimates of Equation 2 for key target variables and staff job satisfaction. Results are similar when we extend the analysis to other non-targeted variables (see Web Appendix W-6). A positive coefficient indicates that a positive deviation from the expected level of a hospital production measure during a CEO's spell at the first hospital is associated with a positive deviation from the expected level of that production measure during the CEO's spell at the second hospital. A statistically significant association would suggest that these deviations can be attributed to the CEO and not to period-hospital-specific effects.

In fact, with one exception, the coefficients are all very small or even negative, indicating there is no association between the residuals at the first and second hospital.

[42]Matching on teaching status implies matching treated major teaching hospitals to control major teaching hospitals and treated minor teaching hospitals to control minor teaching hospitals. For specialist status we match only on the broad definition of specialist hospital. Teaching and specialist status are fixed characteristics, foundation trust status is time-varying.

The one exception is the result for day cases, with a coefficient of 0.18 indicating some portability. When we test for pre-assignment trends in Column 2 of Table 2, however, day cases have a similar sized positive association, but in this case between the mean of the residuals in the second hospital where each CEO is assigned during the three years *before* the CEO arrived and those at the CEO's first hospital. Thus, the positive correlation in Column 1 could be due to selection rather than CEOs imposing their style. Apart from day cases, none of the coefficients in Column 2 is statistically significant, so there is only weak evidence of selection for these targets.

Table 2: Association of (1) means of residuals for CEO spells at first and second hospital and (2) mean of residuals for CEO spell at first hospital and pre-assignment residual at second hospital

	1			2		
	CEO spells at 1st hospital and 2nd hospital			CEO spell at 1st hospital and pre-assignment trend at 2nd hospital		
	Coefficient (std. error)	R^2	Obs.	Coefficient (std. error)	R^2	Obs.
Surplus	-0.05 (0.30)	0	95	0.16 (0.22)	0.01	92
Waiting time	-0.01 (0.08)	0	93	0.01 (0.08)	0	90
Day cases	0.18* (0.09)	0.04	95	0.19** (0.10)	0.04	92
Length of stay	0.05 (0.06)	0.01	94	-0.04 (0.09)	0	91
MRSA rate	0.10 (0.10)	0.01	80	-0.05 (0.12)	0	78
Staff job satisfaction	-0.07 (0.11)	0	73	-0.11 (0.17)	0.01	73

Results in Column 1 are from regressions of mean of residuals for CEO spell in second hospital on mean of residuals for CEO spell in first hospital. Results in Column 2 are from regressions of mean of residuals in second hospital during the three years before CEO was appointed on mean of residuals for CEO spell in first hospital. The residuals are from a regression of the standardised measures on hospital characteristics, financial year effects and hospital effects. *Significant at 10%, **significant at 5%, ***significant at 1%

Within spells correlations One possible concern is that the lack of portability across hospitals for the key target variables is that the residuals of the yearly hospital-level aggregate may be too noisy to meaningfully detect the effects of managerial actions. To assess this possibility, we examined the patterns of correlations between the residuals of the target variables *within* spells. This exercise consists of checking that "expected" correlations among key targets are found in the data within each CEO spell. For example, longer than expected length of stay should in principle be associated with longer than expected waiting times. Similarly, if more than expected treatments are being carried out as day cases, patients staying overnight are likely to require more complex care, and therefore have longer than expected length of stay.

To implement this check, we use the $\frac{1}{n^{i,A}}\sum_{t=t_1^{i,A}}^{t_n^{i,A}} e_{At}$ and the $\frac{1}{n^{i,B}}\sum_{t=t_1^{i,B}}^{t_n^{i,B}} e_{Bt}$ from Equation 2 for each of our key target measures. Therefore, the number of observations is up to $95 \times 2 = 190$. Four out of the 15 correlation coefficients are statistically significant. They are 0.27 (p = 0.00, n = 184) for (waiting time, length of stay); 0.29 (p = 0.00, n = 188) for (day cases, length of stay); 0.25 (p = 0.00, n = 158) for (waiting time, MRSA rate) and 0.16 (p = 0.05, n = 146) for (day cases, staff job satisfaction). Thus, we find the expected correlations between throughput measures, suggesting that the residuals are capturing salient aspects of hospital production within each CEO's spells rather than being pure noise. However, we do not find any patterns beyond throughput measures, suggesting a lack of a broader managerial "style" across different types of target variables.

Comparison with pay For comparison, we repeated the two-step approach using the pay data. We derived the mean residuals from a regression of executive director pay on hospital characteristics, tenure, financial year effects and hospital effects. The regression coefficient is 1.33 (s.e. = 0.17), providing strong evidence of portability of pay: a CEO who is paid more than expected at their first hospital is also paid more than expected in the second.[43]

We also examined the correlations between the spell-specific residual means for the key target measures and the shrunk CEO pay effects. None of the six correlation coefficients is statistically significant and they range in size from -0.09 (p = 0.22, n = 190) for (surplus, pay effect) to 0.07 (p = 0.37, n = 188) for (length of stay, pay effect). Thus there seems to be little relationship between performance of the CEO on these observable and targeted measures and their pay.

[43] The larger variance of the mean residuals at the second hospital distorts the coefficient, making it larger than 1. The equivalent correlation coefficient is 0.63. The correlation coefficients for the production measures are -0.02 for surplus; -0.01 for waiting time; 0.19 for day cases; 0.08 for length of stay; 0.11 for MRSA rate and -0.07 for staff job satisfaction.

5.2 Non-parametric estimates

We now turn to the non-parametric approach and examine whether there is any change in the targeted production measures following a CEO turnover event. This approach uses a larger set of CEOs than the analyses above that uses only mover CEOs, as it includes as potential treated observations all CEO turnover events. Table 3 presents the results, which suggest few changes in the target measures after a new CEO is in post. Length of stay does fall, though the effect size is small, with the estimate showing a drop of 5% of a standard deviation after a turnover event. However, the main impact of a CEO turnover event is to decrease staff satisfaction, which falls by one-fifth of a standard deviation.

Table 3: Changes in hospital production measures following a CEO turnover event compared to matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Surplus	Treated	205	0.096 (0.072)		
	Controls	205	0.139 (0.055)	-0.044 (0.091)	0.63
Waiting time	Treated	200	-0.328 (0.043)		
	Controls	200	-0.291 (0.037)	-0.037 (0.056)	0.51
Day cases	Treated	202	0.109 (0.030)		
	Controls	202	0.084 (0.036)	0.024 (0.047)	0.60
Length of stay	Treated	205	-0.167 (0.023)		
	Controls	205	-0.123 (0.015)	-0.046 (0.028)	0.10
MRSA rate	Treated	197	-0.262 (0.048)		
	Controls	197	-0.275 (0.050)	0.013 (0.069)	0.85
Staff job satisfaction	Treated	163	0.124 (0.076)		
	Controls	163	0.312 (0.071)	-0.189 (0.104)	0.07

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. Controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t-1$ and $t-2$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status, followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

These (null) results persist in a number of robustness tests of the non-parametric analysis. Table W-14 in Web Appendix W-5 repeats the analysis using 1:3 matching and gives similar results. Table B-1 in Appendix: B assesses the balance of the matched samples. We find little difference between the treated and the control samples, with the exception of the number of beds, which is slightly larger for the treated sample. Table W-15 in Web Appendix W-5 checks the common trend assumption by examining changes in hospital production in the two-year period preceding the CEO turnover event. Changes observed in treated hospitals are generally similar to changes in control hospitals. Table W-16 uses only those CEO spells included in the parametric two-step approach. The results are similar to those here: there are no changes in the hospital production measures following a CEO turnover event with the exception of staff job satisfaction. Table W-17 examines changes over the three (rather than two) years following the CEO turnover event. Again, there is a negative impact on staff satisfaction, though due to the smaller sample size this effect is less precisely estimated.[44]

5.3 Extensions

The New Public Management model embodies a focus on key targets for CEOs. We find little evidence of portability of performance with respect to these six measures. However, it may be that NHS CEOs do have a management style, but one that focuses on other aspects of hospital production beyond these that constitute official government targets. We use our rich production data to examine this question. We do this in three ways.

First, we repeat both the parametric and the non-parametric analyses for all hospital production measures in our data (i.e. not just those that were objects of targets). Results, presented in Web Appendix W-6, show a lack of persistent effect across all the measures we considered.

Second, we exploit our rich production data by "stacking" the four input measures, the five throughput measures and the five clinical performance measures and repeating the parametric and non-parametric analyses for each of these three sets of stacked production measures. Stacking allows for correlations between the production measures in a set and reduces potential multiple comparison issues. It also sidesteps the issue of missing data for some of the production measures and therefore increases the sample size, giving us more statistical power to detect CEO effects. Details and results are in Web Appendix W-7. Even in these case, we fail to detect persistent CEO impacts on hospital production.

Third, we extend the parametric analysis by estimating Equation 2 across outcome measures. It might be the case that a CEO focuses on one target measure at one hospital and another at a second, which is not taken into account in the analysis of individual target outcomes across spells. To investigate this possibility we replaced the $\frac{1}{n^{i,B}} \sum_{t=t_1^{i,B}}^{t_n^{i,B}} e_{Bt}$

[44]The sample size is smaller as we only include CEO turnover with the new CEO still in post in $t+2$, to calculate changes between the year before the CEO turnover event $(t-1)$ and three years later $(t+2)$.

in Equation 2 with the spell-specific residual means for one of the five other hospital production measures. Thus, instead of estimating six relationships we estimate $6 \times 6 = 36$ relationships. The results are in Web Appendix W-8. We find that a shorter than expected length of stay at a CEO's first hospital is associated with more day cases at the second hospital and shorter than expected waiting times at the first hospital with lower than expected MRSA rates at the second hospital. These results suggest that CEOs may have a style that is portable across hospitals but its impact, if it exists, is limited to a small number of targets (and possibly in the case of day surgery and waiting times, easier ones to achieve).

The bottom line from this battery of analyses and robustness tests is that there is little evidence that NHS CEOs have a management style related to easily observed aspects of hospital production.

6 Potential Reasons for Lack of Persistent CEO Effects

The New Public Management model is predicated on the assumption that the managers of public service delivery organisations can affect organisational performance along observable dimensions that are established by politicians or regulators as key targets. However, we have found very little evidence supporting this assumption in the context of NHS hospitals. In this section we explore features of the labour market for NHS CEOs that may explain this finding.[45] We investigate (1) whether the lack of persistent CEO effects may be driven by the patterns of CEOs' movements across NHS hospitals, or exits of capable CEOs from the NHS to the private sector; (2) whether there are match effects, such that certain CEOs only have an effect when managing certain types of hospitals; and (3) whether the average length of tenure of NHS CEOs is too short to make a difference.

6.1 Negative Selection of CEOs

We start by studying whether CEOs with above normal performance in their first spell are systematically assigned to manage difficult hospitals in their second spell, or whether they are more likely to exit the NHS system all together.[46] If "better" CEOs have a higher probability of being hired by more problematic hospitals, it could make it harder

[45] We do not examine pay, as we have already shown in Section 2 that NHS CEOs are well paid relative to other senior managers of organisations that deliver public services in the public sector.

[46] In the literature examining worker and firm fixed effects in pay using the Abowd et al. (1999) methodology endogenous assignment of workers to firms would threaten identification of the parameters of interest (Card et al. 2013). The goal of our investigations is not to demonstrate the validity of our identification strategy but to examine potential reasons for the lack of CEO effects in hospital production.

for these CEOs to replicate a positive impact across hospitals. Similarly, the NHS system may be systematically retaining CEOs that have a lower ability of leaving a mark on their organization.

Endogenous Assignment A CEO who experiences a positive shock in one hospital may subsequently be hired by a hospital at which it is difficult to bring about positive changes. In this case, an above expected performance would be followed by a below average performance, resulting in no apparent CEO fixed effect and a negative correlation between performance in the first hospital and performance in the second. Other indications of this type of assignment would be if CEOs who have higher variability in their performance are at some point in a hospital that is hard to manage.

To test these ideas, we begin by identifying hospitals which may be hard-to-manage because they are experiencing problems. We consider four definitions of "problematic". These are (i) having received a poor rating from the government regulator of hospitals before the CEO arrived at the hospital, (ii) having poor financial performance, defined as being in the lowest 25^{th} percentile of surplus in the year before the CEO arrived, (iii) being a "new" hospital that was created through a merger at some point during our sample period, and (iv) holding a contract for a large capital investment at some point during the CEO's tenure.[47]

We first examine whether CEOs who were good performers at their first hospital are subsequently hired by a problematic hospital. For surplus, day cases and staff job satisfaction we define as good performers those CEOs whose mean of the residuals e_{jt} from Equation 1 for the financial years $t_1^{i,A}$ to $t_n^{i,A}$ when CEO i is observed in hospital A is at or above the 75^{th} percentile. For waiting times, length of stay and MRSA rate we define as good performance being at or below the 25^{th} percentile. We estimate linear probability models of the impact of good performance in the first hospital on moving to a problematic hospital. Table 4 presents the results. There are 4 definitions of problematic and 6 hospital production measures, generating 24 coefficient estimates. Only two of these estimates are statistically significant and both are negative, suggesting that good performers are *less* likely to move to a problematic hospital.[48]

[47]NHS hospitals have to borrow for large capital investments from the private market. Borrowing is through vehicles with long-term fixed interest rates and long payback periods, known as private finance initiative (PFI) contracts. Hospitals with these contracts have often struggled to meet financial performance requirements once the payback period has begun. Regulator ratings were not issued each year. Details are in Appendix: A.

[48]In complementary analysis, we examined whether good performers move to more prestigious hospitals. We define as "prestigious" teaching hospitals, hospitals which have foundation trust status, and the biggest hospitals (defined by number of beds). Our definitions of good performance in the first hospital are the same as in Table 4. Table W-37 in Web Appendix W-9 shows that there is some indication that good

Table 4: Linear probability models of the impact of good performance in first hospital on moving to a "problematic" hospital

	Hospital commission rating poor in year before CEO arrived			Hospital with surplus below 25th percentile in year before CEO arrived			'New' hospital created through merger during sample period			Hospital with PFI contract at some point during CEO's tenure		
	Good perf.	Const.	N	Good perf.	Const.	N	Good perf.	Const.	N	Good perf.	Const.	N
Surplus	0.14	0.39	71	0.02	0.46	91	-0.07	0.33	95	0.08	0.40	95
	(0.14)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Waiting time	-0.11	0.44	70	-0.05	0.46	89	-0.20*	0.37	93	0.01	0.43	93
	(0.14)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Day cases	0.07	0.40	71	0.08	0.44	91	-0.09	0.34	95	-0.06	0.44	95
	(0.13)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Length of stay	-0.08	0.43	70	-0.08	0.49	90	0.04	0.31	94	0.07	0.41	94
	(0.14)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.41)	
MRSA rate	-0.26*	0.44	58	-0.08	0.50	77	-0.03	0.33	80	0.03	0.42	80
	(0.14)	(0.07)		(0.13)	(0.07)		(0.12)	(0.06)		(0.13)	(0.06)	
Staff job satisfaction	0.13	0.33	52	0.13	0.40	72	0.04	0.28	73	0.21	0.37	73
	(0.16)	(0.08)		(0.13)	(0.07)		(0.12)	(0.06)		(0.13)	(0.07)	

Each entry in this table refers to a separate regression of an indicator of a CEO moving to a "problematic" hospital on an indicator of good performance at the CEO's first hospital. "Problematic" hospital is defined as either poor hospital commission rating, surplus below 25^{th} percentile, hospital created through merger or hospital with PFI contract. The 25^{th} percentile of surplus is calculated separately for each financial year to ensure the categorisation is net of year effects. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table 5: Impact of ever being observed at a "problematic" hospital on variability in CEO performance as measured by the absolute difference in the mean residuals for the CEO spells at each of their two hospitals for each production measure

	Hospital commission rating poor in year before CEO arrived	Hospital with surplus below 25th percentile in year before CEO arrived	'New' hospital created through merger during sample period	Hospital with PFI contract at some point during CEO's tenure	Mean (st. dev.) [obs.] of dependent variable (Absolute difference in mean residuals at both hospitals)
Surplus	0.18 (0.16) [89]	-0.07 (0.17) [94]	-0.10 (0.16) [95]	0.18 (0.16) [95]	0.46 (0.76) [95]
Waiting time	0.13** (0.06) [88]	0.09 (0.07) [92]	0.04 (0.06) [93]	-0.03 (0.06) [93]	0.36 (0.29) [93]
Day cases	0.01 (0.05) [89]	0.03 (0.05) [94]	-0.00 (0.05) [95]	0.05 (0.04) [95]	0.27 (0.22) [95]
Length of stay	-0.00 (0.03) [88]	0.00 (0.03) [93]	0.03 (0.03) [94]	0.01 (0.03) [94]	0.13 (0.13) [94]
MRSA rate	-0.08 (0.07) [76]	0.03 (0.08) [79]	-0.05 (0.07) [80]	0.03 (0.07) [80]	0.33 (0.31) [80]
Staff job satisfaction	-0.03 (0.06) [70]	-0.02 (0.06) [73]	-0.03 (0.06) [73]	-0.15*** (0.06) [73]	0.31 (0.24) [73]

Each entry in this table refers to a separate regression of a performance variability measure on a dummy variable indicating that the CEO has ever been observed at a "problematic" hospital defined as either poor hospital commission rating, surplus below 25^{th} percentile, hospital created through merger or hospital with PFI contract. The 25^{th} percentile of surplus is calculated separately for each financial year to ensure the categorisation is net of year effects. Standard errors in (parentheses) and number of observations in [brackets]. *Significant at 10%, **significant at 5%, ***significant at 1%

Second, we examine whether variability in CEO performance is associated with being at some point matched with a problematic hospital. We generate a measure of the variability in CEO performance as follows. We use the mean of the residuals e_{jt} from Equation 1 for the financial years $t_1^{i,A}$ to $t_n^{i,A}$ when CEO i is observed in hospital A and the mean of the residuals for the financial years $t_1^{i,B}$ to $t_n^{i,B}$ when CEO i is observed in hospital B. To measure variability in CEO performance, we calculate the absolute value of the difference in these two means. We generate this variability measure for all production measures and examine whether the variability measure is larger for CEOs who at some point in their career work in problematic hospitals. To do this, for each of the four definitions of "problematic", we regress each of the 6 variability measures against a dummy variable indicating that the CEO was ever observed in a problematic hospital. Table 5 presents the results. 2 out of the 24 coefficients are statistically significantly different from zero, which is what we expect at a 10% significance level. However, only one of these coefficients (for waiting times) is positive. A positive coefficient would suggest that being at a "problematic" hospital is associated with higher variability in CEO performance. The negative coefficient (for staff job satisfaction) suggests that CEOs who are at some point at a more problematic hospital actually have *lower* variability in their performance across hospitals.[49]

In essence, we find little evidence that "good" CEOs (as inferred from their performance) end up managing more problematic hospitals. If anything, good performers are *less* likely to be hired by such hospitals during their careers.

Finally, we investigated whether CEOs who have a large positive pay effect – and are therefore presumably perceived by the market as good performers – were more likely to be hired at problematic hospitals. The results of this analysis are presented in Table W-38. CEOs with large positive pay effects are less likely to be hired at hospitals rated as low quality, though the estimate is not statistically significantly different from zero.[50]

performance in the first hospital is associated with being hired at a more prestigious hospital. Specifically, performing well in terms of staff satisfaction is positively associated with moving to a foundation trust. When looking at the targeted hospital production measures, none of the 15 coefficients are statistically significant at the 10% level, though there is some indication that better performance in terms of MRSA rates is associated with moving to one of the biggest hospitals. On the other hand, good performance in terms of day cases seems to reduce the probability of moving to a big hospital.

[49]We re-ran both analyses – the linear probability models and the variability impacts – for the stacked sets of input, throughput and clinical performance measures. The results are in Web Appendix W-7. They show a similar lack of evidence of CEOs who were good performers at their first hospital subsequently being hired by a problematic hospital and of CEO variability in performance being associated with being at a more problematic hospital at some point whilst we observe them.

[50]Rating for some hospitals at which a CEO has worked are not available. If only one rating is available we base our definition of "problematic" on this rating. If no rating is available, the CEO is dropped from this analysis.

Furthermore, CEOs with large positive pay effects are significantly less likely to work at hospitals with low surplus, while we find no statistically significant association with merger status and PFI contracts.[51]

In sum, consistent with the results obtained using performance measures, there is little evidence that highly paid CEOs are more likely to be hired by problematic hospitals.

Negative Selection within the NHS Another flavor of the endogenous assignment hypothesis is that the NHS internal labor market may be systematically unable retain good CEOs within the system (though, as mentioned above, being a CEO of an NHS hospital is a well paid public sector post, and the hospital sector is the more prestigious part of the NHS). In this case, the lack of CEO effects would be due to negative selection of the pool of managers used to estimate the CEO fixed effects, since these managers are by definition stayers, i.e. managers with at least two stints as CEOs within the NHS system.

To investigate this hypothesis, we use all observed CEO spells (rather than the mover sample consisting of the subset of CEO spells included in the parametric two-step approach) to examine the relationship between CEO performance and leaving the sample after being observed in only one hospital. We define CEO performance using the mean of of the residuals e_{jt} from Equation 1 for the financial years $t_1^{i,A}$ to $t_n^{i,A}$ when CEO i is observed in hospital A. For surplus, day cases, staff job satisfaction we classify as good performers CEOs with a residual mean at or above the 75$^{\text{th}}$ percentile and as bad performers CEOs whose residual mean is at or below the 25$^{\text{th}}$ percentile. For waiting times, length of stay and MRSA rate the definitions of good and bad performers is reversed.

Table W-40 presents the association between good and bad performance and the probability of exiting the sample. In general, there is no systematic evidence that managers who perform exceptionally well in their first observed stint as CEO are more likely to exit the sample. Good performance in terms of length of stay is associated with a higher probability of exiting the sample, though the coefficient is statistically significant only at the 10% level. On the other hand, bad performance in terms of a CEO's financial bottom line is strongly associated with exiting the sample. This result suggests that, if anything, stayers are positively selected, at least in terms of their ability to generate a financial surplus.[52]

[51] In complementary analysis (Table W-39), we find that highly paid CEOs are more likely to be matched with teaching and larger hospitals.

[52] We re-run this analysis for the stacked sets of throughput measures and clinical performance measures. The results are in Web Appendix W-7. There is no evidence that manager who perform well in their first observed CEO stint are more likely to exit the sample.

Summary Taken together, these results suggest that the absence of CEO effects is not driven by (negative) endogenous selection, either at the intensive margin (i.e. the allocation of CEOs within the NHS) or extensive margin (i.e. the selection of managers who stay within the NHS system).

6.2 Match Effects

We find no evidence that CEOs can carry good performance from one hospital to another. But perhaps certain individuals perform better in certain environments i.e. there might be match effects, such that some individuals can achieve good outcomes in a certain environment but cannot transfer this performance to another environment.[53]

To estimate the importance of such match effects, we explore the association between our hospital production measures and a range of combinations of CEO characteristics and hospital characteristics, an approach similar to Jackson (2013). The CEO characteristics we examine are those associated with differences in CEO pay effects (see Table W-4 in Web Appendix W-2), i.e. observable managerial characteristics that are rewarded in the labor market for CEOs. These are: being female, being a doctor and having leadership experience in the private sector. We consider a selected set of salient hospital characteristics on which NHS (and other) hospitals differ: teaching status, foundation trust status (hospitals which were granted more autonomy by the regulator), and large hospitals (measured in terms of beds).[54]

We estimate the effect of interactions of CEO characteristics W_i and hospital characteristics W_j on the targeted hospital production measures and staff job satisfaction using the following regression:

$$y_{jt} = \boldsymbol{X}'_{jt}\boldsymbol{\beta} + \lambda_t + \psi_j + \delta W_i + \gamma(W_i \times W_j) + \varepsilon_{jt} \tag{3}$$

y_{jt} is one of targeted hospital production measures or staff job satisfaction, $\boldsymbol{X}_{jt}$ are time-varying observable hospital characteristics, λ_t are a set of financial year effects and ψ_j a set of hospital effects. The parameter of interest is γ, the coefficient on the interaction term $W_i \times W_j$. We also include the constitutive terms of the interaction term: the CEO characteristic directly as W_i and the hospital characteristic indirectly through the hospital fixed effects ψ_j.

[53] Card et al. (2013) discuss the identification threat posed by sorting based on the idiosyncratic match component of wages when estimating worker and firm fixed effects in pay in the tradition of Abowd et al. (1999). The goal of our analysis is not to assess whether CEOs sort into hospitals based on a match effect but to explore if match effects exist in the first place.

[54] These hospital characteristics are also associated with higher CEO remuneration.

Table 6: Estimates of quality of CEO-hospital matches for all observed CEO spells

	Surplus	Waiting time	Day cases	Length of stay	MRSA rate	Staff job satisf.
Female * Teach	−0.66	0.22	−0.47*	0.20***	−0.76	0.12
	(0.57)	(0.16)	(0.27)	(0.07)	(0.47)	(0.09)
Female * FT	−0.05	0.31***	−0.06	−0.00	0.22**	−0.07
	(0.10)	(0.11)	(0.13)	(0.04)	(0.10)	(0.09)
Female * Comp	0.12	−0.09	−0.07	−0.02	0.05	−0.10
	(0.13)	(0.12)	(0.09)	(0.04)	(0.12)	(0.14)
Female * Beds (100s)	−0.02	−0.00	−0.00	0.00	−0.02*	0.00
	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Doctor * Teach	0.70	0.03	0.32**	−0.00	−0.90**	0.21
	(0.58)	(0.18)	(0.14)	(0.14)	(0.45)	(0.15)
Doctor * FT	−0.20	0.02	−0.04	−0.08	−0.22	−0.14
	(0.24)	(0.16)	(0.12)	(0.07)	(0.31)	(0.18)
Doctor * Comp	−0.02	0.02	−0.15	0.01	0.22	0.18
	(0.22)	(0.16)	(0.15)	(0.07)	(0.31)	(0.13)
Doctor * Beds (100s)	0.05	−0.02	0.02	0.00	−0.02	0.02
	(0.05)	(0.01)	(0.02)	(0.01)	(0.05)	(0.02)
Priv. sec. * Teach	0.05	−0.03	−0.21	0.01	−0.68**	0.03
	(0.21)	(0.17)	(0.20)	(0.08)	(0.27)	(0.13)
Priv. sec. * FT	0.37***	0.10	0.22	−0.13	0.16	−0.04
	(0.12)	(0.19)	(0.20)	(0.08)	(0.21)	(0.16)
Priv. sec. * Comp	0.01	1.33*	0.83	−0.44	0.70***	−0.15
	(0.19)	(0.75)	(0.57)	(0.36)	(0.23)	(0.16)
Priv. sec. * Beds (100s)	−0.01	−0.02	−0.05**	0.01	−0.04***	0.01
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Observations	2,396	2,356	2,383	2,386	2,055	1,838

Teach = teaching hospital, FT = foundation trust, Comp = competitive market. Each estimate is from a separate regression of the relevant hospital production measure on hospital characteristics, financial year effects, hospital effects, the relevant CEO characteristic and the interaction of the relevant CEO characteristic and hospital characteristic. All outcome variables are standardised to have a mean of zero and a standard deviation of one. Standard errors in (parentheses), clustered at hospital level. *Significant at 10%, **significant at 5%, ***significant at 1%

Table 6 presents the results. Each entry is the coefficient on the interaction term from a separate regression. The rows indicate the interactions and are ordered by person characteristics. The columns are the production measures. The table indicates that, whilst there are not match effects for all combinations of person and hospital type, there do seem to be hints of match effects which fit with the institutional set-up of the hospitals we study.

First, female CEOs in teaching hospitals have fewer day cases and longer length of stay while female CEOs in foundation trusts have longer waiting times and higher MRSA rates.[55] Second, CEOs who are doctors are associated with more day cases and lower MRSA rates in teaching hospitals. These hospitals are exactly those settings where it might be expected that clinically trained CEOs perform best.[56] Third, CEOs with leadership experience in the private sector appear to respond to external incentives more than those with an NHS background, in that they achieve higher surpluses when paired with a foundation trust, a setting in which surpluses are important because they were required to achieve foundation trust status. When paired with a teaching hospital or a larger hospital, CEOs with private sector experience have lower MRSA rates. Interestingly, they have longer waiting times and higher MRSA rates when operating in a more competitive market. In Web Appendix Table W-41 we repeat this analysis for the subset of CEO spells included in the parametric two-step approach (sample of movers). The results are broadly similar.

Overall, for the CEO characteristics gender, clinical background and private sector experience, it appears that certain CEO-hospital matches are associated with differences in hospital production. This analysis is subject to the caveat that our samples are small and any match effects may therefore be driven by a small number of individuals observed in particular hospital settings.[57]

6.3 Short Tenure

CEOs in the NHS system have short tenure. The average tenure of our overall sample is 3.8 years, and for the mover CEOs observed in two hospitals for at least two years 4.5 years. Whilst rapid turnover of managers is not a necessary feature of the NPM model,

[55] Additional results for the stacked hospital production measures in Web Appendix W-7 suggest that female CEOs in teaching hospitals perform better in terms of non targeted clinical performance but at the expense of lower throughputs (which are more easily observed than clinical outcomes, perhaps explaining the lower pay of female CEOs).

[56] Goodall (2011) and Kakemam & Goodall (2019) present empirical evidence for the hypothesis that domain experts are better leaders in the hospital sector.

[57] Matching may also occur along CEO and hospital characteristics that are unobservable to us (e.g. CEO social or leadership skills).

the focus on meeting key targets and giving more autonomy to senior managers to deliver has meant that the NPM model has been accompanied by a 'turnaround' leader model. These are CEOs who come into organisations and re-focus them on the key priorities of government and then leave for another organisation. But this model, whilst common, may not be appropriate to the hospital environment–our finding of a lack of persistent impact of CEOs may results from their tenure being too short for any impact on hospital production to materialize consistently across hospitals. To examine this possibility, we look at the association between tenure and CEO performance.

Table 7: Association of tenure and residuals for all observed CEO spells and for subset of CEO spells included in parametric two-step approach (sample of movers)

	All CEO spells			CEO spells for sample of movers		
	Coefficient on tenure var. (std. error)	R^2	Obs.	Coefficient on tenure var. (std. error)	R^2	Obs.
Surplus	0.010* (0.006)	0.001	2,541	0.025** (0.012)	0.004	858
Waiting time	-0.001 (0.006)	0.000	2,501	-0.02 (0.012)	0.012	831
Day cases	-0.000 (0.006)	0.000	2,527	-0.012 (0.008)	0.006	852
Length of stay	0.002 (0.003)	0.001	2,528	-0.004 (0.004)	0.003	843
MRSA rate	-0.001 (0.007)	0.000	2,223	-0.029** (0.012)	0.018	709
Staff job satisfaction	0.015** (0.006)	0.005	2,012	0.025* (0.013)	0.012	629

All regressions include a dummy variable indicating that tenure is unsure. The residuals are from a regression of the standardised measures on hospital characteristics, financial year effects and hospital effects. "All CEO spells" excludes spells at hospitals which we observe for only one year since hospital effects predict the outcome variable perfectly. "All CEO spells" also excludes hospital-year observations for which we observe fewer than 2 of the hospital production measures in a stacked set. Standard errors clustered at hospital level. *Significant at 10%, **significant at 5%, ***significant at 1%

To do this we use the residuals e_{jt} generated by Equation (1), i.e. hospital-year deviations from the expected level of our different hospital production measures. We separately examine all CEO spells in our data set and the subset of CEO spells in the mover sample using the regression:

$$e_{jt} = \alpha + \delta tenure_{i(j,t)} + \gamma tenure\_unsure_{i(j,t)} + \varepsilon_{jt} \quad (4)$$

The *tenure* variable takes discrete values between 1 and 14. We allow for left censoring in our data by including the indicator variable *tenure_unsure*, which takes the value 1 for CEO spells that start in 2000 and 0 otherwise. The function $i(j, t)$ maps hospital j to CEO i in financial year t.

Table 7 presents the results. For the full sample, the estimates suggest that longer tenure is associated with higher surpluses and higher job satisfaction. This latter finding mirrors the finding from our non-parametric approach that a CEO turnover event reduces staff job satisfaction. A similar pattern can be seen in the smaller mover sample. In addition, in this sample longer tenure is associated with lower MRSA rates. In terms of size, the coefficient estimates for the smaller sample suggest that 10 years of tenure is associated with surplus and staff satisfaction being one-quarter of a standard deviation above the expected level and MRSA rates being one-quarter of a standard deviation below the expected level.[58]

These results provide some support to the idea that the short tenure of CEOs in the NHS may dampen their ability to systematically achieve the targets they are set or to impress their mark on the organizations they lead. Of course, the possible endogeneity of tenure makes it hard to pin down the direction of causality behind these correlations.

7 Conclusions

In this paper we examine one of the key tenets of the New Public Management model: that CEOs of public sector organizations can have an impact on the performance of those organizations. We focus, in contrast with previous studies, on large and complex organisations: the public hospitals of the English National Health Service, for which successive governments have set key targets against which the performance of CEOs is to be measured.

We find little evidence that CEOs are able to change the performance of hospitals with respect to these key targets, or indeed with respect to almost any of wide range of measures of production which are easily observed by the government. We explore why this may be the case. Our results do not seem to be driven by the endogenous allocation of

[58]The stacked production measures show similar patterns but also small effect sizes. See Web Appendix W-7.

better performing CEOs–as measured either by performance or in terms of their individual pay effect–to worse performing hospitals. Nor do they seem to be driven by the exit of better performing CEOs from the NHS. However, we provide suggestive evidence that NHS CEOs may be in post for too short a time to have an effect, and also that certain CEOs may matter in certain conditions, i.e. that there are CEO-hospital match effects. These match effects, together with the short tenure of hospital CEOs in the English NHS system, may be the reason why we also find that certain CEOs are systematically paid more than others: when the average tenure period is short, the market cannot distinguish between a good match and good performance in all settings.

More broadly, there are at least two possible explanations for our findings. The NHS is central in political discourse in the UK. Its importance means that politicians are very concerned about NHS performance, particularly negative performance, and are also keen to be seen to be doing something, which is generally manifest in a desire to implement new policies. The lack of persistent CEO effects is consistent with a scenario in which top managers simply chase political goals, rather than policies that might actually improve hospital performance (see, for example, qualitative studies in Powell & Davies (2016)). In this context, the rational response of a NHS CEO is not necessarily to improve the long-term performance of the hospital or even hit key targets but, instead, to simply minimize the amount of bad news that ends up on the Secretary of State's desk. This situation may explain why there is a CEO effect in remuneration that is not associated with observed hospital performance but is associated with receiving a public honour. Furthermore, the political nature of the NHS may also lead to reluctance of high performers to seek CEO appointments, thus inducing negative sorting.

A second explanation is that hospitals are large complex organizations, in which highly trained (and hard to monitor) individuals run separate but interconnected production processes. Management at the very top of such organizations may find it difficult to engage in coordination and getting a large number of actors, who traditionally have not worked together, to work cooperatively. Put another way, a possible interpretation of our finding is that the organizational inertia of a large hospital is too strong for a single manager – even if this person is the CEO – to be able to impact performance within the short time period in which they are in office, and consistently across organizations. This situation, of course, is not specific to public sector hospitals. But it may have more of an effect in hospitals, public or private, where there are many dimensions of performance (clinical, access, financial) that can be pursued and can in the short run conflict. This inertia may also be exacerbated by the often much longer contract durations of clinical staff relative to CEOs.

Regardless of the underlying drivers of our results, they raise concerns about the plausibility of the overall New Public Management approach that focuses on the role of top management and advocates the use of transient "turnaround" CEOs to improve the performance of individual hospitals. A leading NHS manager recently argued that it takes five years for a CEO to make a difference but the average time in post is much shorter than that.[59] Coupled with the findings of Tsai et al. (2015) and Bloom, Propper, Seiler & van Reenen (2015) that the management capabilities of middle managers in hospitals are systematically associated with better outcomes, our paper suggests that rather than seeking to rapidly change hospital performance through the appointment of a cadre of "superheads", strategies for improvement should instead focus on nurturing and sustaining the skills of middle managers.

References

Abowd, J., Kramarz, F. & Margolis, D. (1999), 'High wage workers and high wage firms', *Econometrica* **67**(2), 251–333.

Andrews, M. J., Gill, L., Schank, T. & Upward, R. (2008), 'High wage workers and low wage firms: negative assortative matching or limited mobility bias?', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **171**(3), 673–697.

Ashraf, N., Bandiera, O. & Jack, B. K. (2014), 'No margin, no mission? a field experiment on incentives for public service delivery', *Journal of Public Economics* **120**, 1 – 17.

Baggott, R. (1994), *Health and Health Care in Britain*, Macmillan, Basingstoke.

Ballantine, J., Forker, J. & Greenwood, M. (2008), 'The governance of CEO incentives in English NHS hospital trusts', *Financial Accountability & Management* **24**(4), 385–410.

Bamber, L. S., Jiang, J. X. & Wang, I. Y. (2010), 'What's my style? The influence of top managers on voluntary corporate financial disclosure', *The Accounting Review* **85**(4), 1131–1162.

Bell, B. & Van Reenen, J. (2016), CEO pay and the rise of relative performance contracts: A question of governance, CEP Discussion Paper 1439, Centre for Economic Performance.

Bennedsen, M., Perez-Gonzalez, F. & Wolfenzon, D. (2006), Do CEOs matter?, NYU Working Paper FIN-06-032, NYU.

Bertrand, M. & Schoar, A. (2003), 'Managing with style: The effect of managers on firm policies', *The Quarterly Journal of Economics* **118**(4), 1169–1208.

[59] `https://www.hsj.co.uk/workforce/so-what-does-it-take-to-be-a-chief-executive-in-the-nhs/5091689.article`).

Besley, T. & Ghatak, M. (2003), 'Incentives, choice, and accountability in the provision of public services', *Oxford Review of Economic Policy* **19**(2), 235–249.

Bloom, N., Lemos, R., Sadun, R. & Van Reenen, J. (2015), 'Does management matter in schools?', *The Economic Journal* **125**(584), 647–674.

Bloom, N., Propper, C., Seiler, S. & van Reenen, J. (2015), 'The impact of competition on management quality: Evidence from public hospitals', *Review of Economics Studies* **82**, 457–489.

Böhlmark, A., Grönquist, E. & Vlachos, J. (2016), 'The headmaster ritual: The importance of management for school outcomes', *The Scandinavian Journal of Economics* **118**(4), 912–940.

Branch, G. F., Hanushek, E. A. & Rivkin, S. G. (2012), Estimating the effect of leaders on public sector productivity: The case of school principals, NBER Working Paper 17803, National Bureau of Economic Research.

Burgess, S., Propper, C., Ratto, M. & Tominey, E. (2017), 'Incentives in the public sector: Evidence from a government agency', *Economic Journal* **127**(605), F117–F141.

Cadbury, A. (1992), *Report of the committee on the financial aspects of corporate governance*, Gee, London.

Card, D., Heining, J. & Kline, P. (2013), 'Workplace Heterogeneity and the Rise of West German Wage Inequality*', *The Quarterly Journal of Economics* **128**(3), 967–1015.

Chandra, A., Finkelstein, A., Sacarny, A. & Syverson, C. (2016), 'Health care exceptionalism? Performance and allocation in the US health care sector', *American Economic Review* **106**(8), 2110–44.

Coelli, M. & Green, D. A. (2012), 'Leadership effects: School principals and student', *Economics of Education Review* **31**(1).

Dejong, D. & Ling, Z. (2013), 'Managers: Their effects on accruals and firm policies', *Journal of Business Finance & Accounting* **40**(1-2), 82–114.

Department of Health (2000), 'The NHS plan: a plan for investment, a plan for reform'.

Department of Health (2001), 'NHS performance ratings: Acute trusts 2000/01'.

Department of Health (2004), 'The NHS improvement plan: Putting people at the heart of public services'.

Dixit, A. (2002), 'Incentives and organizations in the public sector: An interpretative review', *The Journal of Human Resources* **37**(4), 696–727.

Fee, C., Hadlock, C. & Pierce, J. (2013), 'Managers with and without style: Evidence using exogenous variation', *The Review of Financial Studies* **26**(3), 567–601.

Fenizia, A. (2019), Managers and productivity in the public sector. Mimeo, Available at: `https://sites.google.com/site/alessandrafenizia/home`.

Gaynor, M., Laudicella, M. & Propper, C. (2012), 'Can governments do it better? merger mania and hospital outcomes in the english nhs', *Journal of Health Economics* **31**(3), 528 – 543.

Goodall, A. H. (2011), 'Physician-leaders and hospital performance: Is there an association?', *Social Science & Medicine* **73**(4), 535 – 539.

Gravelle, H., Sutton, M. & Ma, A. (2010), 'Doctor behaviour under a pay for performance contract: Treating, cheating and case finding?', *The Economic Journal* **120**(542), F129–F156.

Grissom, J. A., Kalogrides, D. & Loeb, S. (2015), 'Using student test scores to measure principal performance', *Educational Evaluation and Policy Analysis* **37**(1), 3–28.

Ham, C., Baker, G. R., Docherty, J., Hockey, P., Lobley, K., Lord Tugendhat & Walshe, K. (2011), *The Future of Leadership and Management in the NHS: No more heroes*, The King's Fund, London.

Healthcare Commission (undated), 'The annual health check 2006/07: Directory of performance ratings for nhs trusts'.

Heckman, J. J., Heinrich, C. & Smith, J. (2002), 'The performance of performance standards', *The Journal of Human Resources* **37**(4), 778 – 811.

Jackson, C. K. (2013), 'Match quality, worker productivity, and worker mobility: Direct evidence from teachers', *Review of Economics and Statistics* **95**(4), 1096–1116.

Jha, A. K. & Epstein, A. (2013), 'A survey of board chairs of English hospitals shows greater attention to quality of care than among their US counterparts', *Health Affairs* **32**(4), 677–685.

Joynt, K., Lee, S., Orav, J. & Jha, A. (2014), 'Compensation of chief executive officers at nonprofit us hospitals', *JAMA Internal Medicine* **174**(1), 61–67.

Kakemam, E. & Goodall, A. H. (2019), 'Hospital performance and clinical leadership: new evidence from Iran', *BMJ Leader* **3**(4), 108–114.

Kline, P., Saggio, R. & Sølvsten, M. (2018), 'Leave-out estimation of variance components'.

Lavy, V. & Boiko, A. (2017), Management quality in public education: Superintendent value-added, student outcomes and mechanisms, NBER Working Paper 24028, National Bureau of Economic Research.

Le Grand, J. (2003), *Motivation, agency, and public policy: Of knights and knaves, pawns and queens*, Oxford University Press, Oxford.

Limodio, N. (2018), Bureaucrats allocation in the public sector: Evidence from the World Bank. Mimeo, Available at: `https://www.dropbox.com/s/zexwd94lrhnlgh4/Limodio_WBManagers.pdf?dl=0`.

Monitor (2014), 'The NHS foundation trust code of governance'.

National Audit Office (2004), 'Achieving first-class financial management in the NHS: A sound basis for better healthcare'.

National Audit Office (2009), 'Reducing healthcare associated infections in hospitals in England'.

Naylor, C. & Gregory, S. (2009), *Briefing: Independent Sector Treatment Centres*, The King's Fund, London.

Pollitt, C. & Bouckaert, G. (2000), *Public management reform: A comparative analysis - Into the age of austerity*, 4 edn, Oxford University Press, Oxford.

Powell, A. & Davies, H. (2016), Managing doctors, doctors managing, Research report, Nuffield Trust.

Prendergast, C. (2001), Selection and oversight in the public sector, with the los angeles police department as an example, Working Paper 8664, National Bureau of Economic Research.

Prendergast, C. (2002), 'Consumers and agency problems', *The Economic Journal* **112**(478), C34–C51.

Propper, C., Croxson, B. & Shearer, A. (2002), 'Waiting times for hospital admissions: the impact of gp fundholding', *Journal of Health Economics* **21**(2), 227 – 252.

Propper, C., Sutton, M., Whitnall, C. & Windmeijer, F. (2010), 'Incentives and targets in hospital care: Evidence from a natural experiment', *Journal of Public Economics* **94**(4), 318–335.

Rasul, I. & Rogger, D. (2018), 'Management of bureaucrats and public service delivery: Evidence from the Nigerian Civil Service', *The Economic Journal* **128**(608), 413–446.

The Health Foundation (2015), 'Evidence scan: The impact of performance targets within the NHS and internationally'.

Tsai, T. C., Jha, A. K., Gawande, A. A., Huckman, R. S., Bloom, N. & Sadun, R. (2015), 'Hospital board and management practices are strongly related to hospital performance on clinical quality metrics', *Health Affairs* **34**, 1304–1311.

Appendix: A Variable Definitions and Sources

Tables A-1 and A-2 provide the data sources for all variables. The pay data are available only in bands of £5,000. We use the midpoint for each band as an approximation of the underlying continuous variable. For example, a basic salary reported as £120,000-£125,000 is recorded as £122,500 in our data set. The time-varying observable hospital level variables, $\boldsymbol{X}_{j(i,t)}$ are foundation trust status, year of merger, years since merger, beds, technology index and case mix variables.

Technology Index The technology index can take any value between 0 and 1. It is the weighted average of 7 dummy variables indicating the availability of advanced technologies: a neonatal intensive care unit, a cardiology unit, magnetic resonance imaging, imaging using radio-isotopes, heart or lung transplants, open heart surgery and percutaneous coronary interventions. The weight for each of these technologies is the proportion of hospitals that do not possess that technology at the beginning of our sample in 2000/01. The resulting index value increases over the sample period as hospitals add technologies.

We use data from a wide range of administrative sources to generate the 7 dummy variables. A hospital is defined as having a neonatal intensive care unit if it has at least one bed in a neonatal intensive care unit, as reported in the beds data published annually in the Hospital Activity Statistics by NHS England. A hospital is defined as having a cardiology unit if according to annual Hospital Episode Statistics it delivered at least 10 finished consultant episodes in a cardiology speciality. We define a hospital as offering magnetic resonance imaging if according to the annual imaging data published as part of Hospital Activity Statistics by NHS England it delivered at least 100 examinations or tests using magnetic resonance imaging. Numbers in this data set tend to be around 1,000 to 30,000; so numbers smaller than 100 might be data entry errors. Similarly, we define a hospital as offering imaging using radio-isotopes if the annual imaging data reports at least 100 examinations or tests using radio-isotopes.

Further, we define a hospital as providing heart or lung transplants if the annual Hospital Episode Statistics report at least 2 transplant procedures (HRGs E01 and E02), as providing open heart surgery if the annual Hospital Episode Statistics report at least 10 open heart surgery procedures (HRGs E01 to E04) and as providing percutaneous coronary interventions (OPCS codes K49 and K75) if the annual Hospital Episode Statistics report at least 10 such interventions.

Once a dummy variable takes the value one, we set its value to one in all following years, to avoid fluctuations that are most likely caused by data entry errors rather than

Table A-1: Variable definitions and sources: Outcome variables

Variable	Definition	Source
Basic pay	Basic remuneration, CPI adjusted (£)	IDS Incomes Data Services and remuneration reports in hospitals' annual reports
Total pay	Total remuneration excluding redundancy payments, CPI adjusted (£)	
Surplus	Retained surplus/deficit (£000)	Trust Financial Returns
(Doctors + nurses)/beds	Ratio of all medical staff and nurses (full-time equivalent) to beds	NHS Hospital and Community Health Service in England workforce statistics, Health and Social Care Information Centre, now NHS Digital
Senior doctors/staff	Consultants, associate specialists, staff grade, registrars as proportion of all staff (%)	
Nurses/staff	Qualified nursing, midwifery, health visiting staff as prop. of all staff (%)	
Technology index	Details in text	Various sources
Beds	Average daily number of available beds	NHS England
Admissions	Number of admissions (count)	Hospital Episode Statistics: Admitted Patient Care, Health and Social Care Information Centre, now NHS Digital
Length of stay	Mean of spell duration, excluding day cases (days)	
Day cases	Proportion of finished consultant episodes relating to day cases (%)	
Waiting time	Mean time waited between decision to admit and actual admission (days)	
Cancelled operations	Operations cancelled for non-clinical reasons (count)	NHS England
AMI deaths	Deaths within 30 days of emergency admission for acute myocardial infarction, age 35-74 (%)	Clinical and Health Outcomes Knowledge Base (NCHOD), since relaunched as Compendium of Population Health Indicators
Stroke deaths	Deaths within 30 days of emergency admission for stroke, all ages (%)	
FPF deaths	Deaths within 30 days of emerg. adm. for fractured proximal femur, all ages (%)	
Readmissions	Emerg. readmissions to hospital within 28 days of discharge, age 16+ (%)	
MRSA rate	MRSA bacteraemia rate per 100,000 bed days	Public Health England
Staff job satisfaction	Scores from 1 to 5, 1 = dissatisfied, 5 = satisfied, mean	NHS Staff Survey

Table A-2: Variable definitions and sources: Control variables

Variable	Definition	Source
Foundation Trust	Dummy variable taking value 1 once a hospital has achieved Foundation Trust status, 0 otherwise	Monitor, now NHS Improvement
Year of merger	Dummy variable taking value 1 in year hospital newly created through merger enters sample, 0 otherwise	Information on hospitals' websites and Statutory Instruments (www.legislation.gov.uk)
Years since merger	Variable taking value 1 in year after hospital newly created through merger enters sample, value 2 in following year and so on, 0 otherwise	
Acquisition	Dummy variable taking value 1 once hospital has been involved in merger that is more like acquisition, i.e. following merger hospital keeps its provider code while provider code of other hospital disappears from any records, 0 otherwise	
Beds	Average daily number of available beds	NHS England
Technology index	Details in text	Various sources
Patients aged 0 to 14	Finished Consultant Episodes (FCEs) involving patients aged 0 to 14/Total FCEs	Hospital Episode Statistics: Admitted Patient Care, Health and Social Care Information Centre, now NHS Digital
Patients aged 60 to 74	FCEs involving patients aged 60 to 74/Total FCEs	
Patients aged 75+	FCEs involving patients aged 75+/Total FCEs	
Male patients	FCEs involving male patients/Total FCEs	
Major teaching hospital	Dummy variable taking value 1 if hospital serves medical school as their major NHS partner, 0 otherwise	The Guardian Healthcare Professionals Network, Wikipedia and information on medical schools' websites
Minor teaching hospital	Dummy variable taking value 1 if hospital is not major teaching hospital but member of the Association of UK University Hospitals	Association of UK University Hospitals, now University Hospital Association
Specialist status	Hospital is specialist acute, children's or orthopaedic hospital	NHS Staff Survey
Hospital commission rating	Details in text	Various sources
Competitive market	Details in text	Own calculations

real changes.

Hospital Commission Rating We use ratings for the financial years 2002/03 to 2007/08. Ratings for the years 2002/02 to 2004/05 used stars, with three stars awarded to hospitals with the "highest levels of performance", two stars awarded to hospitals that are "performing well overall, but have not quite reached the same consistently high standards", one star awarded to hospitals "where there is some cause for concern regarding particular key targets" and zero stars awarded to hospitals "that have shown the poorest levels of performance against key targets" (Department of Health 2001). We classify zero stars and one star as a poor rating.

For the years 2005/06 to 2007/08 the Hospital Commission published ratings using a four-point scale of "excellent", "good", "fair" and "weak" (Healthcare Commission undated). Each hospital received two scores, one for quality of services and one for use of resources (Healthcare Commission undated). We use the score for quality of services and classify scores of "fair" and "weak" as a poor rating. Ideally, we want to use the hospital commission rating from the year before the CEO arrived. Because of data limitations we can use this definition only for the financial years 2003/04 to 2008/09. For the financial year 2002/03 we use the contemporanous rating, for 2009/10 the rating from two years before the CEO arrived and for 2010/11 the rating from three years before the CEO arrived.

Competitive Market The competitive market variable is a dummy variable that takes the value 1 if the hospital is located in a competitive market. To measure hospital competition we use a method similar to Bloom, Propper, Seiler & van Reenen (2015). First, we calculate the Euclidean distance (in km) for each pairwise combination of all hospitals in our data. For this calculation we use the geographical coordinates associated with the postcode of the hospital trust's headquarters. Next, we count the number of competitors within a 30 km radius. Finally, we calculate the quartiles of these counts for each region. If a hospital is in the top quartile for its region, we classify it as operating in a competitive market.

Appendix: B Balance of Matched Samples for Non-parametric Approach

Table B-1: Means of matching variables and other hospital characteristics for treated and control groups and means of exactly matched hospital characteristics

				Means of variables measured in $t-1$						Means of vars. measured in t		Exactly matched charact. (%)			
						Prop. in each category (%)						Teaching			
		Obs.	Unique con-trols	Beds	Tech-nology	0-14	60-74	75+	male	Year of merger (%)	Years since merger	Major	Minor	Spec. hosp.	Foun-dation trust in $t-1$
Surplus	T	205		722	0.35	5.31	8.14	8.29	43.5	0	1.22	12.2	9.3	9.8	26.8
	C	205	160	714	0.39	5.05	8.29	8.24	44.3	0.98	0.82				
Waiting time	T	200		731	0.35	5.12	8.23	8.36	43.6	0	1.26	12.5	9.0	9.0	26.5
	C	200	157	725	0.39	5.07	8.29	8.31	44.2	1.00	0.82				
Day cases	T	202		725	0.35	5.33	8.12	8.32	43.5	0	1.24	12.4	9.4	9.4	27.2
	C	202	156	718	0.39	5.05	8.29	8.26	44.2	0.99	0.86				
Length of stay	T	205		722	0.35	5.31	8.14	8.29	43.5	0	1.22	12.2	9.3	9.8	26.8
	C	205	159	715	0.39	5.05	8.29	8.23	44.2	0.98	0.84				
MRSA rate	T	197		729	0.35	5.34	8.12	8.25	43.5	0	1.27	12.2	9.6	9.6	27.9
	C	197	156	721	0.40	5.02	8.36	8.23	44.4	1.00	0.82				
Staff sat-isfaction	T	163		713	0.36	5.18	8.18	8.46	43.5	0	1.48	11.0	9.8	9.2	33.7
	C	163	126	711	0.40	5.06	8.36	8.42	44.5	1.23	0.99				

T = Treated, C = Controls

Web Appendix: For Online Publication

W-1 Differences between SIC 92 and SIC 2007 industry classifications

Figure 2 in Section 2.2 presents pay of top managers in industries with both public and private sector organizations over the period 2000 to 2017. As the industry classification was changed substantially in 2009, we present separate graphs for 2000-2008 and 2009-2017. Here we describe the relevant changes in more detail.

Data for 2000 to 2008 use the "UK Standard Industrial Classification of Economic Activites - SIC 92" while data for 2009 to 2017 use the the "UK Standard Industrial Classification of Economic Activities - SIC 2007". In SIC 92 the section "K - Real estate, renting and business activities" includes "73.10 Research and experimental development on natural sciences and engineering" and "73.20 Research and experimental development on social sciences and humanities". In SIC 2007 these activities have been subsumed into the new section "M - Professional, scientific and technical activities". The SIC 92 section "I - Transport, storage and communication" includes "60.1 Transport via railways", "60.21 Other scheduled passenger land transport" and "64.11 National post activities". These industries are comparable to the SIC 2007 section "H - Transportation and storage", which includes "49.10 Passenger rail transport interurban", "49.31 Urban and suburban passenger land transport" and "53.10 Postal activities under universal service obligation".

The SIC 92 section "J - Financial intermediation" includes "65.11 Central banking" and "66.02 Pension funding". Similarly, the SIC 2007 section "K - Financial and Insurance Activities" includes "64.11 Central banking" and "65.30 Pension funding". The SIC 92 section "L - Public administration and defence, compulsory social security" includes "75.24 Public security, land and order activities", which is comparable to the SIC 2007 section "O - Public administration and defence, compulsory social security", which includes "84.23 Justice and judicial activities" and "84.24 Public order and safety activities". The SIC 92 section "O - Other community, social and personal service activities" includes "92 Recreational, cultural and sporting activities". In SIC 2007 these activities have been subsumed into the new section section "R - Arts, entertainment and recreation".

W-2 CEO Fixed Effects in Pay: Data, Methods and Results

Data We obtained data on CEO pay from the NHS Boardroom Pay Reports published by IDS Incomes Data Services for 2000/01 to 2010/11. We extended these data by hand-collecting data from hospitals' annual reports for 2011/12 to 2013/14. These reports provide data on salary, taxable benefits and total remuneration of executive directors for nearly all NHS hospitals.

The core executive director positions present on all hospital boards are CEO, Medical Director, Nursing Director, Finance Director and HR Director.[60] In the later years of our panel we also regularly observe a Chief Operating Officer. Additionally, there is a range of other positions such as Director of Facilities and Estate Development or Director of Information Management and Technology, which we categorize as "Other". To ensure comparability, we drop from the pay data all observations that refer only to part of the financial year (for example, because an executive director left the hospital at some point during the year).

We combined the CEO pay data with data on time-varying hospital characteristics, which we include as control variables in our pay regressions. They include governance measures, capital measures and case-mix measures.

Table W-1 presents descriptive statistics for the pay data and for the hospital-level control variables. For each variable, we show the overall mean and standard deviation as well as the mean at the beginning, in the middle and at the end of our sample period. Average CEO pay (adjusted for inflation) went from £99,000 in 2000 to £138,000 in 2013. Average pay of executive directors other than the CEO also increased but less steeply. The proportion of hospitals who achieved Foundation Trust status steadily increased over our sample period, reaching 62% in 2013. Consolidation of NHS hospitals fluctuated over our sample period. On average, each year 1.8% of hospitals in our sample had just been created through a consolidation. The number of beds initially increased with hospital consolidation, but then decreased as efficiency improvements in care delivery were achieved (as indicated by the decline in length of stay and the increase in the day case rate). Case-mix became more challenging over our sample period, with the proportion of older patients steadily increasing.

To supplement the CEO pay data, we hand-collected data on the CEO characteristics of gender, educational achievements, clinical background, private sector experience and public honours.[61] Table W-2 summarizes the main characteristics of the 469 CEOs in our

[60]The CEO of an NHS hospital is known as the Chief Executive, but the role is that of a CEO.

[61]The British honours system recognizes people who have made achievements in public life. Titles bestowed on hospital CEOs include Knight, Dame, Commander/Officer/Member of the Order of the

Table W-1: Descriptive statistics for pay data and hospital-level control variables

				Mean of variable in		
	Obs.	Mean	St. Dev.	2000	2006	2013
Executive director pay:						
Basic pay, CPI adjusted (£)	8,749	91,182	26,675	93,672	89,944	96,753
Total pay, CPI adjusted (£)	8,760	92,353	27,454	98,010	90,813	98,069
Executive director total pay, CPI adjusted (£), by position:						
CEO	1,851	126,230	28,612	98,756	126,852	137,917
Finance Director	1,479	94,270	18,993	88,600	94,436	101,300
Chief Operating Officer	779	87,118	18,955	na	85,137	90,666
Nursing Director	1,444	80,428	15,478	71,600	80,078	84,722
HR Director	1,044	76,991	15,060	71,700	75,115	81,531
Other	2,163	79,312	18,271	66,200	78,791	83,686
Governance measures:						
Foundation Trust (%)	2,396	30.1		0	28.8	61.6
Year of merger (%)	2,396	1.75		4.1	0.59	1.26
Years since merger	2,396	1.02	2.67	0	0.89	2.40
Acquisition (%)	2,396	1.38		0	1.18	5.03
Capital measures:						
Beds	2,396	722	402	702	727	683
Technology index	2,396	0.38	0.24	0.29	0.39	0.43
Case-mix measures:						
Patients aged 0 to 14 (%)	2,396	13.5	13.2	14.5	13.4	12.5
Patients aged 60 to 74 (%)	2,396	21.2	6.35	20.3	20.9	22.6
Patients aged 75+ (%)	2,396	20.9	6.85	18.8	20.3	23.6
Male patients (%)	2,396	44.0	5.42	43.3	43.9	44.6

Means of total pay in 2000 are based on only 1 observation each for Finance Director, Chief Operating Officer, Nursing Director, HR Director and Other. Foundation Trust is a dummy variable that takes the value 1 once a hospital has achieved Foundation Trust status. Year of merger takes the value 1 in the year a hospital newly created through a merger enters the sample and 0 otherwise. Years since merger takes the value 1 in the year after a hospital newly created through a merger enters the sample, the value 2 in the following year and so on. Acquisition takes the value 1 once a hospital has been involved in a merger that is more like an acquisition, i.e. following the merger the hospital keeps its provider code while the provider code of the other hospital disappears from any records. Definitions and sources of all variables in Table A-1 in Appendix: A.

sample. About a third are women, many of whom previously had nursing careers. About a quarter have a clinical background, the majority being nurses or allied health professionals rather than doctors. About a quarter hold a postgraduate management qualification.[62] 10% have prior experience in the private sector, sometimes between spells as CEOs in the NHS. Industries range from health care and pharmaceuticals to manufacturing, retail, transport, communications and management consultancy. About 13% of the total sample received a public honour at some point in their career. In terms of tenure as CEO, we observe the vast majority of the sample for more than one year. The maximum is of course restricted by our sample period of 14 years. We observe the majority of CEOs in only one particular CEO job but 30% of CEOs held two or more CEO jobs over our sample period. The median number of years a CEO is observed in a particular CEO job is 3 years and the mean is 3.8 years.

Table W-2: Demographic and sample characteristics of CEOs

	Number	Proportion
Female	147	31%
Clinical background, of which	112	24%
Nurse or allied health professional	79	17%
Doctor	33	7%
Postgraduate management qualification	121	26%
Private sector experience	49	10%
Public honour	60	13%
Number of years observed as CEO:		
1 year	75	16%
2 to 5 years	211	45%
6 to 9 years	105	22%
10 to 13 years	59	13%
14 years	19	4%
Number of CEO jobs observed in:		
1 job	324	69%
2 jobs	105	22%
3+ jobs	40	9%
Observations	469	

British Empire (CBE/OBE/MBE).

[62] These qualifications include straightforward MBAs but many CEOs hold qualifications such as an MSc in Healthcare Management or a Diploma in Health Services Management.

Methods We follow Abowd et al. (1999) to estimate CEO fixed effects in pay. We use the pay data for all executive directors, i.e. including COOs, Finance Directors, HR Directors, Nursing Directors and other directors but excluding Medical Directors as only some of their remuneration is recorded in the executive pay data.[63] We estimate the following wage equation for the largest subset of hospitals that are connected by executive directors moving between them:

$$pay_{ijt} = \boldsymbol{X}'_{jt}\boldsymbol{\beta} + \gamma\, tenure_{ijt} + \boldsymbol{Z}'_{ijt}\boldsymbol{\delta} + \lambda_t + \alpha_i + \psi_j + \varepsilon_{ijt} \quad \text{(W-1)}$$

The left-hand side variable, *pay*, denotes pay of executive director i at hospital j in financial year t. $\boldsymbol{X}_{jt}$ is a vector of the following time-varying observable hospital characteristics: foundation trust status, year of merger, years since merger, beds, technology index, case mix measures (patients aged 0 to 14, patients aged 60 to 74, patients aged 75+, male patients). $tenure_{ijt}$ is the tenure of executive director i at hospital j in financial year t. $\boldsymbol{Z}_{ijt}$ is a vector of dummy variables indicating the board level position of director i at hospital j in financial year t. A full set of financial year effects, λ_t, provides non-parametric control for trends in pay that are national in scope while a full set of hospital effects, ψ_j, controls for non-time varying unobserved differences between hospitals. The estimates of interest are the executive director effects α_i, which capture non-time varying unobserved characteristics that affect directors' pay. ε_{ijt} represents the error term.

As discussed by Abowd et al. (1999), between hospital mobility of the executive directors is essential for the identification of the hospital effects. Including all executive directors, and not just CEOs, increases the size of the set of hospitals connected by worker mobility, and also produces more reliable estimates of the hospital effects. On the other hand, it requires the inclusion of controls for the different board level positions because of the differences in the level of remuneration (see Table 1). The coefficients on these controls are identified by executive directors changing board level position. As we observe a limited amount of movement between positions, we can only control for position effects that are constant over time and across hospitals. Ideally, we would control for hospital-specific position effects and also position-specific tenure effects. Alternatively, we could include only CEOs in our wage equation. Again, observing a limited amount of movement – in this case of CEOs between hospitals – precludes this approach. In our full sample of all executive directors the connected set has 478 movers that connect 196 hospitals, leaving only 162 pay observations in 17 hospitals that are not connected by worker mobility. In the CEO sub-sample the largest connected set has 35 movers con-

[63] Medical Directors salaries in the directors' data are lower than those of other directors. The majority of Medical Director remuneration is for clinical work, which is excluded from the directors' data.

necting 43 hospitals, leaving 1,474 out of the 1,897 pay observations either unconnected or in one of 20 different connected sets, with ten of these connected sets consisting of only two hospitals.

The main purpose of estimating CEO fixed effects in pay is to establish whether – after controlling for observable characteristics such as hospital size and unobservable time-invariant hospital characteristics (ψ_j) – there is dispersion in CEO pay. Such dispersion might suggest that employers perceive CEOs to be differentiated due to some characteristics that are unobservable to us such as managerial effectiveness. Thus, our focus is different from the literature exploiting linked employer-employee data to estimate the impact of worker and firm heterogeneity on labour market outcomes. For one, we are not interested in the correlation between firm and worker effects, estimates of which have been shown to suffer from limited mobility bias (Andrews et al. 2008, Kline et al. 2018). Similarly, we are not interested in estimating the contribution of firms to wage inequality. Thus, we are not overly concerned about potential biases in the estimates of the variance of the hospital fixed effects that have been discussed in the recent literature (Card et al. 2013, Kline et al. 2018).

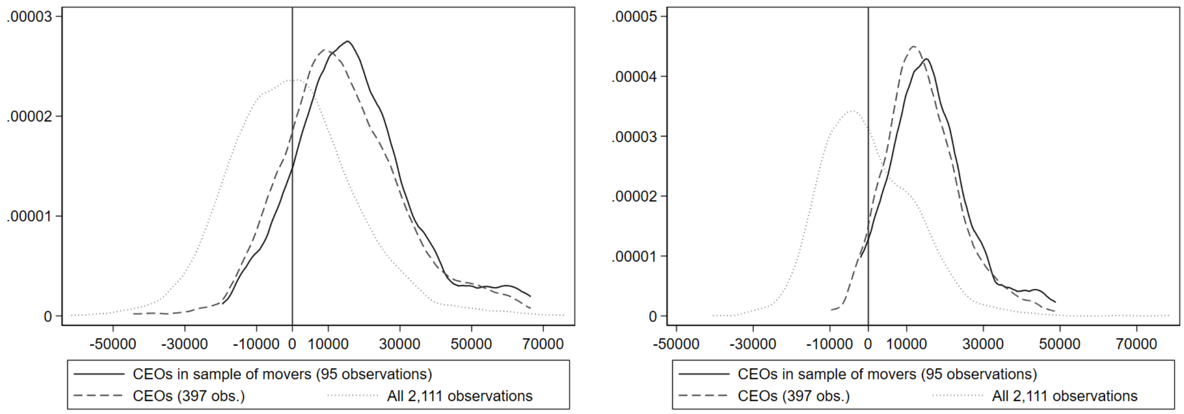
Results Table W-3 reports the results from estimating Equation W-1. We report only the results for total pay as the results for basic pay are very similar. We estimate Equation W-1 only for the pay observations in the largest connected set, which comprises of 478 movers that connect 196 hospitals (the vast majority of the sample). Table W-3 shows that in the connected set the director effects are jointly statistically significant.

Table W-3: Summary statistics for the regression estimating CEO effects in total pay

	F-test of joint significance of director effects (p-value, df1, df2)	R^2	Obs.	Hospitals	Persons	Movers
Connected set	6.52 (0, 2111, 6425)	0.93	8,760	196	2,111	479
Subset of directors observed as CEO at least once	9.84 (0, 397, 6425)		1,851	196	397	171
Subset of CEOs included in parametric two-step approach (sample of movers)	18.3 (0, 95, 6425)		633	122	95	94
Outside connected set			162	17	47	

Table W-3 also presents results for the subset of directors observed in a CEO position

Figure W-1: Kernel density plots of director effects in total pay



(a) Estimates

(b) Shrunk estimates

Notes: Kernel density plots of deviations of estimated director effects in total pay, CPI adjusted, from mean. Shrunk estimates obtained using empirical Bayes shrinkage estimator.

at least once (397 of the 2,111 executive directors in the connected set) and for the further subset of CEOs that are included in the parametric analyses of CEO impacts on hospital production (95 of the 397 CEOs). The director effects are jointly statistically significant in both subsets.

Figure W-1a shows the distribution of the pay effects for all directors in the connected set, the 397 who were ever CEOs and the subset of 95 CEOs. Since the $\hat{\alpha}_i$ are estimated relative to an arbitrary omitted director, we have transformed the estimates into deviations from the mean of all $\hat{\alpha}_i$. The distribution for the 95 CEOs included in the hospital production analyses lies slightly to the right of the distribution for all CEOs. The interquartile range is around £22,000 for the full sample of director effects and around £20,000 for the subsample of directors ever observed as CEO. For the 95 CEOs used in the parametric hospital production analyses (sample of movers) the range is slightly smaller at £18,000. By position, the interquartile range is smallest for Nursing Directors and COOs at around £17,600, followed by HR Directors at £17,900, Finance Directors at £18,100 and directors that we categorize as "Other" at £19,800.

Due to sampling error the distribution of the estimates of the director effects in pay will overstate the true distribution of the effects. We therefore adjust the effect estimates using an empirical Bayes procedure and shrink the effect estimates towards the position means. We employ the empirical Bayes shrinkage estimator described in Chandra et al. (2016). The position means are CEO = £14,212; Finance Director = £8,340; COO = £-5,498; Nursing Director = £-2,806; HR Director = £-6,283; Other = £-8,697.

Figure W-1b presents the distribution of the shrunk effect estimates. As for the unadjusted estimates, the distribution for the 95 CEOs included in the parametric hospital production analyses lies slightly to the right of the distribution for all CEOs. As expected, the distribution is much narrower: at £11,400 the interquartile range for the 95 CEOs is nearly halved compared to the unadjusted estimates, but still sizeable relative to the CEO pay sample mean of £126,230. Considering that our estimation sample is based on a homogeneous set of establishments – acute care hospitals – and that we control for hospital fixed effects as well as time-varying hospital characteristics such as size and case mix, the spread of CEO fixed effects in pay is substantial.

In Table W-4 we examine if any CEO characteristics explain the spread of CEO fixed effects in pay. We estimate the associations between a series of personal and sample-specific characteristics and the CEO pay effects (original estimates as well as shrunk estimates). We also test whether there are differences in these associations between all CEOs in our sample and the 95 CEOs in the parametric hospital production analyses (mover sample). We report results from a series of regressions, each examining one of the CEO characteristics. We find that female CEOs are paid statistically significantly less, with the effect being even stronger for the subset of CEOs included in the hospital production analyses. As being female and some of the other characteristics such as clinical background are highly correlated, we include in the regressions a control for female and for the interaction of female and being in the 95 CEO subset. We find CEOs who received a public honour and CEOs with a clinical background are paid more (the effect being driven by doctors rather than nurses). CEOs with a postgraduate management qualification are paid less. CEOs with leadership experience in the private sector are paid less, but the association is positive for the 95 CEOs subset. CEOs observed in our sample for 10 years or longer are paid more, as are CEOs observed in 3 or more CEO jobs. Importantly, there are few differences in the patterns of these associations between the 95 CEOs included in the parametric analyses of CEO impacts on hospital production and the CEOs not used in these analyses.

Table W-5 presents results from a regression that includes all CEO characteristics simultaneously. The results are broadly similar to the results in Table W-4 but some coefficients are less precisely estimated due to the loss of degress of freedom. Importantly, the R-squared suggests that our measures of CEO characteristics jointly explain only about one-quarter of the variation in the CEO fixed effects in pay, suggesting that the market perceives CEOs to be differentiated by factors that are unobservable to us such as managerial ability or leadership effectiveness.

Figure W-2a shows the distribution of the estimated hospital (firm) effects in pay, $\widehat{\psi_{j(i,t)}}$. Since the $\psi_{j(i,t)}$ are estimated relative to an arbitrary omitted hospital, we trans-

Table W-4: Associations between estimated director effects in total pay and personal characteristics for subset of directors observed in a CEO position at least once

	Original effect estimates	Shrunk effect estimates	Obs. in category
Female	-6,154*** (2,086)	-3,569*** (1174)	125
Female × in 95 CEOs subset	-6,473 (4,223)	-4,311* (2,377)	31
R^2	0.058	0.069	
Nurse or allied health professional	-2,872 (2,906)	-1,921 (1,643)	69
Nurse or AHP × in 95 CEOs subset	-853 (5,870)	-72 (3,318)	20
Doctor	12,836*** (3,700)	5,596*** (2,091)	26
Doctor × in 95 CEOs subset	-3,519 (8,553)	256 (4,834)	5
R^2	0.093	0.096	
Postgraduate managm. qualification	-5,958*** (2,210)	-3,027** (1,246)	107
PG managm. qual. × in 95 CEOs subset	4,852 (4,278)	2,235 (2,412)	31
R^2	0.075	0.083	
Public honour	9,916*** (2,976)	6,074*** (1,664)	51
Public honour × in 95 CEOs subset	815 (5,539)	1,480 (3,097)	16
R^2	0.096	0.118	
Private sector	-5,691* (3,353)	-2,530 (1,889)	42
Private sector × in 95 CEOs subset	11,638** (5,893)	6,199* (3,320)	14
R^2	0.068	0.077	
Observed as CEO for 2 to 9 years	-358 (3,092)	-9.51 (1,726)	290
Observed as CEO for 10 plus years	13,893*** (4,050)	8,578*** (2,260)	76
10 plus years × in 95 CEOs subset	-3,418 (4,557)	-1,595 (2,543)	43
R^2	0.131	0.155	
Observed in 2 CEO jobs	2,593 (3,169)	1,239 (1,778)	102
Observed in 3+ CEO jobs	10,911** (4,572)	6,439** (2,566)	38
3+ CEO jobs × in 95 CEOs subset	1,263 (6,675)	1,020 (3,746)	24
R^2	0.085	0.101	

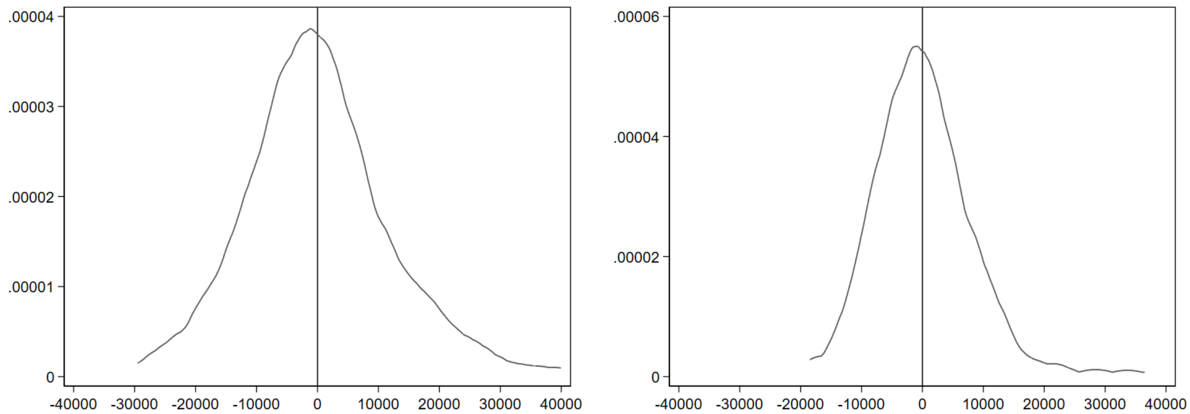
All regressions from the second panel onwards control for female and female × In 95 CEO subset. The executive director effects are extracted from the regressions reported in Table W-3 and transformed into deviations from the mean of all estimated executive director effects. For the shrunk effect estimates column, the deviations are then shrunk towards the overall mean using the empirical Bayes shrinkage estimator described in Chandra et al. (2016). Standard errors in (parentheses). 397 observations in all regressions. *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-5: Joint estimates of the association between estimated director effects in total pay and personal characteristics for subset of directors observed in a CEO position at least once

	Original effect estimates	Shrunk effect estimates	Obs. in category
Female	-6,321*** (2,221)	-3,370*** (1,233)	125
Female × in 95 CEOs subset	1,574 (4,991)	552 (2,770)	31
Nurse or allied health professional	-881 (2,787)	-1,285 (1,546)	70
Nurse or AHP × in 95 CEOs subset	-1,876 (5,651)	-150 (3,136)	20
Doctor	10,592*** (3,623)	5,498*** (2,010)	26
Doctor × in 95 CEOs subset	4,806 (8,522)	4,668 (4,730)	5
Postgraduate managm. qualification	-5,217** (2,113)	-2,582** (1,172)	107
PG managm. qual. × in 95 CEOs subset	3,071 (4,168)	1037 (2,313)	31
Public honour	7,715*** (2,980)	4,844*** (1,654)	51
Public honour × in 95 CEOs subset	1,047 (5,497)	1,478 (3,051)	16
Private sector	-2,928 (3,225)	-1,037 (1,790)	42
Private sector × in 95 CEOs subset	8,467 (5,663)	4,468 (3,143)	14
Observed as CEO for 2 to 9 years	-202 (3,065)	166 (1,701)	290
Observed as CEO for 10 plus years	11,135*** (4,185)	6,793*** (2,322)	76
10 plus years × in 95 CEOs subset	-3,611 (4,907)	-2,027 (2,723)	43
Observed in 2 CEO jobs	1,553 (3,034)	642 (1,684)	102
Observed in 3+ CEO jobs	7,364 (4,477)	4,277* (2,484)	38
3+ CEO jobs × in 95 CEOs subset	1,905 (6,675)	1,430 (3,705)	24
In 95 CEOs subset	-3,786 (4,358)	-1,952 (2,418)	95
Constant	13,564*** (2,967)	13,534*** (1,647)	
R^2 / Observations	0.21	0.24	397

The executive director effects are extracted from the regressions reported in Table W-3 and transformed into deviations from the mean of all estimated executive director effects. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Figure W-2: Kernel density plots of hospital effects in total pay



(a) Estimates (b) Shrunk estimates

Notes: Kernel density plots of deviations of estimated hospital effects in total pay (CPI adjusted) from mean. Shrunk estimates obtained using empirical Bayes shrinkage estimator.

form the estimates into deviations from the mean of all $\psi_{j(i,t)}$. At 25% of hospitals the executive directors are paid an extra £6,400 or more in total pay, holding our basic set of time-varying hospital characteristics and all time-invariant executive director characteristics constant. Similarly, at 25% of hospitals the executive directors receive pay packages that are £7,000 or more below the average pay package. The interquartile range in hospital (firm) effects is around £13,500.

Figure W-2b shows the distribution of the hospital effects in pay after adjusting the estimates for sampling error using the empirical Bayes procedure, with the estimates shrunk towards the overall mean.[64] The distribution is only slightly narrower, with the interquartile range dropping from £13,500 to £10,000 after shrinkage.

Table W-6 explores the determinants of the hospital effects in pay using linear regressions of the estimated hospital effects on a set of dummy variables indicating time-invariant hospital characteristics. Results in the first major column are for the original effect estimates while results in the second major column are for the hospital effect estimates shrunk towards the overall mean. In general, the results are similar but unsurprisingly the coefficients for the shrunk hospital effect estimates are smaller.

The first specification in Table W-6 explores the impact of the region where the hospital is based on executive director pay. We expect the hospital effects to reflect differences in the cost of living across the different regions in England. The omitted region being the North West, the constant of $-$£ 6,815 for the original effect estimates is the North West

[64] We employ the empirical Bayes shrinkage estimator described in Chandra et al. (2016).

Table W-6: Association between estimated hospital effects in pay and time-invariant hospital characteristics

	Original effect estimates		Shrunk effect estimates		Obs. in each category
North West (omitted category)					34
North East	9,141**	7,509**	6,558**	5,442**	9
	(4,119)	(3,789)	(2,829)	(2,566)	
Yorkshire and Humber	7,300**	5,637**	4,469**	3,294*	20
	(3,097)	(2,863)	(2,127)	(1,939)	
West Midlands	5,346*	6,105**	3,209	3,669*	22
	(3,007)	(2,808)	(2,065)	(1,902)	
East Midlands	6,140	2,029	3,799	725	10
	(3,953)	(3,699)	(2,715)	(2,505)	
East of England	5,260	4,840	3,502	3,192	18
	(3,203)	(2,942)	(2,200)	(1,992)	
London	14,450***	12,946***	10,026***	8,902***	38
	(2,594)	(2,398)	(1,782)	(1,624)	
South West	2,816	2,461	1,972	1,762	19
	(3,148)	(2,884)	(2,162)	(1,953)	
South East	8,889***	9,244***	5,939***	6,152***	26
	(2,863)	(2,657)	(1,966)	(1,800)	
Major teaching hospital		10,878***		8,330***	26
		(2,201)		(1,490)	
Minor teaching hospital		5,418**		3,677***	25
		(2,281)		(1,544)	
Specialist acute		-5,539*		-3,489***	12
		(3,122)		(2,114)	
Specialist orthopaedic		-16,007***		-10,770***	4
		(5,189)		(3,514)	
Constant	-6,815***	-7,596***	-4,351***	-4,977***	
	(1,885)	(1,813)	(1,294)	(1,228)	
R^2	0.16	0.32	0.17	0.34	

The hospital effects are extracted from the regressions reported in Table W-3 and transformed into deviations from the mean of all estimated hospital effects. For the shrunk effect estimates column, the deviations are then shrunk towards the overall mean using the empirical Bayes shrinkage estimator described in Chandra et al. (2016). A major teaching hospital serves a medical school as their main NHS partner, a minor teaching hospital is only a member of the Association of UK University Hospitals. 196 observations in each regression. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

average of the deviations from the mean of all estimated hospital effects for total pay. As the coefficient estimates for all other regions are positive, the North West is the region with the lowest hospital effects. The regions with the largest hospital effects are London and the South East, which reflects the higher cost of living in these regions.

The ranking of the coefficients for the remaining regions does not reflect the ranking of the cost of living. The North East dummy, the Yorkshire and Humber dummy and the East Midlands dummy have the next largest coefficients, while the coefficient on the South West dummy is small and not statistically significantly different from zero. Once we add dummy variables indicating whether a hospital is a major teaching hospital, a minor teaching hospital, a specialist acute hospital or a specialist orthopaedic hospital the coefficients on the North East dummy, the Yorkshire and Humber dummy and the East Midlands dummy drop by £2,000 to £4,000, suggesting some of these unexpectedly large hospital effects are driven by the teaching status and specialist status of hospitals in these regions. However, average hospital effects for hospitals in the North East, Yorkshire and Humber and the West Midlands are still larger than for hospitals in the South West and the East of England, which tend to have higher cost of living.

Potentially, factors other than the cost of living drive these regional differences in the estimated hospital effects. For example, hospitals in the North East might have more difficulties in attracting and retaining good managers than hospitals in the South West and therefore have to offer a pay premium.

Hospital effects at teaching hospitals are statistically significantly larger than at non-teaching hospitals while hospital effects at specialist hospitals are statistically significantly smaller than at general hospitals. Combining the two largest coefficient estimates, the hospital effect of a major teaching hospital in London is on average $-£\,7{,}596 + £\,12{,}946 + £\,10{,}878 = £\,16{,}228$ above the sample average of the hospital effects in total pay, well above the 75$^{\text{th}}$ percentile of the distribution of deviations from the mean of all hospital effects (£6,400) displayed in Figure W-2.

Overall, the region dummies, the teaching status dummies and the specialist status dummies jointly explain around 30% of the variation in the hospital effects.

W-3 Predictors of CEO turnover

Table W-7: Association between number of CEOs observed per hospital and time-invariant hospital characteristics

	Coefficient	Obs. in category
North West	1.29** (0.59)	29
North East (omitted category)		
Yorkshire and Humber	1.43** (0.64)	15
West Midlands	1.77*** (0.62)	19
East Midlands	3.20*** (0.75)	7
East of England	1.86*** (0.62)	18
London	1.53*** (0.59)	27
South West	1.59** (0.62)	18
South East	0.95 (0.61)	21
Major teaching hospital	-0.16 (0.36)	19
Minor teaching hospital	-1.11*** (0.35)	23
Specialist acute	-0.47 (0.46)	12
Specialist orthopaedic	-0.67 (0.76)	4
Constant	2.41*** (0.52)	
R^2/Observations	0.20	162

A major teaching hospital serves a medical school as their main NHS partner, a minor teaching hospital is only a member of the Association of UK University Hospitals. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-8: Linear probability model of CEO being one of the 95 CEOs included in two-step approach estimation sample

	Coefficient	Obs. in category
Female	0.003 (0.043)	147
Clinical background	0.019 (0.047)	112
Postgraduate management qualification	0.070* (0.043)	121
Constant	0.18*** (0.025)	
R^2/Observations	0.007	469

Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

W-4 Fixed effects approach to estimating impact of CEOs on hospital production and placebo experiments

The fixed effects approach exploits movement of the same CEO across different hospitals. It involves estimating regressions of the following form:

$$y_{jt} = \boldsymbol{X}'_{jt}\boldsymbol{\beta} + \lambda_t + \psi_j + \alpha_{i(j,t)} + \varepsilon_{jt} \tag{W-2}$$

The left-hand side variable, y_{jt}, is a production measure of hospital j in financial year t. The function $i(j,t)$ maps hospital j to CEO i in financial year t. $\boldsymbol{X}_{jt}$ is a vector of the following time-varying observable hospital characteristics: foundation trust status, year of merger, years since merger, beds, technology index, case mix measures (patients aged 0 to 14, patients aged 60 to 74, patients aged 75+, male patients). We also include a full set of financial year effects, λ_t, which non-parametrically controls for trends in hospital performance that are national in scope while a full set of hospital effects, ψ_j, controls for non-time varying unobserved differences between hospitals. The estimates of interest are the CEO effects $\alpha_{i(j,t)}$. ε_{jt} represents the error term. We estimate standard errors that are clustered at hospital level.

Following Bertrand & Schoar (2003), we estimate CEO effects $\alpha_{i(j,t)}$ only for the subset of CEOs observed in two hospitals for at least two years each. This sampling requirement means effects that matter require that corporate practices be correlated across two hospitals when the same CEO is present and gives CEOs time to "make their mark" (Bertrand & Schoar 2003). For CEOs observed in three or four hospitals for at least two years in each, we use only the two most recent spells to be comparable with the other CEOs. As Bertrand & Schoar (2003) discuss, the estimated CEO effect for a CEO observed in only one hospital for only part of the time for which that hospital is observed would be identified but would be a period-hospital-specific effect rather than a CEO effect.

There are two further issues when determining the CEOs for whom we estimate CEO effects. One, some CEOs are observed in a hospital for two years but they served for only part of each of these two years. We define these observations as not complying with our requirement of being observed for at least two years. Two, the CEO effect for a CEO observed in one hospital for the exact same time period we observe the hospital for would be perfectly collinear with the hospital effect ψ_j. Therefore we ignore such observations when determining which CEO observations meet our estimation requirements.

Bertrand & Schoar (2003) present F-statistics from tests of the joint significance of the estimated CEO effects to assess the statistical significance of these estimates. However, the estimated CEO effects are essentially the mean of the residuals of a regression of y_{jt}

on $\boldsymbol{X}_{jt}$, λ_t and ψ_j over the observations of the hospitals the CEO has been observed in, for the financial years the CEO has been observed there. Thus, a large residual in one hospital might result in a large mean residual and therefore a CEO effect estimate $\hat{\alpha}_{i(j,t)}$ that is statistically significantly different from zero as a consequence of a period-hospital-specific effect rather than as a consequence of a persistent CEO effect.

We conduct placebo experiments using random assignment of CEOs to hospitals to assess the validity of F-tests on CEO effects. Our starting point for the random assignment are those CEO spells used in Equation W-2. For example, a CEO might be observed at Hospital A from 2001/02-2004/05 and at Hospital B from 2005/06-2008/09. We randomly assign this CEO to a hospital for the period 2001/02-2004/05 and to another hospital for 2005/06-2008/09. The set of hospitals used in the random assignment is the same set as that used in the non-random assignment estimates. To ensure that each hospital is assigned to only one CEO at a time, we sample hospitals without replacement and remove a hospital that has been assigned to a CEO spell from the pool for the duration of the CEO spell it has been assigned to. We then estimate Equation W-2 for the sample with the random CEO-hospital matches $i(j,t)$, test the joint significance of the CEO effects using an F-test and count the number of CEO effects that are individually statistically significant. We repeat this process 100 times and compare the means over the 100 replications to the values obtained using the actual CEO-hospital matches $i(j,t)$.

In addition to the significance testing proposed by Bertrand & Schoar (2003), we examine the proportion of the variance in the hospital production measures, y_{jt}, that is explained by the covariates, $\boldsymbol{X}'_{jt}\boldsymbol{\beta}$ and λ_t, the hospital effects, ψ_j, and the CEO effects, $\alpha_{i(j,t)}$. The proportion of the variance explained by the hospital effects and the CEO effects are $[\text{Cov}(y_{jt}, \hat{\psi}_j)/\text{Var}(y_{jt})] \times 100$ and $[\text{Cov}(y_{jt}, \hat{\alpha}_{i(j,t)})/\text{Var}(y_{jt})] \times 100$, respectively. To obtain the proportion explained by the covariates, we calculate $\hat{y}_{jt} = \boldsymbol{X}'_{j(i,t)t}\hat{\boldsymbol{\beta}} + \hat{\gamma} + \hat{\lambda}_t$ and use this prediction to calculate the covariance: $[\text{Cov}(y_{jt}, \hat{y}_{jt})/\text{Var}(y_{jt})] \times 100$. When conducting the placebo experiments using random assignment of CEOs, we also calculate the variance proportions for each random draw and compare the means over the 100 replications to the the values obtained using the actual CEO-hospital matches $i(j,t)$.

Table W-9 presents the results for financial surplus and staff job satisfaction, Table W-10 the results for the input measures, Table W-11 the results for the throughput measures and Table W-12 the results for the clinical performance measures. Results for the actual CEO-hospital matches are presented in the first row of each panel. At around 0.70 to 0.90, the R^2 (Column 3) is high for most of the hospital production measures. The R^2 values are lowest for the clinical performance measures and the financial surplus variable, with values ranging from 0.30 to 0.50. Even at this values, it seems that the hospital effects, the CEO effects, the financial year effects and our measures of time-varying hospital characteristics jointly explain a large proportion of the variation in the hospital production measures. The F-tests (Column 1) suggest that the estimated CEO effects are jointly statistically significantly different from zero for all our hospital production measures. The proportion of CEO effects that are individually significantly different from zero (Column 2) varies from 24.2% for surplus to 34.7% for one of our input measures.

Columns 5 to 8 present, for the subsample of hospital-year observations with at least one CEO effect $\alpha_{i(j,t)}$ (i.e. hospital-year observations when at least one of the 95 CEOs is present), the proportion of the variance in the hospital production measures that is explained by each term in Equation W-2: the covariates (time-varying hospital characteristics + year effects), the hospital effects, the CEO effects and the residuals. For some of the measures the variance proportions explained by the covariates, hospital effects and CEO effects are invalid due to one of the proportions being negative. For the hospital production measures with valid variance proportions a considerable proportion of the variance is accounted for by the observed covariates and the hospital effects, with the exception of surplus. The CEO effects, despite being jointly statistically significant as measured by the F-test, explain only around 10% of the variance in the hospital production measures. Again, surplus is the exception, as the variance proportion explained by the covariates and the variance proportion explained by the hospital effects is less than the variance proportion explained by the CEO effects.

While these results suggest the existence of statistically significant CEO effects, a deeper look at the data suggests otherwise. From the random CEO-hospital matches reported in the second and third row of each panel, it is clear the means of the F-statistics across the 100 replications are as large as for the actual CEO-hospital matches. The F-test rejects the null hypothesis of the randomly generated CEO effects jointly being equal to zero for every one of the 100 replications, a rejection frequency of 100% at a nominal significance level of 1%. Similarly, the mean of the proportion of CEO effects that are individually statistically significantly different from zero is very similar to that for the actual CEO-hospital matches. Finally, the mean variance proportion explained by the CEO effects when CEOs are randomly assigned to hospitals is very similar to the variance proportions explained by the CEO effects using the actual assignments. The largest difference is for surplus, but even for this measure the ratio of variance explained by the CEO effect to the residual is similar across actual and random matches.

Overall, our placebo experiments suggest that the F-tests on the CEO effects are not valid. The CEO effect estimates, and therefore the F-tests, may be capturing period-hospital-specific shocks rather than true CEO effects. Our findings mirror the results in

Table W-9: Estimates of CEO effects for surplus and staff job satisfaction using actual CEO-hospital matches as well as random CEO-hospital matches

	F-test of joint significance of CEO effects (p-value/rejection frequency using 1% signif. level, df1, df2) (1)	Number (prop.) of CEO effects statist. signif. at 5% (2)	R^2 (3)	Total hospital-year obs. (4)	Variance proportions (%) for subsample of obs. with at least one non-zero CEO effect: Co-variates (5)	Hospital effects (6)	CEO effects (7)	Re-siduals (8)	Subsample hospital-year obs. (9)
Surplus									
Actual matches	32.6 (<0.001, 95, 224)	23 (24.2%)	0.31	2,396	4.7	13.9	19.5	61.9	830
Random matches:									
Means	44.6 (100%, 93.8, 224)	23.9 (25.5%)	0.29	2,396	4.9	17.8	13.8	63.5	843.6
(Std. dev.)	(36.9) (n.a., 1.06, 0)	(4.27, 4.56)	(0.01)		(1.42)	(2.99)	(3.42)	(3.16)	(12.1)
Staff job satisfaction									
Actual matches	14.85 (<0.001, 73, 176)	24 (32.9%)	0.76	1,838	44.7	24.1	5.3	25.9	609
Random matches:									
Means	48.8 (100%, 72.9, 176)	23.7 (32.6%)	0.77	1,838	42.2	29.0	7.0	21.7	623.7
(Std. dev.)	(33.9) (n.a., 1.12, 0)	(4.68, 6.46)	(0.004)		(2.32)	(3.36)	(2.49)	(1.67)	(11.0)

df = degrees of freedom. df1 is the number of CEO effects, df2 is the number of hospital clusters. Standard errors used for the statistical significance tests are clustered at hospital level. Variance proportion is the proportion of variance in the outcome variable that is explained by the covariates, the hospital effects, the director effects and the residuals, respectively. Covariates are financial year effects, foundation trust status, year of merger, years since merger, beds, technology index and case mix variables. The results for random CEO-hospital matches are means and standard deviations across 100 replications.

Table W-10: Estimates of CEO effects for input measures using actual and random CEO-hospital matches

	F-test of joint significance of CEO effects (p-value/rejection frequency using 1% signif. level, df1, df2)	Number (prop.) of CEO effects statist. signif. at 5%	R^2	Total hospital-year obs.	Variance proportions (%) for subsample of obs. with at least one non-zero CEO effect: Co-variates	Hospital effects	CEO effects	Re-siduals	Subsample hospital-year obs.
Doctors + nurses/beds									
Actual matches	32.8 (<0.001, 94, 223)	31 (33.0%)	0.90	2,382	38.2	49.6	4.9	7.4	819
Random matches:									
Means	54.8 (100%, 92.0, 223)	27.2 (29.6%)	0.90	2,382	36.4	51.4	4.9	7.3	826.8
(Std. dev.)	(31.8) (n.a., 1.23, 0)	(4.87, 5.22)	(0.002)		(3.05)	(3.92)	(2.27)	(1.07)	(12.4)
Senior doctors/staff									
Actual matches	54.4 (<0.001, 95, 224)	33 (34.7%)	0.89	2,396	44.9	35.3	3.7	16.1	830
Random matches:									
Means	75.2 (100%, 93.7, 224)	29.9 (31.9%)	0.89	2,396	45.6	38.5	3.2	12.7	842.3
(Std. dev.)	(49.6) (n.a., 1.09, 0)	(5.01, 5.32)	(0.002)		(3.21)	(3.82)	(1.75)	(3.14)	(12.2)
Nurses/staff									
Actual matches	67.9 (<0.001, 95, 224)	27 (28.4%)	0.86	2,396	4.8	75.0	9.1	11.2	830
Random matches:									
Means	70.58 (100%, 93.7, 224)	29.1 (31.1%)	0.86	2,396	4.5	79.8	5.4	10.3	842.3
(Std. dev.)	(54.1) (n.a., 1.09, 0)	(4.92, 5.23)	(0.002)		(2.03)	(2.95)	(2.15)	(1.32)	(12.2)
Technology									
Actual matches	256.29 (<0.001, 95, 224)	33 (33.7%)	0.95	2,398	*4.1*	*91.1*	*-0.04*	*4.8*	830
Random matches									
Means	96.4 (100%, 93.8, 224)	33.1 (35.3%)	0.0.95	2,398	4.4	86.9	3.8	4.8	843.6
(Std. dev.)	(71.0) (n.a., 1.09, 0)	(4.97, 5.23)	(0.001)		(1.59)	(1.86)	(1.62)	(0.78)	(12.1)

See notes for Taable W-9.

Table W-11: Estimates of CEO effects for throughput measures using actual and random CEO-hospital matches

	F-test of joint significance of CEO effects (p-value/rejection frequency using 1% signif. level, df1, df2)	Number (prop.) of CEO effects statist. signif. at 5%	R^2	Total hospital-year obs.	Variance proportions (%) for subsample of obs. with at least one non-zero CEO effect: Co-variates	Hospital effects	CEO effects	Re-siduals	Subsample hospital-year obs.
Admissions									
Actual matches	32.4 (<0.001, 95, 224)	30 (31.6%)	0.98	2,392	*25.1*	*73.1*	*-0.39*	*2.2*	826
Random matches:									
Means	70.2 (100%, 93.6, 224)	28.7 (30.6%)	0.98	2,392	20.4	74.6	2.7	2.3	839.3
(Std. dev.)	(43.1) (n.a., 1.16, 0)	(4.69, 5.00)	(0.0005)		(2.02)	(2.32)	(1.23)	(0.54)	(12.5)
Length of stay									
Actual matches	45.5 (<0.001, 94, 224)	31 (33.0%)	0.95	2,386	48.5	38.5	0.5	12.5	815
Random matches:									
Means	58.5 (100%, 92.9, 224)	26.0 (28.0%)	0.95	2,386	47.8	40.3	1.6	10.3	831.8
(Std. dev.)	(40.1) (n.a., 1.32, 0)	(4.90, 5.23)	(0.001)		(9.16)	(8.26)	(2.72)	(2.36)	(12.8)
Day cases									
Actual matches	100.3 (<0.001, 95, 223)	33 (34.7%)	0.86	2,383	27.8	47.9	13.3	10.9	824
Random matches:									
Means	64.2 (100%, 93.4, 223)	28.8 (30.9%)	0.85	2,383	26.1	55.4	7.8	10.8	837.1
(Std. dev.)	(36.9) (n.a., 1.21, 0)	(4.42, 4.68)	(0.003)		(5.94)	(6.80)	(3.18)	(1.73)	(12.7)
Waiting time									
Actual matches	61.2 (<0.001, 93, 223)	29 (31.2%)	0.84	2,356	52.1	25.1	7.8	15	804
Random matches									
Means	79.2 (100%, 91.7, 223)	29.9 (32.6%)	0.84	2,356	53.7	28.0	6.1	12.3	815.8
(Std. dev.)	(62.5) (n.a., 1.64, 0)	(4.29, 4.67)	(0.003)		(2.71)	(3.21)	(2.40)	(1.54)	(15.5)
Cancelled operations									
Actual matches	77.0 (<0.001, 90, 199)	25 (27.8%)	0.73	2,332	*-4.8*	*78.1*	*2.6*	*24.1*	786
Random matches:									
Means	66.4 (100%, 90.4, 199)	25.9 (28.6%)	0.73	2,332	0.11	68.6	9.2	22.0	813.0
(Std. dev.)	(69.5) (n.a., 1.53, 0)	(4.80, 5.29)	(0.005)		(4.30)	(6.09)	(3.90)	(2.01)	(15.7)

See notes for Table W-9.

Table W-12: Estimates of CEO effects for clinical performance measures using actual and random CEO-hospital matches

	F-test of joint significance of CEO effects (p-value/rejection frequency using 1% signif. level, df1, df2)	Number (prop.) of CEO effects statist. signif. at 5%	R^2	Total hospital-year obs.	Variance proportions (%) for subsample of obs. with at least one non-zero CEO effect: Co-variates	Hospital effects	CEO effects	Re-siduals	Subsample hospital-year obs.
AMI deaths									
Actual matches	23.6 (<0.001, 61, 200)	18 (29.5%)	0.48	1,757	21.5	27.6	5.5	45.4	490
Random matches:									
Means	28.4 (100%, 53.4, 200)	16.9 (31.7%)	0.48	1,757	17.9	18.1	12.5	51.5	430.8
(Std. dev.)	(23.0) (n.a., 3.25, 0)	(3.65, 6.72)	(0.005)		(2.90)	(4.24)	(3.73)	(4.40)	(25.9)
Stroke deaths									
Actual matches	25.1 (<0.001, 72, 200)	26 (36.1%)	0.68	1,965	40.1	24.0	9.1	26.8	596
Random matches:									
Means	38.7 (100%, 64.8, 200)	21.1 (32.6%)	0.68	1,965	40.7	20.2	7.3	31.8	552.2
(Std. dev.)	(26.0) (n.a., 2.30, 0)	(3.98, 6.03)	(0.003)		(2.62)	(2.70)	(2.49)	(2.47)	(20.9)
FPF deaths									
Actual matches	23.9 (<0.001, 72, 195)	20 (27.8%)	0.48	1,920	20.9	16.5	10.9	51.7	588
Random matches:									
Means	32.2 (100%, 64.3, 195)	20.4 (31.7%)	0.49	1,920	21.3	17.8	11.2	49.7	544.3
(Std. dev.)	(20.8) (n.a., 2.33, 0)	(3.53, 5.56)	(0.005)		(1.60)	(2.49)	(3.20)	(2.96)	(21.2)
Readmissions									
Actual matches	30.2 (<0.001, 78, 222)	25 (32.0%)	0.78	2,070	39.6	27.0	12.8	20.5	636
Random matches:									
Means	38.2 (100%, 71.0, 222)	23.2 (32.6%)	0.78	2,070	26.9	39.4	9.7	24.0	583.3
(Std. dev.)	(31.2) (n.a., 1.44, 0)	(4.26, 5.93)	(0.004)		(7.04)	(5.91)	(3.14)	(3.37)	(13.1)
MRSA rate									
Actual matches	34.5 (<0.001, 80, 165)	20 (25.0%)	0.77	2,055	54.8	19.6	6.3	19.3	684
Random matches:									
Means	61.8 (100%, 85.5, 165)	26.9 (31.5%)	0.77	2,055	53.3	22.5	5.9	18.3	748.3
(Std. dev.)	(54.1) (n.a., 1.64, 0)	(4.10, 4.68)	(0.004)		(2.46)	(2.46)	(2.30)	(1.58)	(16.2)

See notes for Table W-9.

Table W-13: Association of (1) means of residuals for CEO spells in first and second hospital and (2) pre-assignment trend and mean of residuals for CEO spell in second hospital using actual CEO-hospital matches as well as random CEO-hospital matches

	1			2		
	CEO spells at 1st and 2nd hospital			Pre-assignment trend and CEO spell at 2nd hospital		
	Coefficient (std. error)	p-value/rejection freq. using 10% signif. level	Obs.	Coefficient (std. error)	p-value/rejection freq. using 10% signif. level	Obs.
Surplus						
Actual matches	-0.05 (0.30)	0.87	95	0.16 (0.22)	0.47	92
Random matches:						
Means	0.003 (0.14)	10%	93.8	0.003 (0.13)	12%	92.9
(Std. dev.)	(0.14, 0.04)		(1.09)	(0.14, 0.03)		(1.48)
Waiting times						
Actual matches	-0.01 (0.08)	0.93	93	0.01 (0.08)	0.86	90
Random matches:						
Means	0.004 (0.10)	9%	91.7	0.02 (0.12)	20%	90.0
(Std. dev.)	(0.10, 0.01)		(1.64)	(0.14, 0.02)		(2.03)
Day cases						
Actual matches	0.18* (0.09)	0.06	95	0.19** (0.10)	0.05	92
Random matches:						
Means	0.003 (0.11)	7%	93.4	0.01 (0.11)	17%	92.6
(Std. dev.)	(0.11, 0.02)		(1.21)	(0.13, 0.02)		(1.62)
Length of stay						
Actual matches	0.05 (0.06)	0.47	94	-0.04 (0.09)	0.68	91
Random matches:						
Means	-0.001 (0.09)	7%	92.9	0.01 (0.11)	20%	92.0
(Std. dev.)	(0.09, 0.02)		(1.32)	(0.15, 0.03)		(1.65)
MRSA rate						
Actual matches	0.10 (0.10)	0.33	80	-0.05 (0.12)	0.71	78
Random matches:						
Means	-0.003 (0.11)	11%	85.5	-0.02 (0.12)	17%	84.7
(Std. dev.)	(0.11, 0.02)		(1.64)	(0.14, 0.02)		(1.90)
Staff job satisfaction						
Actual matches	-0.07 (0.11)	0.56	73	-0.11 (0.17)	0.50	73
Random matches:						
Means	-0.003 (0.14)	6%	72.9	-0.03 (0.14)	7%	72.7
(Std. dev.)	(0.13, 0.02)		(1.12)	(0.13, 0.02)		(1.20)

Results in Column 1 are from regressions of mean of residuals for CEO spell in second hospital on mean of residuals for CEO spell in first hospital. Results in Column 2 are from regressions of mean of residuals for CEO spell in second hospital on mean of residuals in second hospital during the three years before CEO was appointed. Residuals are from regressions of hospital production measures on hospital characteristics, financial year effects and hospital effects. Results for random CEO-hospital matches are means and standard deviations across 100 replications. *Significant at 10%, **significant at 5%, ***significant at 1%

Fee et al. (2013). Using data similar to that in Bertrand & Schoar (2003), they estimate statistically significant CEO fixed effects even when they randomly assign each CEO (observed at two firms) to a second firm other than the one they actually joined.

The placebo experiments also allow us to assess the validity of our preferred two-step procedure. Specifically, we estimate Equations 1 and 2 for both actual CEO-hospital matches and random CEO-hospital matches. Table W-13 presents the results. Looking at the results for the regressions using random CEO-hospital matches in the second and third row of each panel, we see that the coefficient estimates $\hat{\delta}_2$ are very small, with the mean coefficient estimates across the 100 replications ranging from -0.003 to 0.004. The proportion of t-tests across our 100 replications that reject the hypothesis that δ_2 is equal to zero when using a significance level of 10% ranges from 6% to 11%, which is close to the nominal level of the test. Overall, the results for random CEO-hospital matches show no impact of CEOs, exactly what we would expect for random matches, suggesting that the two-step procedure is valid.

W-5 Non-parametric estimates: Additional results

Table W-14: Changes in hospital production measures following a CEO turnover event compared to *three* matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Surplus	Treated	205	0.096 (0.072)		
	Controls	596	0.007 (0.048)	0.089 (0.092)	0.33
Waiting time	Treated	200	0.328 (0.043)		
	Controls	583	0.300 (0.022)	0.028 (0.045)	0.53
Day cases	Treated	202	0.109 (0.030)		
	Controls	586	0.134 (0.021)	-0.026 (0.040)	0.53
Length of stay	Treated	205	0.167 (0.023)		
	Controls	596	0.112 (0.010)	0.056 (0.022)	0.01
MRSA rate	Treated	197	0.262 (0.048)		
	Controls	572	0.280 (0.029)	-0.018 (0.057)	0.75
Staff job satisfaction	Treated	163	0.124 (0.076)		
	Controls	468	0.247 (0.042)	-0.123 (0.084)	0.14

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. Up to three controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t-1$ and $t-2$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status, followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-18 presents non-parametric results for the subset of hospitals for whom we observe an average management score in both the 2006 and the 2009 wave of the World Management Survey. Thus, we can include hospitals with a CEO turnover event in 2007 or 2008. There are only 9 treated observations, so the effect estimate is very imprecise. However, there is no indication of a CEO turnover event improving management practices. If anything, a turnover event decreases the average management score. The table also presents estimates of the impact of a CEO turnover event on how much hospitals spend on CEO remuneration. The estimates suggest that as a result of a CEO turnover event hospitals' spending on CEO remuneration increases by about £7,500 more than it would have done in the absence of a turnover event. However, the last panel of Table W-18 shows that the parallel trend assumption for hospital spending on CEO pay is unlikely to

Table W-15: Changes in hospital production measures *before* the CEO turnover events compared to matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Surplus	Treated	183	-0.172 (0.090)		
	Controls	184	-0.136 (0.066)	-0.036 (0.111)	0.74
Waiting time	Treated	177	-0.317 (0.043)		
	Controls	178	-0.335 (0.043)	0.018 (0.061)	0.77
Day cases	Treated	179	0.077 (0.031)		
	Controls	182	-0.014 (0.039)	0.091 (0.050)	0.07
Length of stay	Treated	183	-0.123 (0.019)		
	Controls	183	-0.163 (0.019)	0.040 (0.027)	0.13
MRSA rate	Treated	156	-0.322 (0.056)		
	Controls	157	-0.349 (0.055)	0.027 (0.078)	0.73
Staff job satisfaction	Treated	123	0.084 (0.089)		
	Controls	123	0.035 (0.093)	0.049 (0.129)	0.70

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t-1)} - y_{j(t-3)}$. The number of treated observations is less than the number of treated observations in Table W-23 because for some treated observations we do not observe the lagged change in the outcome variable. For details on selection of treated and control observations refer to notes in Table W-23. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-16: Changes in hospital production measures following a CEO turnover event compared to matched control hospitals with no CEO turnover event with potential treated observations limited to CEO spells included in parametric two-step approach (sample of movers)

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Surplus	Treated	106	0.148 (0.133)		
	Controls	106	0.233 (0.092)	-0.085 (0.162)	0.60
Waiting time	Treated	105	-0.344 (0.059)		
	Controls	105	-0.334 (0.052)	-0.010 (0.079)	0.90
Day cases	Treated	105	0.139 (0.042)		
	Controls	105	0.019 (0.042)	0.121 (0.060)	0.04
Length of stay	Treated	106	-0.176 (0.034)		
	Controls	106	-0.132 (0.020)	-0.044 (0.039)	0.27
MRSA rate	Treated	102	-0.320 (0.076)		
	Controls	102	-0.232 (0.067)	-0.088 (0.101)	0.39
Staff job satisfaction	Treated	84	0.097 (0.116)		
	Controls	84	0.393 (0.106)	-0.296 (0.157)	0.06

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. The maximum number of treated observations is less than 95 CEOs x 2 hospitals = 190 for the following reasons: Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. We cannot use observations for 2000/01 and 2001/02 since we cannot establish whether there was no turnover event in $t-1$ and $t-2$. For details on selection of control observations refer to notes in Table W-23. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-17: Changes in hospital production measures over a period of 3 years instead of 2 years following a CEO turnover event compared to matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Surplus	Treated	151	-0.003 (0.102)		
	Controls	151	-0.020 (0.070)	0.017 (0.124)	0.89
Waiting time	Treated	144	-0.604 (0.056)		
	Controls	144	-0.526 (0.056)	-0.078 (0.079)	0.33
Day cases	Treated	150	0.143 (0.046)		
	Controls	150	0.212 (0.056)	-0.069 (0.072)	0.34
Length of stay	Treated	151	-0.244 (0.027)		
	Controls	151	-0.194 (0.020)	-0.050 (0.034)	0.14
MRSA rate	Treated	145	-0.503 (0.062)		
	Controls	145	-0.412 (0.069)	-0.090 (0.093)	0.34
Staff job satisfaction	Treated	114	0.227 (0.112)		
	Controls	114	0.419 (0.107)	-0.191 (0.155)	0.22

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+2)} - y_{j(t-1)}$. Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and $t+2$ and no CEO turnover event in $t-1$ and $t-2$. Controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t+2$, $t-1$ and $t-2$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status. Some treated observations remain without a match. Exact matching is followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

be satisfied, since it increased by about £7,400 less in treated hospitals over the two-year period before the CEO turnover event.

Table W-18: Changes in average management score and hospital spending on CEO remuneration following a CEO turnover event compared to one or three matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Average management score	Treated	9	0.076 (0.162)		
	Controls	9	0.272 (0.243)	-0.195 (0.292)	0.51
	Controls	23	0.315 (0.156)	-0.239 (0.271)	0.39
Hospital spending on CEO remuneration	Treated	175	8,672 (2,140)		
	Controls	175	827 (1,088)	7,845 (2,400)	0.001
	Controls	509	1,225 (616)	7,448 (1,636)	0.00
Changes in spending on CEO remuneration *before* turnover event	Treated	150	2,117 (2,128)		
	Controls	149	9,532 (1,683)	-7,414 (2,716)	0.01
	Controls	427	9,538 (910)	-7,421 (1,987)	0.00

Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. Up to three controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t-1$ and $t-2$. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status, followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. For changes in spending on CEO remuneration *before* turnover event, the change in outcome variable is $y_{j(t-1)} - y_{j(t-3)}$. The number of treated observations is less than the number of treated observations for hospital spending on CEO remuneration because for some treated observations we do not observe the lagged change in hospital spending on CEO remuneration. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

W-6 Results for additional hospital production measures

Table W-19 presents results for the parametric two-step approach. Similar to the results for the key target measures, the coefficient estimates are very small or negative, suggesting no portability of performance. Tables W-20 and W-21 present results for the non-parametric approach for all hospital production measures in their original units.[65] Again, similar to the results for the key target measures, we find very few changes following a CEO turnover event.

[65] The analyses reported in Table 3 for the key target measures use standardised outcome variables.

Table W-19: Association of (1) means of residuals for CEO spells in first and second hospital and (2) pre-assignment trend and mean of residuals for CEO spell in second hospital for additional hospital production measures

	1			2		
	CEO spells at 1st and 2nd hospital			Pre-assignment trend and CEO spell at 2nd hospital		
	Coefficient (std. error)	R^2	Obs.	Coefficient (std. error)	R^2	Obs.
Doctors + nurses/beds	-0.01	0	94	-0.05	0	91
	(0.15)			(0.09)		
Senior doctors/staff	0.03	0	95	-0.08	0.01	92
	(0.12)			(0.11)		
Nurses/staff	0.08	0.01	95	0.10	0.01	92
	(0.10)			(0.11)		
Technology	0.001	0	95	-0.05	0	92
	(0.10)			(0.10)		
Admissions	0.11	0.01	95	-0.005	0	92
	(0.12)			(0.11)		
Cancelled operations	-0.12	0.01	90	0.32	0.03	87
	(0.17)			(0.21)		
AMI deaths	-0.17	0.04	61	-0.01	0	58
	(0.11)			(0.08)		
Stroke deaths	0.001	0	72	0.02	0	69
	(0.10)			(0.12)		
FPF deaths	-0.08	0.01	72	0.01	0	69
	(0.11)			(0.12)		
Readmissions	0.07	0.01	78	0.03	0	75
	(0.10)			(0.10)		

Results in Column 1 are from regressions of mean of residuals for CEO spell in second hospital on mean of residuals for CEO spell in first hospital. Results in Column 2 are from regressions of mean of residuals for CEO spell in second hospital on mean of residuals in second hospital during the three years before CEO was appointed. The residuals are from a regression of the hospital production measures on hospital characteristics, financial year effects and hospital effects. *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-20: Changes in all hospital production measures in their original units following a CEO turnover event compared to matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Input measures:					
Doctors + nurses/beds	Treated	205	0.20 (0.02)		
	Controls	205	0.20 (0.02)	-0.00 (0.03)	0.82
Senior doctors/staff	Treated	205	0.67 (0.13)		
	Controls	205	0.62 (0.10)	0.05 (0.17)	0.76
Nurses/staff	Treated	205	-0.25 (0.12)		
	Controls	205	-0.12 (0.13)	-0.13 (0.17)	0.46
Technology	Treated	205	0.024 (0.005)		
	Controls	205	0.018 (0.004)	0.007 (0.006)	0.27
Throughput measures:					
Admissions	Treated	205	4,216 (404)		
	Controls	205	4,955 (542)	-739 (676)	0.28
Length of stay	Treated	205	-0.48 (0.07)		
	Controls	205	-0.35 (0.04)	-0.13 (0.08)	0.10
Day cases	Treated	202	0.94 (0.26)		
	Controls	202	0.73 (0.31)	0.21 (0.40)	0.60
Waiting time	Treated	200	-9.83 (1.29)		
	Controls	200	-8.72 (1.10)	-1.11 (1.69)	0.51
Cancelled operations	Treated	202	-15.8 (14.5)		
	Controls	202	-3.15 (11.3)	-12.6 (18.4)	0.49

The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. Controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t-1$ and $t-2$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status, followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-21: Changes in all hospital production measures in their original units following a CEO turnover event compared to one or three matched control hospitals with no CEO turnover event - continued

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Clinical performance measures:					
AMI deaths	Treated	143	-0.64 (0.30)		
	Controls	143	-0.54 (0.27)	-0.10 (0.41)	0.80
Stroke deaths	Treated	168	-2.21 (0.30)		
	Controls	168	-1.07 (0.34)	-1.15 (0.45)	0.01
FPF deaths	Treated	165	-0.16 (0.23)		
	Controls	165	-0.38 (0.24)	0.22 (0.33)	0.51
Readmissions	Treated	172	0.54 (0.09)		
	Controls	172	0.50 (0.08)	0.03 (0.12)	0.78
MRSA rate	Treated	197	-2.19 (0.40)		
	Controls	197	-2.30 (0.42)	0.11 (0.58)	0.85
Surplus	Treated	205	1,444 (1,088)		
	Controls	205	2,105 (829)	-661 (1,368)	0.63
Staff job satisfaction	Treated	163	0.013 (0.008)		
	Controls	163	0.032 (0.007)	-0.019 (0.011)	0.07

See notes for Table W-20

W-7 Stacked analyses

We adapt both the parametric approach and the non-parametric approach to take advantage of our large number of hospital production measures by stacking them into sets. Specifically, we stack our four input measures, our five throughput measures and our five clinical performance measures. Table 1 presents descriptive statistics for the components of the stacked sets.

Stacked version of parametric approach The stacked version of our parametric two-step procedure essentially estimates a common coefficient δ_2 for each set of hospital production measures. The stacked equivalent of Equation (1) takes the following form:

$$y_{jt}^{k_s} = \sum_{k_s=1}^{K_s} z_{k_s}[\boldsymbol{X}'_{jt}\boldsymbol{\beta} + \lambda_t + \psi_j] + \varepsilon_{jt}^{k_s} \tag{W-3}$$

The left-hand side variable, $y_{jt}^{k_s}$ is one of the K_s outcome variables in set s. All K_s outcome variables in set s are stacked, so the sample size is approximately K_s times the number of hospital-year observations jt. Each outcome variable is standardised to have mean zero and a standard deviation of one. Furthermore, variables such as waiting times and stroke deaths are multiplied by (-1), so an increase in an outcome variable can be interpreted as an improvement. The additional (compared to Equation 1) variable z_{k_s} is a dummy variable that takes the value one if the left-hand side variable $y_{jt}^{k_s}$ is outcome variable k_s. For each of the K_s outcome variables in set s, we estimate separate coefficients on the time-varying observable hospital characteristics $\boldsymbol{X}_{jt}$, a separate set of financial year effects λ_t and a separate set of hospital effects ψ_j by interacting the hospital characteristics, financial year effects and hospital effects with the dummy variables z_k.

We extract the residuals $e_{jt}^{k_s}$ and for each hospital-year observation jt generate the mean residual across all k_s outcome variables in set s: $\overline{e_{jt}^{s}} = \frac{1}{K_s}\sum_{k_s=1}^{K_s} e_{jt}^{k_s}$.[66] For each CEO with at least two observations for at least two of the outcome variables in set s in both hospitals, we generate the mean of the mean residuals $\overline{e_{jt}^{s}}$ for each of the CEO's spells and regress them on each other as in (2):

$$\frac{1}{n^{i,B}} \sum_{t=t_1^{i,B}}^{t_n^{i,B}} \overline{e_{Bt}^{s}} = \delta_1 + \delta_2 \frac{1}{n^{i,A}} \sum_{t=t_1^{i,A}}^{t_n^{i,A}} \overline{e_{At}^{s}} + \varepsilon_i^s \tag{W-4}$$

[66] In the case of a missing observation for one or two of the K_s outcome variables, the mean residual could be a mean over $K_s - 1$ or $K_s - 2$ outcome variables.

To test for pre-assignment trends we replace the mean of the mean residuals at hospital B during the time the CEO was observed there, $\frac{1}{n^{i,B}} \sum_{t=t_1^{i,B}}^{t_n^{i,B}} \overline{e_{Bt}^s}$, with the mean of the mean residuals $\overline{e_{jt}^s}$ at hospital B during the three financial years before the CEO arrived at the hospital B.

Table W-22: Association of (1) means of residuals for CEO spells in first and second hospital and (2) mean of residuals for CEO spell at first hospital and pre-assignment trend at second hospital

	1			2		
	CEO spells at 1st hospital and 2nd hospital			CEO spell at 1st hospital and pre-assignment trend at 2nd hospital		
	Coefficient (std. error)	R^2	Obs.	Coefficient (std. error)	R^2	Obs.
Input measures	0.055 (0.11)	0	95	0.13 (0.11)	0.01	92
Throughput measures	0.20* (0.11)	0.03	95	0.10 (0.12)	0.01	92
Clinical performance	-0.04 (0.10)	0	82	-0.10 (0.09)	0.02	79

Results in Column 1 are from regressions of mean of residuals for CEO spell in second hospital on mean of residuals for CEO spell in first hospital. Results in Column 2 are from regressions of mean of residuals in second hospital during the three years before CEO was appointed on mean of residuals for CEO spell in first hospital. The residuals are from a regression of the standardised stacked measures on hospital characteristics, financial year effects and hospital effects. The residuals for the 4 or 5 outcome measures in each group are then averaged by hospital-year before the mean of residuals for each CEO spell is being calculated. *Significant at 10%, **significant at 5%, ***significant at 1%

The stacked version of the two-step method has several advantages in our context. Firstly, it tackles the issue of missing observations for some of the hospital production measures.[67] When estimating Equation 2 for each y_{jt} separately we can use only those CEOs for whom we observe outcome y_{jt} at two hospitals for at least two years. The stacked approach in Equation W-4 allows us to relax this requirement. We require CEOs to have at least two observations for at least two of the outcome variables in set s at two different hospitals. This relaxed requirement increases the number of CEOs for whom we can estimate the coefficient δ_2. This increase in sample size gives us more statistical power to detect CEO effects. Secondly, the stacked approach allows for correlations between the outcome variables in a set and reduces potential multiple comparison issues. Thirdly, it simplifies the exposition of the results.

[67]Observations are missing because certain production measures are not relevant for the particular hospital. For example, some specialist hospitals have no admissions for acute myocardial infarction (AMI), so we have no observations on AMI deaths for these hospitals.

Table W-22 presents the results. A positive coefficient indicates that a positive deviation from the expected level of a set of production measures during a CEO's spell at the first hospital is associated with a positive deviation from the expected level of the set of production measures during the CEO's spell at the second hospital. A substantial association would suggest that these deviations can be attributed to the CEO and not to period-hospital-specific effects. The coefficient for input measures is small while the coefficient for clinical performance measures is small and even negative. The coefficient for the set of throughput measures is relatively large and statistically significant at the 10% level. However, the examination of pre-assignment trends in Column 2 of Table W-22 suggests that the positive association could be due to selection rather than CEOs imposing their style.

Stacked version of non-parametric approach The stacked version of our matching difference-in-difference estimator replaces the mean of the difference in a single production measure, $\frac{1}{n}\sum_{j=1}^{n}\left(y_{j(t+1)}-y_{j(t-1)}\right)$, with the mean of the differences in all hospital production measures in a set, $\frac{1}{\sum n^{k_s}}\sum_{k_s=1}^{K_s}\sum_{j=1}^{n^{k_s}}\left(y_{j(t+1)}^{k_s}-y_{j(t-1)}^{k_s}\right)$. We report this mean difference and its standard error for both the treated and the control hospitals and present the difference between the two means as well as the standard error and p-value from a two-sample t-test with equal variance.

Table W-23 presents the results. They suggest that there are no changes in the stacked hospital production measures after a new CEO is employed. Table W-24 assesses the balance of the matched samples, showing little difference between the treated and the control samples. Table W-25 repeats the analysis using 1:3 matching and gives similar results. Table W-26 checks the common trend assumption by examining changes in hospital production in the two-year period preceding the CEO turnover event. Changes observed in treated hospitals are generally similar to changes in control hospitals. Table W-27 uses only those CEO spells included in the parametric two-step approach. The results are similar to those here. Table W-28 examines changes over the three (rather than two) years following the CEO turnover event.

Table W-23: Changes in hospital production measures following a CEO turnover event compared to matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Input measures	Treated	820	0.136 (0.017)		
	Controls	820	0.135 (0.016)	0.001 (0.023)	0.95
Throughput measures	Treated	1,011	0.152 (0.015)		
	Controls	1,011	0.124 (0.014)	0.028 (0.021)	0.18
Clinical performance measures	Treated	845	0.128 (0.033)		
	Controls	845	0.103 (0.032)	0.025 (0.046)	0.58

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. Controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t-1$ and $t-2$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status, followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-24: Means of matching variables and other hospital characteristics for treated and control groups and means of exactly matched hospital characteristics

				Means of variables measured in $t-1$						Means of vars. measured in t		Exactly matched charact. (%)			
			Unique con-		Tech-	Prop. in each category (%)				Year of merger	Years since	Teaching		Spec.	Foun- dation trust
		Obs.	trols	Beds	nology	0-14	60-74	75+	male	(%)	merger	Major	Minor	hosp.	in $t-1$
Inputs	T	820		721	0.35	5.31	8.14	8.29	43.5	0	1.22	12.2	9.3	9.8	26.8
	C	820	161	714	0.39	5.06	8.28	8.23	44.3	0.98	0.82				
Through-	T	1,011		724	0.35	5.29	8.14	8.32	43.5	0	1.22	12.4	9.3	9.3	26.8
puts	C	1,011	163	717	0.39	5.07	8.28	8.26	44.3	0.99	0.84				
Clinical	T	845		770	0.36	5.05	8.21	8.60	43.4	0	1.12	13.8	7.2	4.0	23.0
perform.	C	845	179	756	0.39	5.03	8.10	8.33	43.8	0.24	0.96				

T = Treated, C = Controls

Table W-25: Changes in hospital production measures following a CEO turnover event compared to *three* matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Input measures	Treated	820	0.136 (0.017)		
	Controls	2,384	0.138 (0.010)	-0.002 (0.019)	0.92
Throughput measures	Treated	1,011	0.152 (0.015)		
	Controls	2,941	0.142 (0.009)	0.010 (0.018)	0.57
Clinical performance measures	Treated	845	0.128 (0.033)		
	Controls	2,499	0.103 (0.018)	0.026 (0.036)	0.48

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. Up to three controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t-1$ and $t-2$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status, followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-26: Changes in hospital production measures *before* the CEO turnover events compared to matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Input measures	Treated	732	0.153 (0.019)		
	Controls	736	0.087 (0.014)	0.066 (0.023)	0.01
Throughput measures	Treated	900	0.133 (0.016)		
	Controls	904	0.145 (0.017)	-0.011 (0.023)	0.62
Clinical performance measures	Treated	719	0.134 (0.038)		
	Controls	721	0.141 (0.037)	-0.007 (0.053)	0.89

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t-1)} - y_{j(t-3)}$. The number of treated observations is less than the number of treated observations in Table W-23 because for some treated observations we do not observe the lagged change in the outcome variable. For details on selection of treated and control observations refer to notes in Table W-23. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-27: Changes in hospital production measures following a CEO turnover event compared to matched control hospitals with no CEO turnover event with potential treated observations limited to CEO spells included in parametric two-step approach (sample of movers)

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Input measures	Treated	424	0.147 (0.019)		
	Controls	424	0.111 (0.021)	0.036 (0.028)	0.20
Throughput measures	Treated	524	0.174 (0.021)		
	Controls	524	0.116 (0.019)	0.059 (0.029)	0.04
Clinical performance	Treated	450	0.123 (0.045)		
measures	Controls	450	0.148 (0.043)	-0.025 (0.063)	0.69

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. The maximum number of treated observations is less than 95 CEOs x 2 hospitals x 5 measures in a set = 950 for the following reasons: Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and no CEO turnover event in $t-1$ and $t-2$. We cannot use observations for 2000/01 and 2001/02 since we cannot establish whether there was no turnover event in $t-1$ and $t-2$. For details on selection of control observations refer to notes in Table W-23. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-28: Changes in hospital production measures over a period of 3 years instead of 2 years following a CEO turnover event compared to matched control hospitals with no CEO turnover event

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Input measures	Treated	604	0.241 (0.024)		
	Controls	604	0.187 (0.020)	0.054 (0.031)	0.08
Throughput measures	Treated	741	0.248 (0.022)		
	Controls	741	0.215 (0.021)	0.034 (0.030)	0.26
Clinical performance	Treated	632	0.261 (0.039)		
measures	Controls	632	0.217 (0.040)	0.044 (0.056)	0.44

All outcome variables are standardised to have a mean of zero and a standard deviation of one. The change in outcome variable is $y_{j(t+2)} - y_{j(t-1)}$. Treated observations are hospital-years with a CEO turnover event in t, the new CEO still in post in $t+1$ and $t+2$ and no CEO turnover event in $t-1$ and $t-2$. Controls are chosen from hospital-years with no CEO turnover event in t, $t+1$, $t+2$, $t-1$ and $t-2$. Controls are matched exactly on year, major teaching hospital, minor teaching hospital, specialist hospital and foundation trust status. Some treated observations remain without a match. Exact matching is followed by closest neighbour matching on beds. In case of ties, closest neighbour matching on beds is followed by closest neighbour matching on technology index. Foundation trust status, beds and technology index as of $t-1$; teaching status and specialist status are permanent characteristics. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Stacked version of investigations into potential reasons for lack of persistent CEO effects To examine the possibility of CEOs who were good performers at their first hospital subsequently being hired by a problematic hospital, we define CEO performance using the mean of the mean residuals $\overline{e^s_{jt}}$ from Equation W-3 for the financial years $t_1^{i,A}$ to $t_n^{i,A}$ when CEO i is observed in hospital A. We classify as good performers those CEOs whose mean of the mean residuals is at or above the 75$^{\text{th}}$ percentile. The results are in Table W-29. Out of the 8 coefficients, one is statistically significant and negative, suggesting that good performers in terms of throughputs are *less* likely to subsequently be hired by a hospital that has experienced a merger event.

Table W-30 reports results from our complementary analysis that examines whether good performers move to more prestigious hospitals. There is some indication that good performance in the first hospital is associated with being hired at a more prestigious hospital. Specifically, performing well in terms of throughput measures is positively associated with moving to a teaching hospital.

To investigate whether variability in CEO performance is associated with being at some point matched with a problematic hospital, we use the mean of the mean residuals $\overline{e^s_{jt}}$ from Equation W-3 for both CEO spells and calculate the absolute value of the difference in these two means. Table W-31 presents the results. 3 out of 12 coefficients are statistically significantly different from zero at a 10% significance level. However, only one of these coefficients (for the set of throughput measures) is positive. A positive coefficient would suggest that being at a "problematic" hospital is associated with higher variability in CEO performance. The two negative coefficients (for the set of clinical performance measures) suggest that CEOs who are at some point at a more problematic hospital actually have *lower* variability in their performance across hospitals.

To examine the relationship between CEO performance and leaving the sample after being observed in only one hospital we define CEO performance using the mean of the mean residuals $\overline{e^s_{jt}}$ from Equation W-3 for a CEO's first observed spell. We classify as good performers CEOs with a mean at or above the 75$^{\text{th}}$ percentile and as bad performers CEOs whose mean is at or below the 25$^{\text{th}}$ percentile. Table W-32 presents the results. There is no evidence that managers who perform exceptionally well in their first observed stint as CEO are more likely to exit the sample.

To estimate the importance of match effects, we run the following regression:

$$y_{jt}^{k_s} = \sum_{k_s=1}^{K_s} z_{k_s}[\boldsymbol{X'}_{jt}\boldsymbol{\beta} + \lambda_t + \psi_j] + \delta W_i + \gamma(W_i \times W_j) + \varepsilon_{jt}^{k_s} \qquad \text{(W-5)}$$

$y_{jt}^{k_s}$ is one of the K_s outcome variables in set s and z_{k_s} is a dummy variable that takes the value one of the the left-hand side variable $y_{jt}^{k_s}$ is outcome variable k_s.

Table W-33 presents the results. First, female CEOs have better clinical outcomes in teaching hospitals, but at the expense of lower throughputs (which are more easily observed than clinical outcomes, perhaps explaining the lower pay of female CEOs). Second, CEOs who are doctors are associated with better throughput and clinical performance in teaching hospitals. These hospitals are exactly those settings where it might be expected that clinically trained CEOs perform best. Third, CEOs with private sector experience have better clinical performance when paired with a teaching hospital or a larger hospital. In Table W-34 we repeat this analysis for the subset of CEO spells included in the parametric two-step approach (sample of movers). The results are broadly similar.

To examine the possibility that our finding of a lack of persistent impact of CEOs is due to their tenure being too short, we use the residuals from Equation (W-3), i.e. hospital-year deviations from the expected level of our different hospital production measures. We separately examine all CEO spells in our data set and the subset of CEO spells in the mover sample using the regression:

$$\overline{e_{jt}^{s}} = \alpha + \delta tenure_{i(j,t)} + \gamma tenure\_unsure_{i(j,t)} + \varepsilon_{jt}^{s} \tag{W-6}$$

The dependent variable is $\overline{e_{jt}^{s}} = \frac{1}{K_s}\sum_{k_s=1}^{K_s} e_{jt}^{k_s}$, the mean of residuals across all k_s outcome variables in set s for hospital year jt.

Table W-35 presents the results. For the sample of movers there is some indication that longer tenure is associated with more throughputs and better clinical performance. The effect sizes, however, are very small. For example, the statistically significant coefficient estimate for throughputs suggests that 10 years of tenure would result in these outcomes being only one-sixteenth of a standard deviation above the expected level.

Table W-29: Linear probability models of the impact of good performance in first hospital on moving to a "problematic" hospital

	Hospital commission rating poor in year before CEO arrived			Hospital with surplus below 25th percentile in year before CEO arrived			'New' hospital created through merger during sample period			Hospital with PFI contract at some point during CEO's tenure		
	Good perf.	Const.	N	Good perf.	Const.	N	Good perf.	Const.	N	Good perf.	Const.	N
Throughput measures	0.03	0.42	71	-0.04	0.47	91	-0.20*	0.36	95	-0.12	0.45	95
	(0.14)	(0.07)		(0.12)	(0.06)		(0.11)	(0.05)		(0.12)	(0.06)	
Clinical performance measures	0.02	0.42	70	-0.19	0.53	78	-0.19	0.38	82	-0.21	0.49	82
	(0.14)	(0.07)		(0.13)	(0.07)		(0.12)	(0.06)		(0.13)	(0.06)	

Each entry in this table refers to a separate regression of an indicator of a CEO moving to a "problematic" hospital on an indicator of good performance at the CEO's first hospital. "Problematic" hospital is defined as either poor hospital commission rating, surplus below 25^{th} percentile, hospital created through merger or hospital with PFI contract. The 25^{th} percentile of surplus is calculated separately for each financial year to ensure the categorisation is net of year effects. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-30: Linear probability models of the impact of good performance in first hospital on moving to a "prestigious" hospital

	Teaching hospital			Foundation trust			Hospital with beds above 75^{th} percentile before CEO arrived		
	Good perform.	Const.	N	Good perform.	Const.	N	Good perform.	Const.	N
Throughput measures	0.19* (0.10)	0.23 (0.05)	95	0.04 (0.11)	0.30 (0.06)	95	0.06 (0.12)	0.33 (0.06)	90
Clinical performance	-0.06 (0.11)	0.30 (0.06)	82	-0.03 (0.12)	0.31 (0.06)	82	-0.13 (0.12)	0.41 (0.07)	77

Each entry in this table refers to a separate regression of an indicator of a CEO moving to a "prestigious" hospital on an indicator of good performance at the CEO's first hospital. "Prestigious" hospital is defined as either teaching hospital (major or minor), foundation trust status or number of beds above 75^{th} percentile in the year before the CEO arrived. The 75^{th} percentile of number of beds is calculated separately for each financial year to ensure the categorisation is net of year effects. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-31: Impact of ever being observed at a "problematic" hospital on variability in CEO performance as measured by the absolute difference in the mean residuals for the CEO spells at each of their two hospitals for each production measure for stacked hospital production measures

	Hospital commission rating poor in year before CEO arrived	Hospital with surplus below 25th percentile in year before CEO arrived	'New' hospital created through merger during sample period	Hospital with PFI contract at some point during CEO's tenure	Mean (st. dev.) [obs.] of dependent variable (Absolute difference in mean residuals at both hospitals)
Input measures	0.03 (0.02) [89]	0.01 (0.02) [94]	-0.02 (0.02) [95]	-0.03 (0.02) [95]	0.14 (0.11) [95]
Throughput measures	0.003 (0.02) [89]	0.02 (0.02) [94]	0.05** (0.02) [95]	0.02 (0.02)	0.13 (0.10) [95]
Clinical performance measures	-0.005 (0.04) [79]	-0.07 (0.05) [81]	-0.09** (0.04) [82]	-0.08* (0.04) [82]	0.20 (0.19) [82]

Each entry in this table refers to a separate regression of a performance variability measure on a dummy variable indicating that the CEO has ever been observed at a "problematic" hospital defined as either poor hospital commission rating, surplus below 25^{th} percentile, hospital created through merger or hospital with PFI contract. The 25^{th} percentile of surplus is calculated separately for each financial year to ensure the categorisation is net of year effects. Standard errors in (parentheses) and number of observations in [brackets]. *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-32: Linear probability models of the impact of good performance or bad performance on exiting the sample

	Good performance	Bad performance	Constant	N
Throughput measures	0.028 (0.062)	-0.057 (0.064)	0.61 (0.04)	359
Clinical performance	0.056 (0.065)	0.074 (0.063)	0.57 (0.04)	356

Each row in this table refers to a separate regression. The dependent variable is a dummy taking the value 1 if a CEO leaves the sample after being observed in only one hospital and the value 0 if a CEO is observed in at least two different hospitals. The dummy variable is set to missing for CEOs observed at only one hospital if this spell includes the year 2013, as 2013 is the last year in our sample. The independent variables are an indicator of good performance and an indicator of bad performance at the CEO's first or only hospital. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-33: Estimates of quality of CEO-hospital matches for all observed CEO spells

	Input measures	Throughput measures	Clinical performance measures
Female * teaching hospital	−0.02	−0.27***	0.39**
	(0.06)	(0.04)	(0.16)
Female * foundation hospital	−0.01	−0.13***	0.03
	(0.04)	(0.04)	(0.05)
Female * competitive	0.00	0.02	−0.04
	(0.05)	(0.04)	(0.05)
Female * beds (100s)	−0.01	−0.00	0.01**
	(0.00)	(0.01)	(0.01)
Doctor * teaching hospital	−0.02	0.16**	0.16
	(0.06)	(0.08)	(0.13)
Doctor * foundation	0.08	−0.02	−0.01
	(0.06)	(0.06)	(0.09)
Doctor * competitive	−0.11	−0.06	−0.15
	(0.08)	(0.07)	(0.09)
Doctor * beds (100s)	−0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
Private sector * teaching hospital	−0.04	−0.14	0.22**
	(0.05)	(0.11)	(0.09)
Private sector * foundation hospital	0.01	0.04	0.02
	(0.06)	(0.05)	(0.09)
Private sector * competitive	−0.14	−0.05	−0.14
	(0.20)	(0.05)	(0.10)
Private sector * beds (100s)	−0.00	−0.01	0.02***
	(0.01)	(0.01)	(0.01)
Observations	9,573	11,849	9,767

Each estimate is from a separate regression of the stacked measures on hospital characteristics, financial year effects, hospital effects, the relevant CEO characteristic and the interaction of the relevant CEO characteristic and hospital characteristic. All outcome variables in a stacked set as well as the individual outcome variables are standardised to have a mean of zero and a standard deviation of one. Some outcome variables in the stacked sets are multiplied by (-1), so that "more" means "better". Standard errors in (parentheses), clustered at hospital level. *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-34: Estimates of quality of CEO-hospital matches for subset of CEO spells included in parametric two-step approach (sample of movers)

	Input measures	Throughput measures	Clinical performance measures
Female * teaching hospital	−0.06	−0.26***	0.17***
	(0.09)	(0.04)	(0.04)
Female * foundation hospital	0.07	−0.06	0.05
	(0.06)	(0.04)	(0.06)
Female * competitive	−0.03	0.04	0.10
	(0.06)	(0.06)	(0.06)
Female * beds (100s)	−0.01	−0.01	0.01*
	(0.01)	(0.01)	(0.01)
Doctor * teaching hospital	0.11	0.29***	0.32***
	(0.09)	(0.07)	(0.09)
Doctor * foundation	0.12	−0.05	0.04
	(0.09)	(0.09)	(0.09)
Doctor * competitive	−0.08	−0.06	−0.15
	(0.06)	(0.14)	(0.12)
Doctor * beds (100s)	0.01	−0.00	0.02
	(0.01)	(0.02)	(0.02)
Private sector * teaching hospital	−0.07	−0.03	0.25**
	(0.06)	(0.12)	(0.09)
Private sector * foundation hospital	0.03	0.08	−0.01
	(0.09)	(0.09)	(0.16)
Private sector * competitive	0.10***	−0.09	−0.11
	(0.04)	(0.05)	(0.07)
Private sector * beds (100s)	−0.00	−0.01	0.02*
	(0.01)	(0.01)	(0.01)
Observations	9,573	11,849	9,767

Each estimate is from a separate regression of the stacked measures on hospital characteristics, financial year effects, hospital effects, the relevant CEO characteristic and the interaction of the relevant CEO characteristic and hospital characteristic. All outcome variables in a stacked set as well as the individual outcome variables are standardised to have a mean of zero and a standard deviation of one. Some outcome variables in the stacked sets are multiplied by (-1), so that "more" means "better". Standard errors in (parentheses), clustered at hospital level. *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-35: Association of tenure and residuals for all observed CEO spells and for subset of CEO spells included in parametric two-step approach (sample of movers)

	All CEO spells			CEO spells for sample of movers		
	Coefficient on tenure var. (std. error)	R^2	Obs.	Coefficient on tenure var. (std. error)	R^2	Obs.
Input measures	-0.001 (0.002)	0.001	2,538	0.002 (0.005)	0.017	854
Throughput measures	0.000 (0.002)	0.000	2,534	0.006* (0.003)	0.006	854
Clinical performance measures	-0.002 (0.003)	0.000	2,307	0.008 (0.007)	0.009	733

All regressions include a dummy variable indicating that tenure is unsure. The residuals are from a regression of the standardised stacked measures on hospital characteristics, financial year effects and hospital effects. The residuals for the 4 or 5 outcome measures in each set are then averaged by hospital-year. "All CEO spells" excludes spells at hospitals which we observe for only one year since hospital effects predict the outcome variable perfectly. "All CEO spells" also excludes hospital-year observations for which we observe fewer than 2 of the hospital production measures in a stacked set. Standard errors clustered at hospital level. *Significant at 10%, **significant at 5%, ***significant at 1%

W-8 Parametric estimates of across target variables associations

Table W-36: Association of means of residuals for CEO spells at first and second hospital for all combinations of target variables

Depend. variable \ Explanat. var.	Surplus at 1st hospital	Waiting time at 1st hosp.	Day cases at 1st hosp.	Length of stay at 1st hospital	MRSA rate at 1st hosp.	Staff job satis. at 1st hosp.
Surplus at 2nd hospital	**−0.05**	-0.22	0.28	-0.24	-0.002	-0.02
	(0.30)	(0.23)	(0.30)	(0.54)	(0.17)	(0.40)
	[95]	[93]	[95]	[94]	[80]	[73]
Waiting time at 2nd hospital	-0.07	**−0.01**	-0.02	0.27	0.05	-0.01
	(0.10)	**(0.08)**	(0.10)	(0.18)	(0.08)	(0.12)
	[93]	**[93]**	[93]	[92]	[79]	[72]
Day cases at 2nd hospital	0.10	-0.03	**0.18***	-0.31**	-0.12	0.11
	(0.09)	(0.07)	**(0.09)**	(0.15)	(0.08)	(0.11)
	[95]	[93]	**[95]**	[94]	[80]	[73]
Length of stay at 2nd hospital	-0.02	0.01	0.05	**0.05**	0.03	0.02
	(0.04)	(0.03)	(0.04)	**(0.06)**	(0.03)	(0.04)
	[94]	[92]	[94]	**[94]**	[80]	[72]
MRSA rate at 2nd hospital	0.08	0.23**	-0.01	0.37	**0.10**	0.38
	(0.13)	(0.10)	(0.14)	(0.24)	**(0.10)**	(0.15)
	[80]	[79]	[80]	[80]	**[80]**	[67]
Staff job satisfaction at 2nd hospital	0.10	-0.09	-0.10	-0.50	0.14	**−0.07**
	(0.13)	(0.11)	(0.13)	(0.28)	(0.11)	**(0.11)**
	[73]	[72]	[73]	[72]	[67]	**[73]**

Results are from regressions of mean of residuals for CEO spell in second hospital on mean of residuals for CEO spell in first hospital. The residuals are from a regression of the standardised measures on hospital characteristics, financial year effects and hospital effects. *Significant at 10%, **significant at 5%, ***significant at 1%

W-9 Potential reasons for lack of persistent CEO effects: Additional results

Table W-37: Linear probability models of the impact of good performance in first hospital on moving to a "prestigious" hospital

	Teaching hospital			Foundation trust			Hospital with beds above 75th percentile before CEO arrived		
	Good perform.	Const.	N	Good perform.	Const.	N	Good perform.	Const.	N
Surplus	0.10	0.25	95	0.06	0.29	95	-0.11	0.37	90
	(0.11)	(0.05)		(0.11)	(0.05)		(0.12)	(0.06)	
Waiting time	0.09	0.26	93	0.05	0.30	93	-0.03	0.35	88
	(0.11)	(0.05)		(0.11)	(0.06)		(0.12)	(0.06)	
Day cases	0.02	0.27	95	0.04	0.30	95	-0.17	0.39	90
	(0.11)	(0.05)		(0.11)	(0.06)		(0.11)	(0.06)	
Length of stay	-0.06	0.28	93	-0.11	0.32	94	0.03	0.34	90
	(0.11)	(0.05)		(0.11)	(0.05)		(0.12)	(0.06)	
MRSA rate	0.07	0.28	80	0.00	0.35	80	0.18	0.29	77
	(0.12)	(0.06)		(0.12)	(0.06)		(0.12)	(0.06)	
Staff job satisfaction	0.09	0.28	73	0.25*	0.33	73	0.01	0.32	71
	(0.12)	(0.06)		(0.13)	(0.07)		(0.13)	(0.07)	

Each entry in this table refers to a separate regression of an indicator of a CEO moving to a "prestigious" hospital on an indicator of good performance at the CEO's first hospital. "Prestigious" hospital is defined as either teaching hospital (major or minor), foundation trust status or number of beds above 75th percentile in the year before the CEO arrived. The 75th percentile of number of beds is calculated separately for each financial year to ensure the categorisation is net of year effects. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-38: Linear probability models of the impact of a large positive CEO fixed effect in pay ($\geq 75^{th}$ percentile) on ever being observed in a "problematic" hospital

	Hospital commission rating poor in year before CEO arrived	Hospital with surplus below 25^{th} percent. in year before CEO arrived	'New' hospital created through merger during sample period	Hospital with PFI contract at some point during CEO's tenure
CEO fixed effect in total pay $\geq$ £24,360	-0.054	-0.22**	0.022	0.19
	(0.12)	(0.11)	(0.12)	(0.12)
Constant	0.58	0.76	0.39	0.52
	(0.06)	(0.05)	(0.06)	(0.06)
R^2	0.002	0.04	0.00	0.03
Observations	89	94	95	95

Each column refers to a separate regression of an indicator of a CEO ever being observed at a "problematic" hospital on an indicator of a large positive CEO fixed effect in total pay. "Problematic" hospital is defined as either poor hospital commission rating, surplus below 25^{th} percentile, hospital created through merger or hospital with PFI contract. The 25^{th} percentile of surplus is calculated separately for each financial year to ensure the categorisation is net of year effects. The CEO fixed effects in total pay are the estimated executive director effects from the total pay regression in Table W-3, transformed into deviations from the mean of all estimated executive director effects and extracted for the subset of CEOs included in the analyses of CEO fixed effects in hospital production. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-39: Linear probability models of the impact of a large positive CEO fixed effect in pay ($\geq 75^{th}$ percentile) on ever being observed in a "prestigious" hospital

	Teaching hospital	Foundation trust	Hospital with beds above 75^{th} percentile before CEO arrived
CEO fixed effect in total pay $\geq$ £24,360	0.41***	-0.10	0.38***
	(0.11)	(0.11)	(0.11)
Constant	0.25	0.35	0.32
	(0.05)	(0.06)	(0.06)
R^2	0.14	0.009	0.11
Observations	95	95	95

Each column refers to a separate regression of an indicator of a CEO ever being observed at a "prestigious" hospital on an indicator of a large positive CEO fixed effect in total pay. "Prestigious" hospital is defined as either teaching hospital (major or minor), foundation trust status or number of beds above 75^{th} percentile in the year before the CEO arrived. The 25^{th} percentile of surplus is calculated separately for each financial year to ensure the categorisation is net of year effects. The CEO fixed effects in total pay are the estimated executive director effects from the total pay regression in Table W-3, transformed into deviations from the mean of all estimated executive director effects and extracted for the subset of CEOs included in the analyses of CEO fixed effects in hospital production. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-40: Linear probability models of the impact of good performance or bad performance on exiting the sample

	Good performance	Bad performance	Constant	N
Surplus	0.031 (0.067)	0.156*** (0.061)	0.56 (0.04)	359
Waiting time	0.007 (0.063)	0.048 (0.063)	0.59 (0.04)	359
Day cases	0.038 (0.062)	-0.089 (0.064)	0.62 (0.04)	359
Length of stay	0.103* (0.062)	0.072 (0.064)	0.56 (0.04)	359
MRSA rate	-0.022 (0.069)	-0.058 (0.071)	0.56 (0.04)	305
Staff job satisfaction	0.100 (0.070)	0.037 (0.074)	0.49 (0.04)	289

Each row in this table refers to a separate regression. The dependent variable is a dummy taking the value 1 if a CEO leaves the sample after being observed in only one hospital and the value 0 if a CEO is observed in at least two different hospitals. The dummy variable is set to missing for CEOs observed at only one hospital if this spell includes the year 2013, as 2013 is the last year in our sample. The independent variables are an indicator of good performance and an indicator of bad performance at the CEO's first or only hospital. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

Table W-41: Estimates of quality of CEO-hospital matches for subset of CEO spells included in parametric two-step approach (sample of movers)

	Surplus	Waiting time	Day cases	Length of stay	MRSA rate	Staff job satisf.
Female * Teach	0.09	−0.03	−0.64***	0.13	−0.27	−0.07
	(0.11)	(0.09)	(0.19)	(0.10)	(0.32)	(0.08)
Female * FT	−0.08	−0.02	−0.06	0.01	0.10	−0.14
	(0.16)	(0.12)	(0.10)	(0.05)	(0.12)	(0.17)
Female * Comp	0.33	−0.13	0.06	−0.04	0.18	−0.14
	(0.27)	(0.21)	(0.11)	(0.06)	(0.12)	(0.18)
Female * Beds (100s)	0.02	−0.01	−0.01	0.00	−0.01	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Doctor * Teach	0.38	−0.16	0.56***	−0.06	−1.41***	−0.14
	(0.43)	(0.21)	(0.16)	(0.21)	(0.47)	(0.21)
Doctor * FT	−0.33	−0.02	−0.05	−0.15	−0.69	−0.29
	(0.43)	(0.21)	(0.20)	(0.10)	(0.11)	(0.22)
Doctor * Comp	−0.61	0.19	−0.28	−0.08	0.22	0.46
	(0.70)	(0.30)	(0.24)	(0.13)	(0.46)	(0.31)
Doctor * Beds (100s)	−0.02	0.02	0.04	−0.00	−0.00	−0.01
	(0.07)	(0.02)	(0.03)	(0.02)	(0.06)	(0.05)
Priv. sec. * Teach	0.12	−0.11	−0.19	−0.04	−0.42	−0.02
	(0.29)	(0.15)	(0.29)	(0.10)	(0.38)	(0.18)
Priv. sec. * FT	0.34**	0.11	0.33	−0.16	−0.14	0.20
	(0.16)	(0.19)	(0.32)	(0.10)	(0.46)	(0.32)
Priv. sec. * Comp	−0.17	0.37***	0.20	0.01	0.61***	−0.40***
	(0.16)	(0.10)	(0.15)	(0.06)	(0.19)	(0.14)
Priv. sec. * Beds (100s)	0.00	−0.01	−0.04*	−0.00	−0.04**	−0.00
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Observations	2,396	2,356	2,383	2,386	2,055	1,838

Teach = teaching hospital, FT = foundation trust, Comp = competitive market. Each estimate is from a separate regression of the relevant hospital production measure on hospital characteristics, financial year effects, hospital effects, the relevant CEO characteristic and the interaction of the relevant CEO characteristic and hospital characteristic. All outcome variables are standardised to have a mean of zero and a standard deviation of one. Standard errors in (parentheses), clustered at hospital level. *Significant at 10%, **significant at 5%, ***significant at 1%