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 and University of Bristol

Raffaella Sadun
Harvard Business School

Working Paper 18-075

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

Katharina Janke[a], Carol Propper[b], and Raffaella Sadun[c]

[a]Lancaster University
[b]Imperial College and University of Bristol
[c]Harvard University

March 14, 2018

Abstract

We investigate whether top managers affect the performance of large public sector organizations. As our case study we examine CEOs of English public hospitals, which are large, complex organizations with multi-million turnover. We study the impact of individual CEOs on a wide set of measures of hospital performance, intermediate operational outcomes and inputs. We adopt two econometric approaches: a parametric approach that exploits the movement of CEOs across different hospitals and a non-parametric difference-in-difference matching estimator. Overall, we find little evidence that individual CEOs have an impact on a large set of measures of hospital performance. This result is not due to the allocation of good performers to poorly performing hospitals.

Corresponding author: Katharina Janke, Furness Building, Lancaster, LA1 4YG, UK.
Email: k.jankemarie@lancaster.ac.uk

1 Introduction

The effect of CEOs in private organizations has been explored in a number of influential studies beginning with Bertrand and Schoar (2003).[1] More recently, a number of papers have shown that CEOs can also impact the performance of public sector organizations. However, this result has so far been documented for relatively small public organizations–for example, schools and development projects–where top managers may have a greater chance of having an impact.[2] 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 these contexts? Addressing this question is important because a popular reform model in the public sector is to give greater autonomy to CEOs to run their organizations, accompanied by the use of manager-specific compensation policies, performance-related pay and dismissals (e.g. Besley and Ghatak (2003), LeGrand (2003)).

We contribute to the literature by looking at CEOs of very large and complex organizations in the public sector. The setting of our study is public sector hospitals in the English NHS. These organizations have on average 4,500 employees, multi-million turnover, with labour accounting for around 70% of costs of production.

This setting is an ideal test bed for several reasons. First, in the late 1980s the English government embarked on a large reform programme which replaced an 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 a fully decentralized fashion.[3] This context led to frequent movements of the same CEOs across NHS hospitals, thus providing an ideal setting to study whether individual CEOs are indeed associated with systematic differences in hospital performance. Second, data are available for these organisations on a wide set of measures of production including key financial targets, clinical outcomes and intermediate outputs and operational variables, allowing us to ask which aspects of performance CEOs can and cannot affect. Third, the NHS requires trusts to publish the pay awarded to their top managers, thus allowing us to complement the performance analysis with complementary evidence on CEO

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

[2]A number of papers investigate the impact of principals on student performance, for example, Böhlmark et al. (2016) presents evidence of significant principal fixed effects in students' outcomes and Lavy and Boiko (2017) also find school principals affect student performance. Bloom, Lemos, Sadun and Van Reenen (2015) examine managerial practices in schools and find they are correlated with school performance. Rasul and Rogger (2018) examine the behaviour of government bureaucrats. Other papers include Branch et al. (2012), Coelli and Green (2012), Dhuey and Smith (2014) and Grissom et al. (2015).

[3]Very similar reforms were adopted in a number of public health care systems and in public administration more generally. See for example Pollitt and Bouckaert (2000).

compensation.

We begin by examining whether the movement of individual CEOs is associated with significant differences in hospital performance. To estimate whether CEOs make a difference, we undertake two complementary approaches, both of which utilise the movement we observe across hospitals. In the first–a parametric approach–we examine whether CEOs have a 'style', i.e. whether they are able to affect hospital outcomes in the same way across different organizations. We adopt the approach pioneered by Bertrand and Schoar (2003) which examines whether there are CEO fixed effects. We do this in two ways to overcome potential statistical issues. We first assess whether CEO movement across hospitals is characterized by systematic within-hospital variation in performance. We then examine whether deviations in any one measure of hospital performance during a CEO's tenure at one hospital are positively correlated with deviations in the same measure in the CEO's tenure at a second hospital.[4] Our second approach is non-parametric, and compares changes in hospital performance after a CEO turnover event to changes experienced by matched hospitals without such an event.

We find little consistent evidence of any effect of the CEO on the large set of production metrics that we can examine. In the parametric approach, we find that the estimated CEO fixed effects are jointly statistically significant. However, the CEO fixed effects are essentially period-hospital-specific shocks rather than true CEO effects, and therefore large deviations in (an aspect of) performance in one hospital are typically not replicated by the same CEO in another hospital. Using the non-parametric approach, hospitals with a CEO turnover event differ from matched hospitals without a CEO turnover event in terms of only a small number of inputs in hospital production (growth in beds, patient length of stay and job satisfaction of staff). No other difference can be found in any of the numerous financial, operational and clinical metrics that we consider.

We contrast the null findings on hospital performance with results examining differences in pay across CEOs (following Abowd et al. (1999)). While the level and the dispersion in compensation across NHS managers is considerably smaller than the one documented in the private sector, we find considerable and persistent differences in managerial pay. Moving from the 25^{th} percentile to the 75^{th} percentile of the CEO pay effects distribution represents a 12% increase in pay relative to mean CEO pay. These results, combined with the lack of systematic differences across CEOs in terms of hospital performance, suggests that NHS board may overestimate the ability of individual managers to affect hospital performance, or compensate CEOs for non-performance related factors.

[4] We also examine correlations across all the different measures of production we examine to study whether there are systematic differences across clusters of outcomes.

Finally, we show that the lack of a CEO effect in hospital performance is not driven by the endogenous assignment of CEOs. CEOs who perform well in one hospital do not then move systematically to hospitals which are performing poorly or have structural features which mean that achieving good performance is more difficult.

Overall, our results indicate that the CEOs of large public hospitals such as those included in the NHS do not bring about changes in hospital performance, a result that stands in stark contrast with earlier findings relating to the private sector and to smaller public sector organizations. In the conclusion we discuss various structural factors which may account for this lack of effect including the public sector nature of the NHS, which may force NHS CEOs to pursue political targets rather than performance enhancing policies. However, the lack of a CEO effect may also be due, more broadly, to the complexity of hospital production, which transcends the fact that the NHS is publicly owned.[5] From this perspective, the results presented in this paper cast some doubts on the effectiveness of a "turnaround CEO" approach–i.e. a model in which top managers frequently rotate across hospitals to induce meaningful changes in performance–for large public sector organizations.

2 Institutional Background

From the 1980s the English government followed a programme of giving greater autonomy to the management of public sector organizations coupled with a series of reforms designed to subject these organizations to the discipline of the market (Le Grand 1991). From the mid-1990s, English public hospitals operated as free-standing organisations, earning revenue from contracts won in competition with other public hospitals and increasingly, from the mid-2000s, private sector hospitals. From the mid-1990s hospitals were also subject to corporate governance reforms similar to ones brought into private sector firms in the UK in 1992 (Cadbury 1992). These reforms required English hospitals to establish boards with executive and non-executive directors whose responsibility was to run the hospital. Political oversight remained at both a regional and central level, but the day-to-day operation of the hospital and responsibility for meeting government targets was vested with the trust board.

With these changes came a greater emphasis on the role of the executive directors and the Chief Executive (the NHS term for the CEO).[6] CEOs are appointed by the board.

[5] See Chandra et al. (2016) for evidence of large performance differentials across hospitals in the USA.

[6] The emphasis on the top manager began much earlier in 1983 following the Griffiths report, which recommended replacing the prevailing consensus management system with a general manager who had overall responsibility for service performance and management (Baggott 1994). During the mid-1990s

In making their choice, the appointment committee will almost alway use private sector headhunters to help them select potential candidates and will typically also consult (usually the regional arms of) the NHS Executive (the national level government organisation responsible for overseeing the NHS).[7]

From 2003 hospitals that met key performance targets set by government were granted greater autonomy and were free to set CEO pay, which was decided upon by the remuneration committee of the hospital as in any private company.[8] The remuneration committee can decide if a proportion of executive directors' remuneration should be linked to corporate and individual performance.[9] In contrast, the pay of clinical staff (including physicians) and lower level managerial staff is set at national level (with some regional uplifts) by a public sector pay review body.

The devolution of responsibility for performance to the hospital level has been accompanied by an increase in publicly available data on hospital performance, including measures of financial performance, access to care (waiting times, which are important in a system where care is rationed) and, since the 2000s, measures of the quality of clinical care. The key performance targets have varied over the period we examine, but have typically included measures of financial performance (with a focus on deficits), waiting times, length of stay (as a measure of efficiency) and more recently avoidance of poor clinical care. Chief Executives are answerable to their boards, but are also subject to close scrutiny by central government (during the period we study this was by the NHS Executive). Missing key performance targets set by central government can 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, reflecting the view that top managers were responsible for the performance of their hospital.

The belief in the importance of senior managers to hospital performance is also reflected in the growth of 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. Figure 1 illustrates this growth. It shows the level of CEO pay over our sample period from 2000 to 2013 as

market reforms, these general managers were renamed Chief Executives and the role of the hospital board strengthened.

[7]In hospitals which have not been granted Foundation Trust status – which gives hospitals greater autonomy from NHS Executive control – one of the appointment panel will be from the regional NHS Executive.

[8]Guidance states that the board of directors must establish a remuneration committee composed of non-executive directors, which should include at least three independent non-executive directors and which decides on pay of all executive directors (Monitor 2014).

[9]Guidance states that the remuneration committee should judge where to position its NHS Foundation Trust relative to other NHS Foundation Trusts and comparable organisations (Monitor 2014).

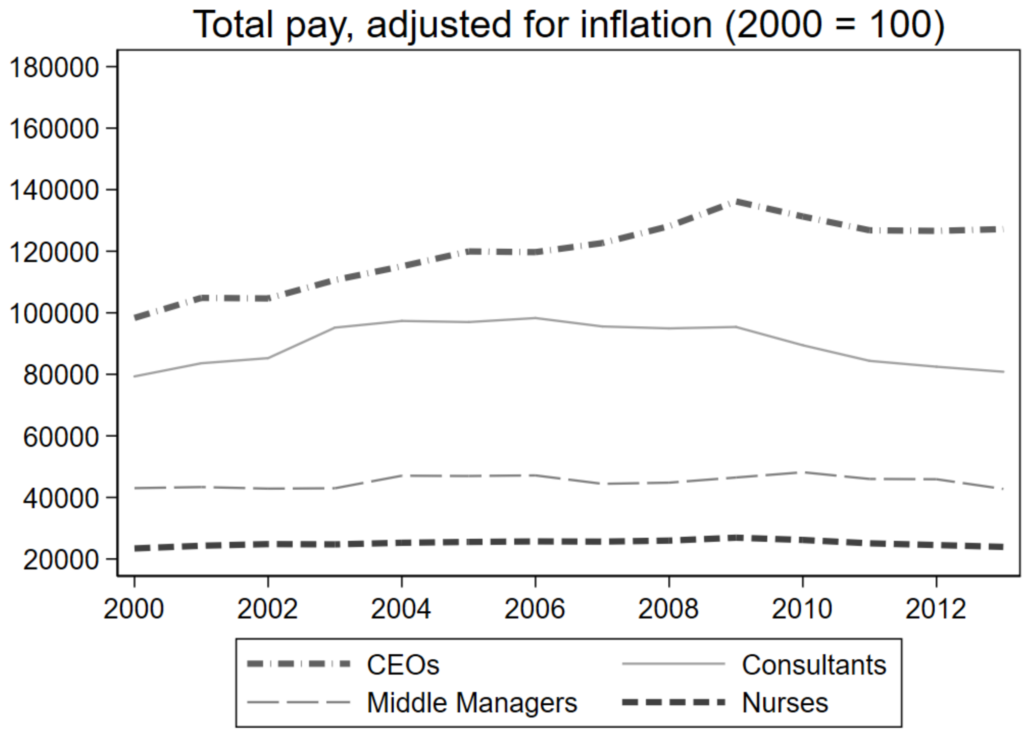


Figure 1: Annual means of pay for NHS staff by job type

well as mean pay of nurses, consultants (senior physicians) and middle managers. Figure 2 shows that over our sample period CEO pay increased faster at the top than at the bottom, with the difference between the 10^{th} and the 90^{th} percentile increasing from £40,000 in 2000 to £60,000 in 2013. At the top of the distribution CEO pay increased from £120,000 in 2000 to £160,000 in 2013.

However, despite the larger increases in CEO pay relative to other hospital staff, CEO remuneration packages are still dwarfed by the financial rewards earned by their counterparts in the UK corporate sector. Bell and Van Reenen (2016) report that 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. Even taking into account that the figures in Bell and Van Reenen (2016) are in 2014 prices whilst our figures are in 2000 prices, these remuneration packages are larger by an order of magnitude.

The remuneration packages are also small compared to the figures reported by Joynt et al. (2014) for CEOs of US non-profit hospitals. For 2009, they report mean compensation of $596,000 (approximately £400,000), but the majority of CEOs in their sample served at hospitals with fewer than 300 beds while in our sample even the 25^{th} percentile is 446 beds. Focusing on the figures Joynt et al. (2014) report for the highest decile of the compensation distribution, which has the largest mean number of beds (though still only

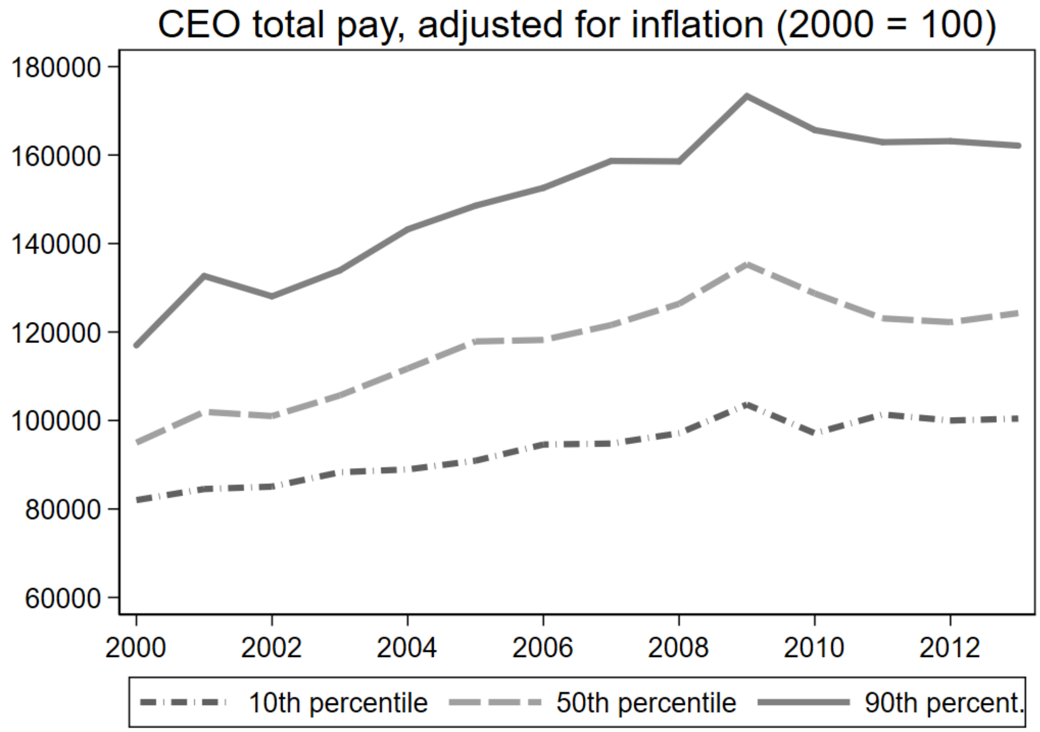


Figure 2: Annual percentiles of pay for CEOs of NHS hospitals

310), mean compensation is $2,100,000 (approximately £1,400,000), which is an order of magnitude larger than compensation of CEOs of NHS hospitals.

In the context of the UK public sector, however, the relatively small remuneration packages of NHS CEOs are at the high end of the compensation distribution for public service managers. The Prime Minister's salary of around £145,000 is often used as a benchmark in public debate and salaries higher than this attract considerable (negative) attention from the popular media.[10]

3 Data

3.1 Data Sources

Our analysis is based on data from various administrative data sources, which have been combined together for the first time. Our starting point are the NHS Boardroom Pay

[10]For example, it is common to find British media articles about "NHS fat cats" receiving "six-figure salaries" or "earning more than the Prime Minister". A report by an important UK health policy "think tank" documents politicians' attacks on the "pen pushers" and "men in grey suits" and the public support for reducing the number managers in the NHS (Ham et al. 2011).

Reports published by IDS Incomes Data Services for 2001 to 2011, which provide information on where each CEO worked and when. We extended this series by hand collecting data from NHS hospital trusts' annual reports for the financial years 2011/12 to 2013/14. We identify CEO turnover by combining into a single data set information on 14 financial years between 2000/01 and 2013/14. To reliably identify moves of CEOs across hospitals, we manually checked the personal identifiers for all executive directors in the data.[11]

We additionally hand collected data on CEO characteristics, such as gender, educational achievements, clinical background and public honours.[12] Table 1 summarizes the main demographic and sample characteristics of the CEOs.

Table 1: Demographic and sample characteristics of CEOs

	Number	Proportion
Female	147	31%
Clinical background	112	24%
MBA or similar qualification	121	26%
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	

We combine the turnover data with a rich set of measures at hospital level for the financial years 2000/01 to 2013/14. From a range of sources we have brought together input measures such as number of beds or number of nurses as a proportion of all staff, throughput measures such as waiting times or length of stay, clinical performance measures such as deaths within 30 days of emergency admission for myocardial infarction or MRSA

[11]For example, we checked all executive directors with the same surname or slightly different spellings of the same surname. We also checked for name changes following marriage.

[12]The British honours system recognises people who have made achievements in public life and who have committed themselves to serving and helping Britain. For example, one of the authors (CP) has received a CBE for services to Social Science. Titles bestowed upon hospital CEOs include Knight, Dame, Commander of the Order of the British Empire (CBE), Officer of the Order of the British Empire (OBE) and Member of the Order of the British Empire (MBE).

bacteremia rates, and surplus as a financial performance measure. For brevity, we will classify all these measures as "hospital performance".[13]

Finally, the IDS data also provide data on salary, taxable benefits and total remuneration of executive directors for nearly all NHS hospital trusts.[14] Because of changes to reporting rules, the 2000/01 pay data are limited to CEOs but from 2001/02 onwards the pay data cover all executive directors. The core executive director positions present on all hospital boards are CEO, Medical Director, Nursing Director, Finance Director and HR Director. 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".

Table 2 presents descriptive statistics for the pay and the hospital performance data. 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. Rows 1 and 2 report statistics for basic pay and total pay of all executive directors. Since the pay data are limited to CEOs in 2000/01, the mean for 2000/01 is larger than the means for 2006/07 and 2013/14. Figures 1 and 2 in Section 2 show the time series for total pay for the subset of CEOs. To ensure comparability, we have dropped 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 financial year). The number of observations for the hospital-level variables is determined by their availability and reflects the observations used in the estimations reported below.

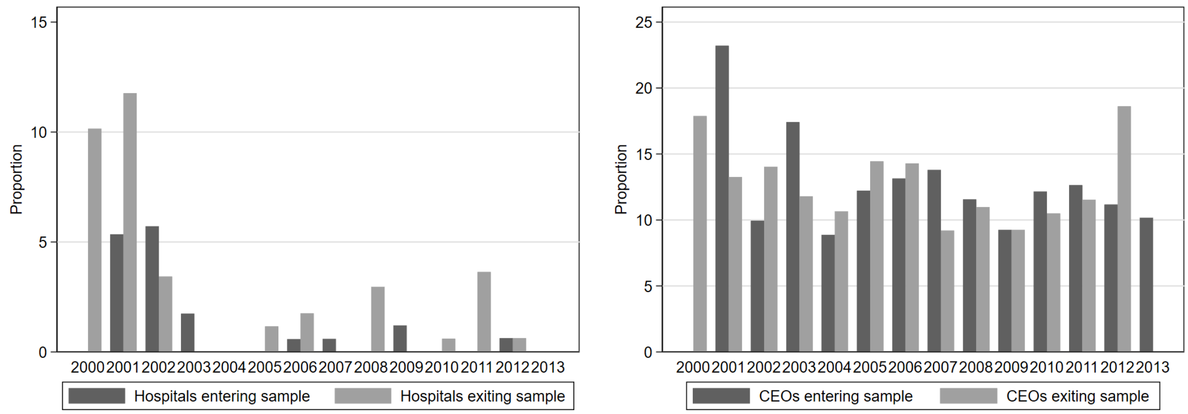
3.2 Entry and Exit of Hospitals and CEOs

Figure 3a shows sample entry and exit of hospitals and Figure 3b shows sample entry and exit of CEOs. Figure 3a shows considerable sample exit and entry of hospitals at the beginning of our sample period: in 2000 and 2001 over 10% of hospitals exit our sample. The reason is a period of intense hospital consolidation in the NHS. Between 1997 and 2003, over half the stock of NHS acute hospitals in 1997 were involved in some kind of merger or reconfiguration with other NHS hospitals (Gaynor et al. 2012). There is also an uptick in consolidation activity at the end of our sample period. These mergers reflected a worldwide trend for consolidation in the hospital sector and meant that NHS hospitals grew in size and in the number of sites in which they provide services. Whilst we refer to these organisational units as hospitals, they are formally known as Hospital Trusts, which

[13]Details on the sources of these data are in Appendix A.
[14]The financial year runs from 1 April to 31 March.

Table 2: Descriptive statistics for executive director pay and hospital behaviour and performance measures

	Obs.	Mean	St. dev.	Mean of variable in 2000	2006	2013
Basic pay, RPI adjusted (£)	8,749	86,276	24,828	93,672	84,963	89,235
Total pay, RPI adjusted (£)	8,760	87,389	25,575	98,010	85,784	90,448
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
Contracted out (%)	1,645	34.7	28.7	33.3 (2004)	35.2	35.0
Technology index	2,398	0.38	0.23	0.29	0.39	0.43
Beds (count)	2,398	722	402	702	727	683
Beds growth	2,165	-0.017	0.085	0.008 (2001)	-0.048	-0.001
Senior doctors growth	2,171	0.06	0.114	0.027 (2001)	0.041	0.030
Nurses growth	2,171	0.020	0.097	0.015 (2001)	0.004	0.023
Admissions (count)	2,392	74,488	42,778	54,000	74,229	92,422
Admissions growth	2,351	0.024	0.075	-0.004	0.030	0.026
Length of stay, mean (days)	2,386	5.23	2.87	7.29	4.80	4.33
Day cases (%)	2,383	31.3	8.7	29.5	30.0	34.9
Waiting time, mean (days)	2,356	70.5	30	93.5	73.9	48.9
Cancelled operations (count)	2,328	373	290	401	301	404
Staff job satisfaction	1,838	3.47	0.10	3.47 (2003)	3.39	3.61
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	2,055	10.2	8.36	15.7 (2001)	16.6	2.4
Surplus	2,396	-1,965	15,101	259	-796	-4,975



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

Figure 3: Number of CEOs observed per hospital for hospitals observed for at least 11 years and number of CEO spells at different hospitals for executive directors that are observed in a CEO position at least once

reflects the fact that many are formed from consolidations across two or more hospital sites. All these sites, however, are in the same geographical area: there are no hospital chains within the NHS.[15] These consolidations were accompanied by changes in CEOs, as at the very least only one of the CEOs of the formerly separate hospitals continued in post. Frequently a new CEO was appointed to lead the consolidated hospital.[16]

Figure 3b shows the extent of CEO sample entry and exit. While the pattern of hospital sample entry and exit is fairly low and stable from 2003 onwards, CEO sample entry and exit is on average considerably higher than hospital sample entry and exit, on average around 14% for the whole period. CEO sample entry and exit are highest during the period of consolidation at the beginning of our sample, then fall and remain relatively stable after 2004, but are still both over 10% at the end of the period.[17]

3.3 CEO Turnover

The market for hospital CEOs in England is characterized by very high separation rates. Hospitals which are in our data for at least 11 years have on average 3.5 CEOs. Figure 4 shows the annual proportion of hospitals with a CEO turnover event in our sample.[18]

[15] All mergers/consolidations are within the NHS; there are none with private hospitals. Private hospitals predominantly provide services for which there are long NHS waiting lists.

[16] Following a merger, the new hospitals were generally given a new name and NHS code. We treat each new code as a separate hospital.

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

[18] As our data start in 2000, we report turnover events only from 2001 onwards. Some hospitals experience more than one CEO turnover event in a financial year, a fact that would not be visible in

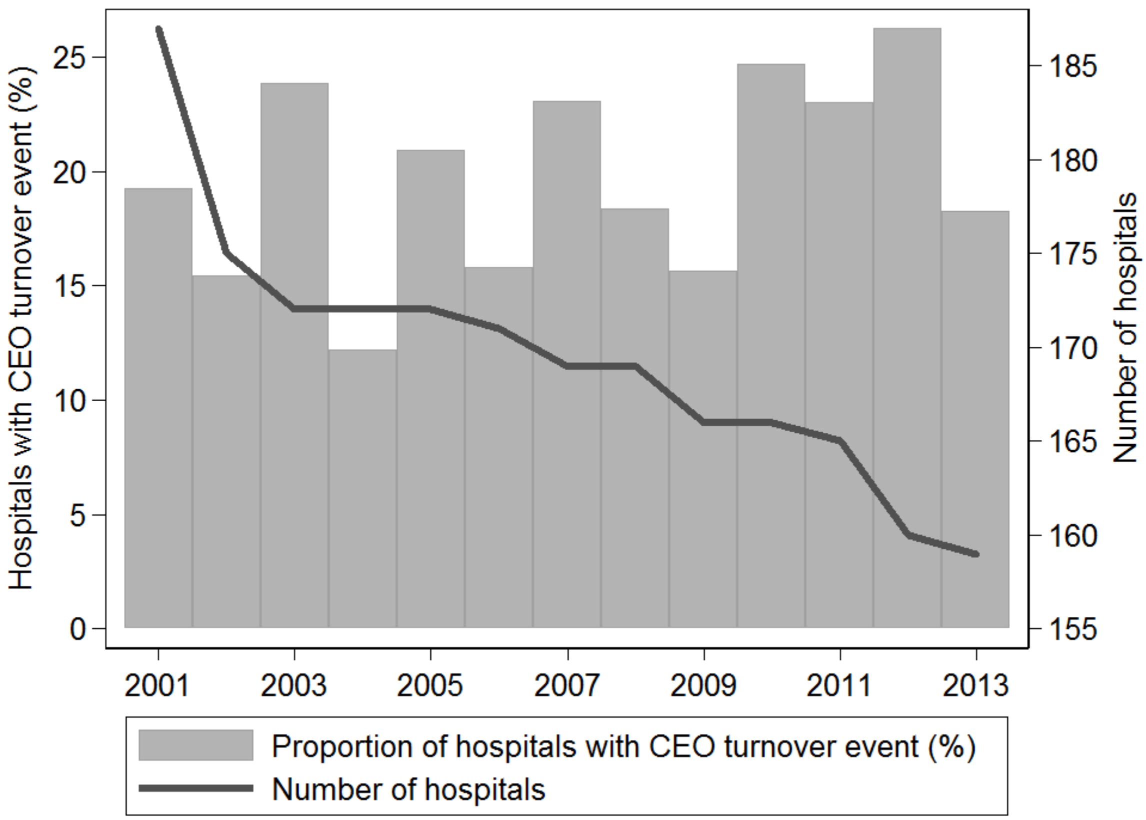


Figure 4: Annual proportion of hospitals with CEO turnover event

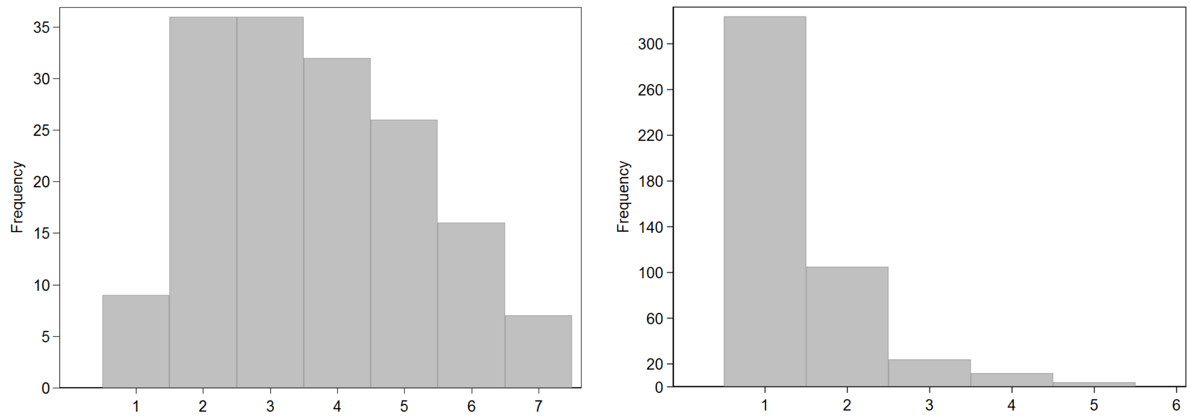
Between 12 to 25% of hospitals in our sample have a turnover event in any year. Figure 5a shows CEO turnover per hospital for the subset of hospitals observed for at least 11 years.[19] Only a minority of hospitals have the same CEO for the whole sample period. The majority have two to five CEOs over the sample period of 11 to 14 years, with a minority of hospitals having more than this. Hospitals with more CEOs over the sample period tend to be in certain broad regions of England – the North East has the lowest turnover and the East Midlands the highest – but few other time-invariant characteristics such as being a large teaching hospital or being a specialist hospital are associated with the number of CEOs a hospital has over our sample period.[20]

Figure 5b shows, for the sample of executive directors that were observed at least once in the position of CEO (N = 469), the number of CEO spells at different hospitals. More than 100 directors served as a CEO in at least two different hospitals. This subset

Figure 4 since only the first turnover event determines the hospitals classified as having experienced a CEO turnover event in the particular year.

[19]As our data set does not include some of the CEOs that served for less than a year, the number of CEOs per hospital could be a lower bound.

[20]For details see Table W-1 in the Web Appendix.



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

Figure 5: Number of CEOs observed per hospital for hospitals observed for at least 11 years and number of CEO spells at different hospitals for executive directors that are observed in a CEO position at least once

of CEOs is our starting point for the sample for which we investigate the existence of CEO fixed effects. Over all CEOs, the median number of years a CEO is observed in a particular CEO job is 3 years and the mean is 3.7 years.[21] For the subset of CEO spells we use to estimate CEO fixed effects the number of years they are observed in a CEO spell is a minimum of 2 years by construction, but still the median is only 4 years and the mean 4.5 years.[22]

To examine whether turnover CEOs are different from others we regressed fixed characteristics of the CEO against whether a CEO ever moved, whether they held a job for longer than the median and whether they were in our fixed effects estimation sample. The characteristics we examined were gender, whether the CEO has a clinical qualification, whether they have an MBA type post-graduate qualification and whether they ever received a national honour. CEOs who never move (which may include those who are in post for only a short duration) are less likely to have an MBA but otherwise do not differ from all other CEOs. CEOs with tenure longer than the median of 3 years are more likely to be female and less likely to have a clinical qualification (and to have received a national honour). The 95 CEOs we use to estimate the CEO fixed effects are more likely to have an MBA type qualification (reflecting the fact that they do move and those who do not move are less likely to have such a qualification) but otherwise do not differ in terms of

[21] The number of years a CEO is observed in a CEO job is not necessarily the job duration since the data often report that a CEO served only for part of the financial year, i.e. the CEO served for less than 12 months. Unfortunately, we do not know the number of months for which CEOs served who served for less than the full financial year.

[22] For the sample of CEOs for whom we estimate pay effects, the mean number of years observed per CEO spell is 4.1 years.

gender or clinical background from the rest of the CEOs in the NHS in our sample period.

4 CEO Fixed Effects: Hospital Performance

We employ two different approaches to estimate the impact of individual CEOs on hospital performance. The first one is parametric and exploits movement of the same CEO across different hospitals. We use the fixed effects approach pioneered by Bertrand and Schoar (2003). We regress measures of hospitals' performance on observable hospital characteristics, hospital effects and CEO effects for the subset of CEOs observed in at least two hospitals for at least two years in each. To assess the validity of this approach we estimate CEO fixed effects for random CEO-hospital matches and compare these estimates to the estimates for the actual CEO-hospital matches.[23] We also apply an alternative two-step procedure proposed by Bertrand and Schoar (2003), which is based on the examination of CEO-spell fixed effects. Our second approach is non-parametric and resembles a difference-in-difference matching estimator.

4.1 Basic Approach

The fixed effects approach proposed by Bertrand and Schoar (2003) involves estimating regressions of the following form:

$$y_{jt} = \boldsymbol{X}'_{jt}\boldsymbol{\beta} + \lambda_t + \psi_j + \alpha_{i(j,t)} + \varepsilon_{jt} \tag{1}$$

The left-hand side variable, y_{jt}, is one of several measures of inputs, throughputs or clinical and financial performance 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 time-varying observable hospital characteristics that includes merger status, number of beds, a technology index and case mix measures; more details are in Appendix A. We also include a full set of financial year effects, λ_t, 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. Standard errors are clustered at hospital level.

We estimate CEO effects $\alpha_{i(j,t)}$ only for the subset of CEOs observed in two hospitals for at least two years each.[24] The CEO effect for a CEO observed in only one hospital, but

[23]Fee et al. (2013) investigate the validity of F-tests on the CEO fixed effects by randomly assigning CEOs to a different second firm than the one they actually joined. We randomly assign CEOs to both the first and the second firm and our analysis looks beyond F-tests.

[24]A few CEOs are observed in three or four hospitals for at least two years each. For these CEOs we

for only part of the time period we observe the hospital for, would be identified but would capture a period-hospital-specific effect rather than a CEO effect. In estimating CEO effects only for CEOs observed in two hospitals, any effects that matter would require that corporate practices be correlated across two hospitals when the same CEO is present (Bertrand and Schoar 2003). The requirement that CEOs have to be observed in each hospital for at least two years ensures they are given time to "imprint their mark".

A number of complications arise when determining which CEOs comply with our requirements. Firstly, because of limited data availability several of our hospital performance variables are missing for many of our hospital-year observations.[25] Therefore, we determine the CEOs complying with our requirements separately for each of the hospital performance variables y_{jt} by first dropping the CEO-year observations for which the hospital performance variable is missing. Second, 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. Third, the CEO effect for a CEO observed in one hospital for the 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 CEOs comply with our requirements.

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 two hospitals the CEO has been observed in, for the financial years the CEO has been observed there. Following Bertrand and Schoar (2003), we present F-statistics from tests of the joint significance of the CEO effects.

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. This issue is illustrated in Fee et al. (2013). Using data similar to the data used by Bertrand and 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. F-tests for the CEO effects derived using these random CEO-firm matches suggest highly statistically significant CEO effects.

use only their two most recent spells because using all three or four spells would "require that corporate practices have to be correlated" (Bertrand and Schoar 2003) across three or four hospitals for these CEOs, whereas for all other CEOs the requirement is only to have practices correlated across two hospitals.

[25]For some variables observations are missing because these 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.

We assess the validity of F-tests on CEO effects in the context of our data by randomly assigning CEOs to both their first and second hospital.

Our starting point are the CEO spells that we use for estimating CEO effects in Equation 1. For example, a CEO might be observed at Hospital A from 2001/02 to 2004/05 and at Hospital B from 2005/06 to 2008/09. We randomly assign this CEO to a hospital for the period 2001/02 to 2004/05 and we randomly assign this CEO to a hospital for the period 2005/06 to 2008/09. The pool of hospitals for the random assignment is made up of the hospitals that actually hosted one of the CEOs observed in two hospitals for at least two years each. 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 1 for the sample with the random CEO-hospital matches $i(j,t)$, test the joint significance of the CEO effects using an F-test, count the number of CEO effects that are individually statistically significant, and calculate the proportion of the variance of the left-hand side variable y_{it} that is explained by the CEO effects. We repeat this process 100 times and compare the means over the 100 replications to the values obtained using the actual CEO-hospital matches $i(j,t)$.

4.2 Two-step Procedure

Bertrand and Schoar (2003) propose an alternative two-step procedure for assessing the effect of a CEO. To implement this, in the first step we regress our measures of inputs, throughputs or clinical and financial performance, y_{jt} on the vector of time-varying observable hospital characteristics $\boldsymbol{X}_{jt}$, the financial year effects λ_t and the hospital effects ψ_j:

$$y_{jt} = \boldsymbol{X}'_{jt}\boldsymbol{\beta} + \lambda_t + \psi_j + \varepsilon_{jt} \tag{2}$$

We extract the residuals e_{jt} from Equation 2. For each CEO observed in two hospitals for at least two years each, we generate 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 we regress the mean for CEO i's spell in hospital B on the mean for CEO i's spell in hospital A:

$$\frac{1}{n^{i,B}} \sum_{t=t_1^{i,B}}^{t_n^{i,B}} e_{Bt} = \delta_1 + \delta_2 \frac{1}{n^{i,A}} \sum_{t=t_1^{i,A}}^{t_n^{i,A}} e_{At} + \varepsilon_i \tag{3}$$

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. To check the validity of this two-step procedure, we estimate Equations 2 and 3 for the simulation data with random CEO-hospital matches $i(j,t)$ and compare the means over the 100 replications to the values obtained using the actual CEO-hospital matches $i(j,t)$.

As a robustness test, we also run a placebo regression proposed by Bertrand and Schoar (2003). Instead of using the mean of the residual at hospital B during the time the CEO was observed there, we use the mean of the residual at hospital B during the three financial years before the CEO arrived there. The idea is that a positive δ_2 in Equation 3 might wrongly suggest that individual CEOs have an impact on hospital performance. 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 move to more day case procedures at hospital A might be recruited to oversee a similar move 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 therefore suggestive of selection of the CEO rather than of the CEO imposing their style. On the other hand, if hospital B's deviations from the expected proportion during the three years before CEO i arrived are completely unrelated to the deviations during CEO i's spell at hospital B, we are more confident that a positive δ_2 in Equation 3 indicates the impact of the CEO on hospital performance.

4.3 Non-parametric Approach

Both the fixed effects approach and the two-step procedure rely heavily on our statistical model of hospital performance, since we use the residuals from this statistical model to estimate the impact of individual CEOs. They also rely on CEOs having an impact on the same dimension of hospital performance across two hospitals. A non-parametric approach avoids both problems. It resembles a difference-in-difference estimator combined with matching.[26] Essentially, we compare the changes in hospital performance following a CEO turnover event to changes in hospital performance at matched hospitals without a CEO turnover event. If there is any impact of CEOs on hospitals' performance, we expect to see different changes after a CEO turnover event compared to otherwise similar hospitals with no CEO turnover event.

We start by identifying hospitals that had a CEO turnover event that resulted in stable leadership for at least two years. Next, we select from this set of observations those CEO turnovers that followed stable leadership in the previous two years. 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. Next, we match these 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 $t-2$ to hospitals with no CEO turnover from $t-2$ to $t+1$. Finally, we compare the difference in our hospital performance measures between the year before the CEO turnover and the end of the two-year period, i.e. between $t-1$ and $t+1$, to the equivalent difference in the matched hospitals.

We match–with replacement–treated hospitals to control hospitals exactly on year, teaching status, specialist status and foundation trust status in $t-1$.[27] This tends to result in more than one match for each treated hospital. Therefore, in the next step we use closest neighbor matching on beds in $t-1$ to choose one or three control hospitals from among the exactly matched hospitals. Where closest 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 closest neighbor in terms of the technology index in $t-1$. Matching exactly on year implies that we compare, for example, the difference in waiting times between 2006 and 2009 for a hospital with a CEO turnover event in 2007 to the difference in waiting times between 2006 and 2009 for a hospital with no CEO turnover event in 2007. Thus, our results will not be confounded by period effects (for example, the general decline in waiting times during the early 2000s (Gaynor et al. 2012)).

For all of our measures of inputs, throughputs, clinical and financial performance y_{jt} we report the mean of the change in the treated hospitals $\frac{1}{n^T}\sum_{j=1}^{n^T}\left(y^T_{j(t+1)}-y^T_{j(t-1)}\right)$ and its

[26] The difference-in-difference matching estimator was introduced by Heckman et al. (1997) and Heckman et al. (1998) and further developed by Abadie (2005). Fee et al. (2013) use a similar approach but make a distinction between exogenous and endogenous CEO turnover events.

[27] Matching on hospital 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 rather than the three different specialties acute, children and orthopedic. Teaching status and specialist status are permanent characteristics while foundation trust status is a time-varying characteristic.

standard error, the mean of the change in the control hospitals $\frac{1}{n^C}\sum_{j=1}^{n^C}\left(y^C_{j(t+1)} - y^C_{j(t-1)}\right)$ and its standard error, the difference between the two means as well as the standard error and p-value from a two-sample t-test with equal variance.

5 Results

We start by estimating Equation 1. We do this for actual CEO-hospital matches and then for the random CEO-hospital matches. We examine a large set of measures of hospital production but as the results are similar across measures we only present results for a selection of our measures. In Table 3 we examine input and throughput measures. These are two measures of inputs (the ratio of the most skilled staff to number of beds as a measure of the labour-to-capital ratio and the ratio of senior doctors tor staff as a measure of the labour skills ratio) and two throughput measures (waiting times and length of stay) which have been used as key performance measures for NHS hospitals during all of our sample period. In Table 4 we examine four measures of output, three of which are measures of clinical quality (AMI deaths, readmissions and MRSA rates), and one of which is a measure of financial performance (financial surplus). Financial surplus has been used as part of the assessment of performance of NHS hospitals during the whole of the sample period. The measures of clinical quality have been used towards the end of our sample period. Results for the remaining measures are in the Web Appendix.

Statistical Significance of CEO Fixed Effects

We begin by examining the results for the actual CEO-hospital matches, presented in the first row of each panel in Tables 3 and 4. The R^2 in Column 3 is large for the input and throughput measures and also for two of the clinical performance measures, suggesting 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 these measures. The F-tests in Column 1 suggest that the estimated CEO effects are jointly statistically significantly different from zero for all our input, throughput and performance measures. The proportion of CEO effects that are individually statistically significantly different from zero varies from 24.2% for surplus to 34.7% for the skill share (ratio of senior doctors to all staff).

The last five columns of Tables 3 and 4 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 performance

Table 3: Estimates of CEO effects for a subset of input and throughput measures for actual CEO-hospital matches as for 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)	25.0 (27.2%)	0.90	2,382	36.4	51.4	4.9	7.3	826.8
(Std. dev.)	(31.8) (n.a., 1.23, 0)	(4.66, 4.99)	(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)	28.1 (30.0%)	0.89	2,396	45.6	38.5	3.2	12.7	842.3
(Std. dev.)	(49.6) (n.a., 1.09, 0)	(4.84, 5.13)	(0.002)		(3.21)	(3.82)	(1.75)	(3.14)	(12.2)
Waiting times									
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)	28.1 (30.6%)	0.84	2,356	53.7	28.0	6.1	12.3	815.8
(Std. dev.)	(62.5) (n.a., 1.64, 0)	(4.15, 4.49)	(0.003)		(2.71)	(3.21)	(2.40)	(1.54)	(15.5)
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)	23.9 (25.8%)	0.95	2,386	47.8	40.3	1.6	10.3	831.8
(Std. dev.)	(40.1) (n.a., 1.32, 0)	(4.51, 4.83)	(0.001)		(9.16)	(8.26)	(2.72)	(2.36)	(12.8)

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 pay 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 (except for (doctors + nurse)/beds), technology index and case mix variables. The results for random CEO-hospital matches are means and standard deviations across 100 replications.

Table 4: Estimates of CEO effects for a subset of our performance measures 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)	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)	15.7 (29.5%)	0.48	1,757	17.9	18.1	12.5	51.5	430.8
(Std. dev.)	(23.0) (n.a., 3.25, 0)	(3.54, 6.55)	(0.005)		(2.90)	(4.24)	(3.73)	(4.40)	(25.9)
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)	21.6 (30.3%)	0.78	2,070	26.9	39.4	9.7	24.0	583.3
(Std. dev.)	(31.2) (n.a., 1.44, 0)	(4.17, 5.81)	(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)	25.4 (29.6%)	0.77	2,055	53.3	22.5	5.9	18.3	748.3
(Std. dev.)	(54.1) (n.a., 1.64, 0)	(4.00, 4.55)	(0.004)		(2.46)	(2.46)	(2.30)	(1.58)	(16.2)
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)	21.8 (23.2%)	0.29	2,396	4.9	17.8	13.8	63.5	843.6
(Std. dev.)	(36.9) (n.a., 1.06, 0)	(4.30, 4.58)	(0.01)		(1.42)	(2.99)	(3.42)	(3.16)	(12.1)

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 pay 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 (unless beds is the input measure), technology index (unless technology index is the input measure) and case mix variables. The results for random CEO-hospital matches are means and standard deviations across 100 replications.

measures that is explained by each term in Equation 1: the covariates (time-varying hospital characteristics + year effects), the hospital effects, the CEO effects and the residuals. The residual variance proportion is generally larger for outcome measures than for either inputs or throughputs and is largest for AMI deaths and surplus. This reflects the more general and widely documented problem of large unexplained variation in outcome measures in hospital production. A considerable proportion of the variance is accounted for by the observed covariates with one exception, surplus, where the covariates have little role. Hospital fixed effects account for a large fraction of the variance across all dependent variables, ranging from nearly 50% for skill mix to 15% for surplus. More generally, the hospital effects are larger for inputs and throughputs, reflecting the fact that different types of hospital employ different mixes of capital and labour and serve different patient groups, and smallest for outcomes, reflecting again the variation in hospital output across observably similar firms.

The CEO effects, while jointly statistically significant as measured by the F-test, explain less of the variance than either the covariates or the hospital effects. On average in the subsample of hospital-year observations with at least one CEO effect $\alpha_{i(j,t)}$, the CEO effects explain around 6% of the variance in the performance measures. The proportion ranges from 0.5 for length of stay to 19.5 for surplus. For surplus alone, 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. These results suggest that there are statistically significant CEO effects and that the CEO may have a larger impact on outputs than inputs.

However, in the random CEO-hospital matches reported in the second and third row of each panel the means of the F-statistics across the 100 replications are as large as they are 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 around 30%, just as it is for the actual CEO-hospital matches. Finally, the mean variance proportion explained by the CEO effects when CEOs are randomly assigned to hospitals is similar to the variance proportions explained by the CEO effects using the actual assignments, ranging from 1.6% for length of stay to 13.8% for surplus.

These results suggest that the CEO effect estimates, and therefore the F-tests, may be capturing period-hospital-specific shocks rather than true CEO effects. Estimating CEO effects only for CEOs that are observed in two hospitals does not seem to ensure that the estimates do not simply capture period-hospital-specific effects.[28]

Results for the Two-step Method

We now turn to the results from the alternative two-step method of Equations 2 and 3. We have assessed the validity of the two-step method by applying it to our random CEO-hospital matches. Table 5 presents the results for the input and throughput measures, and in Table 6 are the results for the performance measures. A positive coefficient indicates that a positive deviation from the expected level of a production measure during a CEO's spell at the first hospital is associated with a positive deviation from the expected level of that measure during the CEO's spell at the second hospital and vice versa. A statistically significant association would suggest that these deviations can be attributed to the CEO and not to period-hospital-specific effects.

We find that, regardless of the input, throughput or performance measure, the means of the coefficient estimates $\widehat{\delta_2}$ across the 100 replications are very small, the rejection frequencies of t-tests of $\widehat{\delta_2}$ are close to the nominal level of the test, and the explanatory power of the regressions as measured by the mean R^2 is very low.[29] Thus, the results for the random CEO-hospital matches show no impact of CEOs, exactly what we would expect for random matches.

More specifically, while there are a few positive coefficients, but most are not statistically significant. However, even the statistically significant coefficients are problematic. For example, the positive and statistically significant coefficient of day cases is mirrored by a statistically significant positive coefficient of the same size in the placebo regression, suggesting that larger than expected day case proportions were already happening in the second hospital before a CEO with larger than expected day case proportions at their first hospital arrived. The only statistically significant positive coefficient that is not mirrored in the placebo regression is for nurses growth. Furthermore, several coefficients are in fact negative, suggesting, for example, that more than expected beds growth during a CEO's spell at the first hospital is associated with lower than expected beds growth during the CEO's spell at the second hospital. However, these negative coefficients are small and not statistically significant.

Overall, these results suggest that the statistical significance of the CEO fixed effects

[28]Fee et al. (2013) argue that standard asymptotic theory does not apply to tests on CEO dummy variables, because the number of dummies increases as the sample grows larger. They also claim that high serial correlation of measures of firm behaviour lead to inference issues. Our finding that estimates of the variance proportion explained by the CEO effects are also not valid suggests that the problem is not only caused by non-applicability of standard asymptotic theory.

[29]More details are in the Web Appendix.

is driven by hospital-period-specific shocks, and not by persistent CEO effects.

Non-parametric Estimates

Finally, we present results for our non-parametric approach which seeks to establish whether there is a CEO effect by comparing changes in hospital performance following a CEO turnover event to changes in hospital performance at matched hospitals without a CEO turnover event.[30]

Information on the quality of our matching is presented in Tables B-2 and B-3 in Appendix B. The tables report, for the treated and the control observations, the means of the hospital characteristics that we include as control variables in the wage equation (5) and in the regressions (1) and (2) to assess the balance of the matched samples. The tables also show the means of the characteristics we match exactly on (teaching status, specialist status and foundation trust status). Since we generate two sets of controls– one derived from 1:1 matching and one derived from 1:3 matching–there are two sets of statistics for controls for each production measure. For almost all the measures there is little difference between the treated and the two control samples, with the exception of beds whose number is slightly large in the treated sample. Thus, the matching produces a good balance.[31]

Using both matched samples, Table 7 presents the results for the input and throughput measures and Table 8 presents the results for the throughput and clinical and financial performance.

Table 7 shows that, in the main, inputs do not change after a new CEO is in post. However, there is one exception–the number of beds–which falls. There is also a fall in one key throughput measures, length of stay, which may be a result of the fall in beds. Table 8 shows that clinical and financial performance do not on balance improve, with improvements on some performance measures matched by reductions in other measures. Staff satisfaction falls following a CEO turnover event.

In a robustness test we apply the non-parametric estimator to the subset of the 95 CEOs that we use in our parametric approach. The results are in Web Appendix Tables W-8 and W-9. They show very similar results to those for the larger sample: some indication that there was a faster drop in the number of beds and length of stay after a CEO

[30] These estimates are for a smaller sample as we need to have information on changes between years.

[31] In the Web Appendix we report in Tables W-6 and W-7 a check of the common trend assumption which examines changes in the outcome variables for the two-year period before the CEO turnover event. We find very few differences between the treated and the control hospitals. Nurses as proportion of all staff seem to have dropped less fast in treated hospitals between $y_{j(t-3)}$ and $y_{j(t-1)}$ but otherwise the trajectories seem to be very similar, providing support for the parallel trend assumption.

Table 5: Association between mean of residuals for CEO's spell in first hospital and mean of residuals for CEO's spell in second hospital for input and throughput measures

	Real regressions			Placebo regressions		
	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)		
Contracted out	-0.04	0	68	0.17	0.03	68
	(0.11)			(0.11)		
Technology	0.001	0	95	-0.05	0	92
	(0.10)			(0.10)		
Beds	0.05	0	95	-0.01	0	92
	(0.17)			(0.17)		
Beds growth	-0.13	0.01	86	0.09	0	82
	(0.11)			(0.16)		
Senior doctors growth	0.09	0.02	86	-0.15	0.02	83
	(0.08)			(0.12)		
Nurses growth	0.16**	0.07	86	-0.16*	0.04	83
	(0.07)			(0.09)		
Admissions	0.11	0.01	95	-0.005	0	92
	(0.12)			(0.11)		
Admissions growth	0.04	0	92	-0.16*	0.03	88
	(0.09)			(0.10)		
Length of stay	0.05	0.01	94	-0.04	0	91
	(0.06)			(0.09)		
Day cases	0.18*	0.04	95	0.19**	0.04	92
	(0.09)			(0.10)		

The residuals are from a regression of the input or throughput measure on hospital characteristics, financial year effects and hospital effects. The results in the Placebo regressions column are from regressions of the mean of the residuals in the second hospital during the three years before the CEO arrived there on the mean of the residuals for the CEO's spell at the first hospital. *Significant at 10%, **significant at 5%, ***significant at 1%

Table 6: Association between mean of residuals for CEO's spell in first hospital and mean of residuals for CEO's spell in second hospital for throughput and performance measures

	Real regressions			Placebo regressions		
	Coefficient (std. error)	R^2	Obs.	Coefficient (std. error)	R^2	Obs.
Waiting times	-0.01	0	93	0.01	0	90
	(0.08)			(0.08)		
Cancelled operations	-0.12	0.01	90	0.32	0.03	87
	(0.17)			(0.21)		
Staff satisfaction	-0.07	0	73	-0.11	0.01	73
	(0.11)			(0.17)		
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)		
MRSA rate	0.10	0.01	80	-0.05	0	78
	(0.10)			(0.12)		
Surplus	-0.05	0	95	0.16	0.01	92
	(0.30)			(0.22)		

The residuals are from a regression of the performance measure on hospital characteristics, financial year effects and hospital effects. The results in the Placebo regressions column are from regressions of the mean of the residuals in the second hospital during the three years before the CEO arrived there on the mean of the residuals for the CEO's spell at the first hospital. *Significant at 10%, **significant at 5%, ***significant at 1%

Table 7: Changes in input and throughput measures 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
Doctors + nurses/beds	Treated	205	0.20 (0.02)		
	Controls	205	0.20 (0.02)	-0.00 (0.03)	0.82
	Controls	596	0.21 (0.01)	-0.01 (0.02)	0.68
Senior doctors/staff	Treated	205	0.67 (0.13)		
	Controls	205	0.62 (0.10)	0.05 (0.17)	0.76
	Controls	596	0.75 (0.07)	-0.07 (0.14)	0.59
Nurses/staff	Treated	205	-0.25 (0.12)		
	Controls	205	-0.12 (0.13)	-0.13 (0.17)	0.46
	Controls	596	-0.24 (0.07)	-0.01 (0.14)	0.95
Contracted out	Treated	145	-0.12 (1.33)		
	Controls	145	0.23 (1.21)	-0.35 (1.80)	0.85
	Controls	413	0.71 (0.73)	-0.83 (1.46)	0.57
Technology	Treated	205	0.024 (0.005)		
	Controls	205	0.018 (0.004)	0.007 (0.006)	0.27
	Controls	596	0.016 (0.002)	0.008 (0.005)	0.08
Beds	Treated	205	-28.2 (4.83)		
	Controls	205	-15.6 (4.86)	-12.7 (6.86)	0.07
	Controls	596	-19.7 (2.76)	-8.66 (5.49)	0.12
Admissions	Treated	205	4,216 (404)		
	Controls	205	4,955 (542)	-739 (676)	0.28
	Controls	596	5,098 (367)	-882 (668)	0.19
Length of stay	Treated	205	-0.48 (0.07)		
	Controls	205	-0.35 (0.04)	-0.13 (0.08)	0.10
	Controls	596	-0.32 (0.03)	-0.16 (0.06)	0.01
Day cases	Treated	202	0.94 (0.26)		
	Controls	202	0.73 (0.31)	0.21 (0.40)	0.60
	Controls	586	1.16 (0.18)	-0.22 (0.35)	0.53

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 $t-2$. One or 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. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table 8: Changes in throughput and performance measures 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
Waiting times	Treated	200	-9.83 (1.29)		
	Controls	200	-8.72 (1.10)	-1.11 (1.69)	0.51
	Controls	583	-8.98 (0.66)	-0.85 (1.36)	0.53
Cancelled operations	Treated	202	-15.8 (14.5)		
	Controls	202	-3.15 (11.3)	-12.6 (18.4)	0.49
	Controls	589	-13.0 (8.28)	-2.74 (16.5)	0.87
Staff satisfaction	Treated	163	0.013 (0.008)		
	Controls	163	0.032 (0.007)	-0.019 (0.011)	0.07
	Controls	468	0.025 (0.004)	-0.013 (0.009)	0.14
AMI deaths	Treated	143	-0.64 (0.30)		
	Controls	143	-0.54 (0.27)	-0.10 (0.41)	0.80
	Controls	424	-0.50 (0.15)	-0.14 (0.31)	0.65
Stroke deaths	Treated	168	-2.21 (0.30)		
	Controls	168	-1.07 (0.34)	-1.15 (0.45)	0.01
	Controls	505	-1.33 (0.19)	-0.88 (0.37)	0.02
FPF deaths	Treated	165	-0.16 (0.23)		
	Controls	165	-0.38 (0.24)	0.22 (0.33)	0.51
	Controls	495	-0.31 (0.12)	0.14 (0.25)	0.57
Readmissions	Treated	172	0.54 (0.09)		
	Controls	172	0.50 (0.08)	0.03 (0.12)	0.78
	Controls	503	0.54 (0.04)	0.001 (0.09)	0.99
MRSA rate	Treated	197	-2.19 (0.40)		
	Controls	197	-2.30 (0.42)	0.11 (0.58)	0.85
	Controls	572	-2.34 (0.24)	0.15 (0.48)	0.75
Surplus	Treated	205	1,444 (1,088)		
	Controls	205	2,105 (829)	-661 (1,368)	0.63
	Controls	596	103 (720)	1,340 (1,384)	0.33

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. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

move and a smaller increase in staff satisfaction and a faster drop in stroke deaths, with the estimates not being statistically significant consistently. A similar picture emerges when we examine changes over the three years following the CEO turnover event rather than two years. This robustness test is presented in Web Appendix Tables W-10 and W-11.

We conclude from these analyses that–with the possible exception of changes in bed numbers and length of stay and a negative impact on staff satisfaction–incoming CEOs do not appear to have a statistically significant effect on hospital production.[32]

6 CEO Fixed Effects: Pay

The performance results suggest that individual CEOs–or, more generally, simply the event of a change in CEO–are not associated with systematic differences in hospital performance. We now turn to study whether and how the lack of performance differentials across CEOs examined in the previous section is also found when examining CEO remuneration. To do so, we use the Abowd et al. (1999) approach to estimate CEO fixed effects in pay.

We use pay data for all executive directors, i.e. including COOs, Finance Directors, HR Directors, Nursing Directors and other directors but excluding Medical Directors.[33] 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. However, since different types of executive directors receive markedly different pay packages, we employ a two-step estimation procedure. We first regress executive directors' pay on a set of dummy variables indicating their board level position:

$$\begin{aligned} pay_{it} = \delta_1 + \delta_2 COO_{it} + \delta_3 finance\_director_{it} + \delta_4 HR\_director_{it} \\ + \delta_5 nursing\_director_{it} + \delta_6 other\_director_{it} + \varepsilon_{it} \end{aligned} \quad (4)$$

pay_{it} denotes pay of executive director i in financial year t. COO_{it} is an indicator

[32] We also examine the impact of a CEO turnover event on the quality of middle management, using data from the 2006 and 2009 wave of the World Management Survey. There are only 9 treated observations, so the effect estimate is imprecise. However, if anything it suggests that a turnover event decreases management quality. More details are in the Web Appendix.

[33] We exclude Medical Directors because their salaries in the directors' remuneration data sets are lower than the salaries of other executive directors since for many Medical Directors a major part of their income is remuneration for clinical work, which is not included in the directors' remuneration data sets.

variable that takes the value one if the job title of executive director i during financial year t was Chief Operating Officer and zero otherwise. Similarly, the other variables indicate a board position as Finance Director, HR Director, Nursing Director and Other type of executive director, respectively. CEO is the omitted category. We estimate this regression using the same observations that we include in the second step and extract the residuals to use them as the outcome variable in the wage equation.[34]

In the second step we estimate the following wage equation:

$$pay\_residual_{it} = \boldsymbol{X}'_{j(i,t)t}\boldsymbol{\beta} + \gamma\, tenure_{ij(i,t)t} + \lambda_t + \alpha_i + \psi_{j(i,t)} + \varepsilon_{it} \tag{5}$$

The left-hand side variable, $pay\_residual_{it}$, is the residual from the regression in Equation 4, i.e. the pay of executive director i in period t net of the impact of their board level position.[35] The function $j(i,t)$ maps executive director i to hospital j in financial year t. $\boldsymbol{X}_{j(i,t)t}$ is the same set of time varying hospitals variables included in Equation 1, $tenure_{ij(i,t)t}$ is the tenure of executive director i at hospital $j(i,t)$ 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, $\psi_{j(i,t)}$, 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_{it} represents the error term.

As discussed by Abowd et al. (1999), for observations not connected by worker mobility it is not possible to identify separate executive director effects α_i and hospital effects $\psi_{j(i,t)}$. Therefore, we estimate Equation 5 using all pay observations for the largest subset of hospitals that are connected by executive directors moving between them. We calculate the proportion of the variance in the pay variable, $pay\_residual_{it}$, that is explained by the covariates, $\boldsymbol{X}_{j(i,t)t}$, $tenure_{ij(i,t)t}$ and λ_t, the hospital effects, $\psi_{j(i,t)}$, and the executive director effects, α_i, respectively. For the hospital effects and the executive director effects, this proportion is simply $[\text{Cov}(pay\_residual_{it}, \widehat{\psi}_{j(i,t)})/\text{Var}(pay\_residual_{it})] \times 100$ and $[\text{Cov}(pay\_residual_{it}, \widehat{\alpha}_i)/\text{Var}(pay\_residual_{it})] \times 100$. To obtain the proportion explained by the covariates, we first calculate the pay residual predicted by the coefficient estimates for the covariates, $\widehat{pay\_residual}_{it} = \boldsymbol{X}'_{j(i,t)t}\widehat{\boldsymbol{\beta}} + \widehat{\gamma} tenure_{ij(i,t)t} + \widehat{\lambda}_t$, and then use this prediction to calculate the covariance: $[\text{Cov}(pay\_residual_{it}, \widehat{pay\_residual}_{it})/\text{Var}(pay\_residual_{it})] \times 100$.

[34] In the second step we only include observations for which we can separately identify executive director effects and hospital effects - more details below.

[35] A one-step estimator that includes the indicator variables for the board level position in Equation 5 does not fully remove the impact of the board level position on pay as the coefficients on the indicator variables are identified only by the handful of executive directors changing board level position.

In terms of sample selection, we drop from the pay data set 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 financial year) to ensure comparability.

Table 9 reports the results from estimating Equation 5 and shows the proportions of the variance in the pay variables that are explained by the covariates, the hospital effects, the director effects and the residuals, respectively. We estimate Equation 5 only for the pay observations in the largest connected set. In fact, there is only one connected set and only 162 pay observations in 17 hospitals that are not connected by worker mobility. The connected set has 478 movers that connect 196 hospitals. Table 9 shows that in the connected set the director effects are jointly statistically significant. The hospital effects, director effects and covariates jointly explain more than 85% of the variation in executive director pay, with the covariates accounting for around 20% of the variation and the hospital and director effects each accounting for around 30%.

Table 9 also presents results for the subset of directors observed in a CEO position at least once (397 of the 2,111 executive directors in the connected set) and for the further subset of CEO who are included in the management style estimations (95 of the 397 CEOs). The director effects are jointly statistically significant in both subsets; the variance decompositions are similar across all the different sets.

The interquartile range in hospital (firm) pay effects is around £15,000. In the Web Appendix we present correlates of this variation. We find pay effects are higher in teaching hospitals and smaller in specialist hospitals and there is is also considerable regional pay variation, reflecting regional differences in the cost of living.

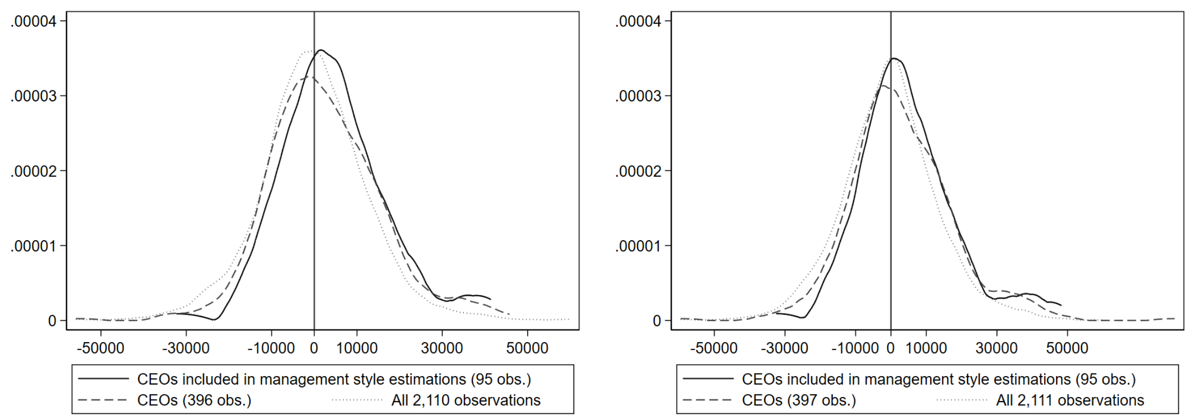
Figure 6 shows the distribution of the pay effects for all directors in the connected set, the 397 who were ever CEOs and the subset of 95 CEOs. Since the $\widehat{\alpha}_i$ are estimated relative to an arbitrary omitted director, we have transformed the estimates into deviations from the mean of all $\widehat{\alpha}_i$. The distribution for the 95 CEOs included in the management style estimation lies slightly to the right of the distribution for all CEOs, though the distribution for all CEOs has longer right and left tails. For both basic pay and total pay the interquartile range is around £17,000 for all CEOs and around £14,500 for the 95 CEOs for whom we estimate managerial effects. For the full sample of director pay effects the interquartile range is £15,000 for basic pay and £16,400 for total pay.

In Table 10 we examine which personal and sample-specific characteristics are associated with the CEO pay effects and test whether there are differences in these associations between all CEOs in our sample and the 95 CEOs for whom we can estimate managerial effects. We find that while pay effects are positively associated with being observed in our sample for 10 years and more, which could indicate longer tenure, and being observed

Table 9: Summary statistics for the pay regressions

	F-test of joint significance of director effects (p-value, df1, df2)	R^2	Obs.	Variance proportions (%) Co-variates	Hospital effects	Director effects	Re-siduals	Hospitals	Persons	Movers
Connected set										
Basic pay	6.03 (0, 2109, 6420)	0.87	8,749	22.9	36.7	27.8	12.6	196	2,110	478
Total pay	5.77 (0, 2110, 6430)	0.86	8,760	21.4	35.3	29.4	13.9	196	2,111	479
Subset of directors observed in a CEO position at least once:										
Basic pay	6.81 (0, 396, 6420)		1,845	22.9	29.5	32.5	15.1	196	396	170
Total pay	7.52 (0, 397, 6430)		1,851	20.6	28.6	34.2	16.5	196	397	171
Subset of CEOs included in management style estimations:										
Basic pay	9.96 (0, 95, 6420)		629	18.8	32.3	32.2	16.7	121	95	93
Total pay	11.27 (0, 95, 6430)		633	16.7	30.1	33.1	20.1	122	95	94
Outside connected set										
Basic pay			162					17	47	
Total pay			162					17	47	

The dependent variable is the residual from a regression of the pay variable (RPI adjusted) on job title dummies. Variance proportion is the proportion of variance in the pay 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, case mix variables and tenure.



(a) Basic pay (b) Total pay

Figure 6: Kernel density plots of deviations of estimated director effects in pay from mean of all estimated director effects in pay for all directors, subset of directors observed in a CEO position at least once and subset of CEOs included in management style estimations

in 3 or more CEO jobs, which could indicate more mobility of CEOs across hospitals, these associations are not statistically significant. However, personal characteristics are associated statistically significantly with the pay effects. Individuals who have received a public honour and those who have a clinical background are paid more, women and those with an MBA are paid less. Importantly, there are very few differences in the patterns of these associations for the 95 CEOs who are the focus of our examination of managerial effects, suggesting that the determinants of remuneration for this group are the same as those for all the other hospital CEOs that we observe.

These results show that there are significant and persistent differences in the pay that different CEOs in the NHS receive.

7 Endogenous Assignment?

Our results show little persistence in the CEOs effects in performance, i.e. the period-hospital-specific effects vary considerably within CEO, but signficant and persistent differences in pay across CEOs. These results might be driven by the endogenous assignment of CEOs to hospitals. For example, a CEO who experiences a positive shock in one hospital may subsequently be hired by a hospital in which it is difficult to bring about positive changes, so that an above expected performance would be followed by a below average performance. We test this hypothesis in two ways.

First, we generate a measure of the variability in CEO performance across the two

Table 10: Association between estimated director effects in pay and personal characteristics for subset of directors observed in a CEO position at least once

	Basic pay		Total pay	
	Coefficient	Obs. in category	Coefficient	Obs. in category
Female	-4,289**	124	-5,292***	125
	(1,737)		(1,942)	
Female × In 95 CEOs subset	177	31	-535	31
	(3,680)		(4,119)	
Clinical background	3,184*	95	4,959**	96
	(1,886)		(2,105)	
Clinical background × In 95 CEOs subset	-2,559	25	-4,273	25
	(3,889)		(4,351)	
MBA or similar qualification	-1,864	107	-3,758**	107
	(1,696)		(1,897)	
MBA × In 95 CEOs subset	1,835	31	3,290	31
	(3,319)		(3,715)	
Public honour	5,218**	51	4,781*	51
	(2,388)		(2,672)	
Public honour × In 95 CEOs subset	2,942	16	1,973	16
	(4,432)		(4,961)	
Observed as CEO for 2 to 9 years	1,189	289	-2,129	290
	(2,462)		(2,756)	
2 to 9 years × In 95 CEOs subset	-1,354	52	-1,324	52
	(3,447)		(3,859)	
Observed as CEO for 10 plus years	6,796**	76	4,080	76
	(3,386)		(3,790)	
10 plus years × In 95 CEOs subset	-2,205	43	-2,806	43
	(4,418)		(4,945)	
Observed in 2 CEO jobs	1,215	102	1,407	102
	(2,445)		(2,737)	
Observed in 3+ CEO jobs	5,949	38	5,332	38
	(3,623)		(4,056)	
3+ CEO jobs × In 95 CEOs subset	-883	24	1,014	24
	(5,373)		(6,015)	
Constant	164		4,142	
	(2,374)		(2,657)	
R^2 / Observations	0.12	396	0.11	397

The executive director effects are extracted from the regressions reported in Table 9 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%

hospitals we observe a CEO in. The starting point for this measure are the mean of the residuals from Equation 2 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 calculate this variability measure for all of our 22 production measures.

We examine whether the variability measure is larger for CEOs who are at some point in their career assigned to "problematic" hospitals. We use four definitions of "problematic" hospitals: (i) having received a poor rating from the government regulator of hospitals for the year before the CEO arrived at the hospital,[36] (ii) having a financial surplus below the 25th percentile 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 large capital investment–a PFI contract–at some point during the CEO's tenure.[37]

For each of our four definitions of "problematic", we regress each of our 22 variability measures against a dummy variable indicating that the CEO was ever observed in a "problematic" hospital. Thus, we run 88 separate regressions and obtain 88 coefficients on a "problematic" hospital dummy variable. The results are in Table 11. Nine out of the 88 coefficients, i.e. 10%, are statistically significantly different from zero at the 10% significance level, a result we would expect just by chance. Furthermore, only three of them (for waiting times, cancelled operations and admissions) are positive, suggesting that being at a "problematic" hospital is associated with higher variability in CEO performance. For the other six statistically significant coefficients the estimated association is negative, suggesting that CEOs who are at some point at a more problematic hospital actually have lower variability in their performance across hospitals.

Second, we examine whether CEOs who did well at their first hospital are subsequently hired at a problematic hospital. We define doing well relative to a CEO's peers using the

[36]Because of data limitations we cannot always use the rating for the year before the CEO arrived at the hospital. For 2002, we use the contemporaneous rating, for 2003 to 2008 the rating for the year before the CEO arrived, for 2009 the rating from two years before the CEO arrived and for 2010 the rating from three years before the CEO arrived. As ratings are available only for parts of our sample period, we do not always observe a rating for both hospitals a CEO has served at and for some CEOs we do not observe any rating. 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. More details on the regulator ratings are in Appendix A.

[37]NHS hospitals have to borrow for large capital investments from the private market. Borrowing is through vehicles with long-term fixed interest rates and 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.

Table 11: 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)
Doctors + nurses/beds	-0.06 (0.04) [88]	-0.06 (0.04) [91]	-0.10 (0.04) [94]	-0.05 (0.04) [94]	0.19 (0.18) [94]
Senior doctors/staff	-0.04 (0.08) [89]	-0.02 (0.11) [92]	-0.11 (0.10) [95]	0.03 (0.10) [95]	0.58 (0.49) [95]
Nurses/staff	0.05 (0.20) [89]	-0.06 (0.21) [92]	0.15 (0.20) [95]	0.08 (0.19) [95]	1.21 (0.93) [95]
Contracted out	0.40 (2.31) [65]	0.23 (2.39) [68]	2.72 (2.27) [68]	-0.74 (2.21) [68]	8.17 (9.04) [68]
Technology	-0.008 (0.01) [89]	0.013 (0.01) [92]	0.002 (0.01) [95]	0.00 (0.01) [95]	0.052 (0.046) [95]
Beds	7.68 (12.3) [89]	6.14 (11.5) [92]	13.8 (11.8) [95]	15.5 (11.6) [95]	56.5 (56.4) [95]
Beds growth	-0.003 (0.009) [82]	-0.013 (0.009) [84]	0.004 (0.009) [86]	-0.020** (0.008) [86]	0.039 (0.04) [86]
Senior doctors growth	0.007 (0.006) [82]	0.008 (0.007) [84]	0.00 (0.006) [86]	0.007 (0.006) [86]	0.031 (0.028) [86]
Nurses growth	-0.004 (0.008) [82]	0.001 (0.008) [84]	0.001 (0.008) [86]	-0.008 (0.008) [86]	0.032 (0.035) [86]
Admissions	869 (933) [89]	-959 (980) [92]	-832 (907) [95]	1,603* (886) [95]	5,343 (4,326) [95]
Admissions growth	-0.64 (0.58) [86]	-0.48 (0.60) [89]	-0.98* (0.55) [92]	0.88 (0.54) [92]	2.90 (2.58) [92]
Length of stay	-0.001 (0.08) [88]	-0.02 (0.07) [91]	0.10 (0.08) [94]	0.03 (0.08) [94]	0.38 (0.36) [94]
Day cases	0.06 (0.41) [89]	0.41 (0.42) [92]	-0.01 (0.39) [95]	0.44 (0.39) [95]	2.37 (1.87) [95]
Waiting time	4.00** (1.82) [88]	2.89 (1.95) [90]	1.10 (1.83) [93]	-1.02 (1.81) [93]	10.8 (8.66) [93]
Cancelled ops.	2.77 (25.5) [84]	25.4 (25.2) [87]	73.8*** (24.0) [90]	29.1 (24.6) [90]	124 (116) [90]
Job satisfaction	-0.003 (0.006) [70]	-0.003 (0.006) [73]	-0.003 (0.006) [73]	-0.015*** (0.006) [73]	0.032 (0.025) [73]
AMI deaths	0.14 (0.39) [59]	-0.33 (0.46) [58]	-0.35 (0.37) [61]	-0.79** (0.38) [61]	1.34 (1.46) [61]
Stroke deaths	0.19 (0.34) [69]	-0.19 (0.37) [69]	-0.34 (0.32) [72]	-0.10 (0.33) [72]	1.83 (1.35) [72]
FPF deaths	-0.13 (0.26) [69]	0.25 (0.30) [69]	-0.44* (0.25) [72]	-0.34 (0.26) [72]	1.06 (1.06) [72]
Readmissions	-0.02 (0.11) [76]	-0.09 (0.12) [75]	-0.25** (0.11) [78]	-0.15 (0.11) [78]	0.56 (0.47) [78]
MRSA rate	-0.69 (0.60) [76]	0.25 (0.64) [78]	-0.44 (0.59) [80]	0.27 (0.59) [80]	2.73 (2.59) [80]
Surplus	2,780 (2,471) [89]	-728 (2,615) [92]	-1,542 (2,404) [95]	2,697 (2,367) [95]	6,941 (11,443) [95]

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 25th percentile, hospital created through merger or hospital with PFI contract. Standard errors in (parentheses) and number of observations in [brackets]. *Significant at 10%, **significant at 5%, ***significant at 1%

mean residual from Equation 2 for the financial years $t_1^{i,A}$ to $t_n^{i,A}$ when CEO i is observed in hospital A. For length of stay, waiting time, canceled operations, AMI deaths, stroke deaths, FPF deaths, readmissions and MRSA rate we classify as good performers CEOs whose mean residual is at or below the 25th percentile. For technology, job satisfaction, day cases, surplus, admissions and admissions growth we define as good performers CEOs whose mean residual is at or above the 75th percentile. We omit from this analysis input variables (such as beds and labour skills ratios) because it is unclear what would be considered good performance along these dimensions.

Table 12 presents results for linear probability models regressing an indicator of moving to a "problematic" hospital on an indicator of good performance at a CEO's first hospital. There are 14 production measures $\times$ 4 definitions of "problematic", generating a total of 56 coefficient estimates. 7 of these estimates, i.e. 12.5% are statistically significant at the 10% level, but again there is no clear pattern in the direction of association. For 2 production measures better performance immediately prior to arrival is associated with being at problematic hospital, but for 5 measures the association is negative.

We also examined whether CEOs who are viewed by the market as good performers, as measured by their pay effect, were allocated to "problematic" hospitals. Table 13 presents results for our four definitions of "problematic" hospital. The first panel shows that CEOs with large positive pay effects are less likely than CEOs with average pay effects to be assigned to hospitals rated as low quality and more likely to be assigned to hospital rated as at least medium quality.

The second panel of Table 13 compares CEOs by the financial state of the hospitals they are joining. For each year, we determine the 25th and the 75th percentile of the financial surplus variable and then categorise hospitals as low, medium or high surplus. We report the surplus category of the hospital in the year before the CEO arrived there. We see that CEOs with large positive pay effects are less likely than CEOs with with average pay effects to be assigned to hospitals with low surplus and more likely to be assigned to hospitals with medium surplus.

The third panel of Table 13 explores whether more highly paid CEOs are assigned to hospitals created through a merger. We see that CEOs with large positive pay effects are less likely than CEOs with average pay effects to be assigned to a merged hospital. The fourth panel examines whether any of the hospitals in which highly paid CEOs are observed had PFI contracts. In this case there is a some evidence that CEOs with a high pay effect were more likely to be at hospitals which had PFI contracts.

Overall, we find that CEOs viewed by the market as good performers are not more likely to be allocated to "problematic" hospitals. If anything, CEOs with large positive

Table 12: 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
Technology	0.18	0.37	71	0.02	0.46	91	0.04	0.31	95	-0.04	0.43	95
	(0.13)	(0.07)		(0.13)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Admissions	-0.15	0.46	71	0.14	0.43	91	0.19*	0.27	95	-0.01	0.42	95
	(0.13)	(0.07)		(0.12)	(0.06)		(0.11)	(0.05)		(0.12)	(0.06)	
Admissions growth	0.32**	0.33	68	0.15	0.42	88	0.06	0.29	92	0.09	0.39	92
	(0.13)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Length of stay	-0.11	0.44	70	-0.04	0.48	90	0.02	0.31	94	0.10	0.40	94
	(0.14)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Day cases	0.10	0.40 (0.07)	71	0.05	0.45	91	-0.07	0.33	95	-0.10	0.44	95
	(0.14)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Waiting time	-0.11	0.44	70	-0.02	0.45	89	-0.21*	0.38	93	0.04	0.42	93
	(0.14)	(0.07)		(0.12)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
Canceled ops.	0.004	0.42	67	-0.08	0.48	86	0.11	0.28	90	-0.22*	0.48	90
	(0.14)	(0.07)		(0.13)	(0.06)		(0.11)	(0.06)		(0.12)	(0.06)	
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)	
AMI deaths	0.11	0.43	55	0.037	0.53	57	0.14	0.36	61	-0.16	0.53	61
	(0.15)	(0.08)		(0.16)	(0.08)		(0.14)	(0.07)		(0.15)	(0.07)	
Stroke deaths	-0.03	0.47	61	-0.10	0.54	68	-0.17	0.39	72	0.06	0.44	72
	(0.15)	(0.08)		(0.14)	(0.07)		(0.13)	(0.06)		(0.14)	(0.07)	
FPF deaths	-0.08	0.48	61	-0.43***	0.62	68	-0.09	0.37	72	-0.17	0.50	72
	(0.15)	(0.07)		(0.14)	(0.07)		(0.13)	(0.07)		(0.14)	(0.07)	
Readmissions	0.03	0.41	76	-0.01	0.48	75	-0.23*	0.39	78	-0.28	0.49	78
	(0.13)	(0.07)		(0.13)	(0.07)		(0.12)	(0.06)		(0.13)	(0.06)	
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)	
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)	

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 25th percentile, hospital created through merger or hospital with PFI contract. Good performance at the CEO's first hospital is defined as the mean residual for the CEO spell at the first hospital being at or below the 25th percentile for length of stay, waiting time, cancelled operations, AMI deaths, stroke deaths, FPF deaths, readmissions and MRSA rate and as the mean residual being at or above the 75th percentile for technology, job satisfaction, day cases, surplus, admissions and admissions growth. Standard errors in (parentheses). *Significant at 10%, **significant at 5%, ***significant at 1%

pay effects are less likely to be hired by more difficult to manage hospitals during their careers.

Table 13: Allocation of CEOs to different types of hospitals by CEOs' total pay effects

	Total pay effect ≤ £3,705.66	£3,705.66 < effect < £11,370.06	Total pay effect ≥ £11,370.06	Total
Regulator rating (before CEO's arrival) of CEO's lowest rated hospital				
Low quality (0 + */weak + fair)	11 (48%)	28 (65%)	11 (48%)	50 (56%)
Medium quality (**/good)	6 (26%)	11 (26%)	9 (39%)	26 (29%)
High quality (***/excellent)	6 (26%)	4 (9%)	3 (13%)	13 (15%)
Total	23	43	23	89
Surplus category (before CEO's arrival) of CEO's lowest surplus category hospital				
Low surplus (≤ 25th percentile)	17 (71%)	33 (75%)	13 (54%)	63 (68%)
Medium surplus (25th to 75th perc.)	7 (29%)	9 (20%)	10 (42%)	26 (28%)
High surplus (≥ 75th percentile)	0 (0%)	2 (5%)	1 (4%)	3 (3%)
Total	24	44	24	92
CEO ever at a merged hospital				
No	20 (83%)	22 (47%)	15 (63%)	57 (60%)
Yes	4 (17%)	25 (53%)	9 (37%)	38 (40%)
Total	24	47	24	95
CEO ever at hospital with PFI contract				
No	14 (58%)	19 (40%)	8 (30%)	41 (43%)
Yes	10 (42%)	28 (60%)	16 (67%)	54 (57%)
Total	24	47	24	95

The total pay effects are the estimated executive director effects from the total pay regression in Table 9, transformed into deviations from the mean of all estimated executive director effects. The percentiles used to categorise surplus are calculated separately for each financial year to ensure the categorisation is net of year effects.

From these tests we infer that the lack of persistence in the CEO effects in performance does not appear to be due to allocation of good performers to poor hospitals.

8 Conclusions

In this paper we have examined whether CEOs of large public sector organizations have an impact on the performance of those organizations, focusing in particular on large public hospitals. We adopt two approaches to testing whether CEOs have an effect: one that is parametric and exploits CEOs with tenure in at least two organizations and a second one which is an event study looking at the effect on hospital performance of a new CEO compared to hospitals that do not experience this change in the relevant time period.

We find little evidence of CEOs being systematically able to change the performance of these organizations. We also do not find evidence that a change in CEO brings about an improvement (or even just a change) in performance. Our results are robust to several alternative econometric approaches and robustness tests. In contrast, we find evidence of systematic and persistent differences across CEOs in terms of pay. These results do not seem to be due to allocation of better performing CEOs–as measured either by performance in terms of production variables or in terms of their individual pay effect–to worse performing hospitals.

This raises the question of why we find no effect. There seem to be at least two possible explanations for our findings. The first is public sector specific. 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 CEO effects is consistent with a scenario in which top managers simply chase political goals, rather than policies that might actually improve hospital performance, as documented in in-depth qualitative studies (Powell and Davies 2016).[38] In this context, the rational response of an appointed NHS CEO is not necessarily to improve the long-term performance of the hospital, but instead to minimize the amount of bad news that ends up on the Secretary of State's desk: this may explain why there is a CEO effect in remuneration, which is not associated with observed hospital performance, but is associated with receiving public honours. Finally, 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

[38] A recent trade press article argued that bureaucracy and political pressure are the most important negatives for CEOs and more so than for other staff `https://www.hsj.co.uk/workforce/so-what-does-it-take-to-be-a-chief-executive-in-the-nhs/5091689.article`. The article states: "High regulatory burden and external pressures were cited by 60% and 58% of organisation leaders respectively as negative pulls on job satisfaction. External pressures and the burden of regulation remained the top two negative factors on job satisfaction when all senior NHS staff were questioned. However, they were cited as negative influences by fewer than half of respondents in each case, suggesting they weigh more heavily on chief executives than on other staff. ... Sir Robert Naylor [a leading NHS CEO] said recent legislation had ramped up the pressure on NHS chief executives. 'There is a huge process you have to follow so making change is really difficult,' he said. 'If you have to make change to adapt to a new environment, but you are stopped by bureaucracy, then you have to be pretty powerful to drive that through.' "

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 co-ordination and getting a large number of actors, who traditionally have not worked together, to work co-operatively. Put another way, a possible interpretation of our finding is that the organizational inertia of a large hospital is too strong for a CEO to be able to impact performance within the short time period in which they are in office. This, 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 measures of performance (clinical, access, financial) that can be pursued and can in the short-run conflict. It may also be exacerbated in the public hospital sector by the fact that the contracts of clinical staff tend to be much longer than the contracts of the CEOs and by changes to budgets that are the result of changes in the tax base, rather then the underlying demand for the service.

Regardless of what is the underlying driver of our results, they raise concerns about the plausibility of policy approaches that focus on 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 less than two.[39] Coupled with the findings of Tsai et al. (2015) and Bloom, Propper, Seiler and van Reenen (2015) that the management capabilities of middle managers in hospitals are systematically associated wtih better outcomes, our paper suggests that rather than seeking to rapidly change hospital performance through the appointment of a cadre of "superheads", alternative strategies for improvement should instead focus on nurturing and sustaining the skills of middle managers.

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

Appendix A Description of main dataset

Table A-1 provides 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. Foundation trust status takes the value one from the year onwards in which the hospital achieved foundation trust status and zero otherwise. Year of merger takes the value one in the year the hospital was established through merger and zero otherwise. Years since merger takes the value one in the first year after the merger, the value two in the second year after the merger and so on and zero otherwise. Beds is the number of beds.

Table A-1: Variable definitions and data sources

Variable	Definition	Source
Basic pay	Basic remuneration, RPI adjusted (£)	IDS Incomes Data Services and
Total pay	Total remuneration excluding redundancy payments, RPI adjusted (£)	remuneration reports in hospitals' annual reports
(Doctors + nurses)/beds	Ratio of all medical staff and nurses (full-time equivalent) to beds	NHS Hospital and Community Health Service in England
Senior doctors/staff	Consultants, associate specialists, staff grade, registrars as proportion of all staff (%)	workforce statistics, Health and Social Care Information Centre, now NHS Digital
Nurses/staff	Qualified nursing, midwifery, health visiting staff as prop. of all staff (%)	
Contracted out	Contracted out estates and hotel services (%)	Hospital Estates and Facilities Statistics
Technology index	Details in text	Various sources
Beds	Average daily number of available beds	NHS England
Beds growth	$\ln(\text{beds}_t)$ - $\ln(\text{beds}_{t-1})$	
Senior doctors growth	$\ln(\text{sen. docs}_t)$ - $\ln(\text{sen. docs}_{t-1})$	Workforce statistics
Nurses growth	$\ln(\text{nurses}_t)$ - $\ln(\text{nurses}_{t-1})$	
Admissions	Number of admissions (count)	Hospital Episode Statistics:
Admissions growth	$\ln(\text{adm}_t)$ - $\ln(\text{adm}_{t-1})$	Admitted Patient Care
Length of stay	Mean of spell duration, excluding day cases (days)	Hospital Episode Statistics: Admitted Patient Care, Health
Day cases	Proportion of finished consultant episodes relating to day cases (%)	and Social Care Information Centre, now NHS Digital
Waiting time	Mean time waited between decision to admit and actual admission (days)	
Cancelled operations	Operations cancelled for non-clinical reasons (count)	NHS England
Staff satisfaction	Scores from 1 to 5, 1 = dissatisfied, 5 = satisfied, mean	NHS Staff Survey
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
Stroke deaths	Deaths within 30 days of emergency admission for stroke, all ages (%)	Compendium of Population Health Indicators
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
Surplus	Retained surplus/deficit (£000)	Trust Financial Returns

Appendix B Matching quality for non-parametric estimates of CEOs' impact on hospital behaviour and performance

Table B-2: Means of matching variables and other hospital characteristics for treated and control groups and means of exactly matched hospital characteristics: input and throughput measures

				Means of variables measured in $t-1$						Means of vars. measured in t		Exactly matched characteristics			
						Prop. in each category						Teaching			
		Obs.	Unique controls	Beds	Tech-nology	0-14	60-74	75+	male	Year of merger	Years since merger	Major	Minor	Spec. hosp.	Foun-dation trust in $t-1$
(Docs. + nurses)/ beds	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	160	715	0.390	0.051	0.083	0.082	0.443	0.01	0.82				
	Controls	596	347	719	0.376	0.049	0.081	0.083	0.438	0.01	0.98				
Senior docs./staff	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	160	714	0.389	0.051	0.083	0.083	0.443	0.01	0.82				
	Controls	596	347	718	0.376	0.049	0.081	0.084	0.438	0.01	0.98				
Nurses/ staff	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	160	714	0.389	0.051	0.083	0.082	0.443	0.01	0.82				
	Controls	596	347	718	0.376	0.049	0.081	0.084	0.438	0.01	0.98				
Contract-ed out	Treated	145		703	0.373	0.052	0.082	0.085	0.436	0	1.56	0.11	0.11	0.10	0.38
	Controls	145	112	698	0.409	0.050	0.084	0.085	0.447	0.01	1.06				
	Controls	413	239	719	0.400	0.048	0.083	0.087	0.440	0.01	1.35				
Technol-ogy	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	160	714	0.389	0.051	0.083	0.082	0.443	0.01	0.82				
	Controls	596	347	718	0.376	0.049	0.081	0.084	0.438	0.01	0.98				
Beds	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	160	714	0.389	0.051	0.083	0.082	0.443	0.01	0.82				
	Controls	596	347	718	0.376	0.049	0.081	0.084	0.438	0.01	0.98				
Admis-sions	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	159	715	0.389	0.051	0.083	0.082	0.442	0.01	0.84				
	Controls	596	344	719	0.376	0.049	0.081	0.084	0.438	0.01	0.98				
Length of stay	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	159	715	0.389	0.051	0.083	0.082	0.442	0.01	0.84				
	Controls	596	344	719	0.376	0.049	0.081	0.084	0.438	0.01	0.98				
Day cases	Treated	202		725	0.353	0.054	0.081	0.083	0.435	0	1.24	0.12	0.09	0.09	0.27
	Controls	202	156	718	0.393	0.051	0.083	0.083	0.442	0.01	0.86				
	Controls	586	335	723	0.380	0.049	0.081	0.084	0.437	0.01	1.00				

Table B-3: Means of matching variables and other hospital characteristics for treated and control groups and means of exactly matched hospital characteristics: throughput and performance measures

				Means of variables measured in $t-1$						Means of vars. measured in t		Exactly matched characteristics			
			Unique		Tech-	Prop. in each category				Year of	Years since	Teaching		Spec.	Foundation trust
		Obs.	controls	Beds	nology	0-14	60-74	75+	male	merger	merger	Major	Minor	hosp.	in $t-1$
Waiting times	Treated	200		731	0.354	0.051	0.082	0.084	0.436	0	1.26	0.13	0.09	0.09	0.27
	Controls	200	157	725	0.388	0.051	0.083	0.083	0.442	0.01	0.82				
	Controls	583	338	726	0.376	0.049	0.081	0.084	0.438	0.01	1.01				
Cancelled ops.	Treated	202		731	0.355	0.054	0.080	0.083	0.435	0	1.24	0.012	0.09	0.08	0.26
	Controls	202	158	723	0.395	0.051	0.083	0.083	0.443	0.01	0.83				
	Controls	589	342	726	0.381	0.049	0.081	0.084	0.438	0.01	0.99				
Staff satisfaction	Treated	163		713	0.364	0.052	0.082	0.085	0.435	0	1.48	0.11	0.10	0.09	0.34
	Controls	163	126	711	0.403	0.051	0.084	0.084	0.444	0.01	0.99				
	Controls	468	272	727	0.396	0.049	0.082	0.086	0.439	0.01	1.23				
AMI deaths	Treated	143		823	0.370	0.048	0.083	0.088	0.434	0	1.13	0.15	0.05	0	0.21
	Controls	143	107	805	0.395	0.050	0.080	0.084	0.435	0	1.36				
	Controls	424	234	797	0.391	0.048	0.079	0.084	0.433	0	1.30				
Stroke deaths	Treated	168		790	0.369	0.048	0.083	0.089	0.434	0	1.10	0.13	0.07	0	0.23
	Controls	168	130	773	0.393	0.049	0.080	0.085	0.435	0	0.93				
	Controls	505	281	770	0.382	0.048	0.080	0.086	0.433	0	1.09				
FPF deaths	Treated	165		791	0.366	0.048	0.083	0.089	0.434	0	1.12	0.13	0.06	0	0.22
	Controls	165	127	772	0.390	0.049	0.081	0.085	0.435	0	0.95				
	Controls	495	274	769	0.377	0.048	0.080	0.085	0.433	0	1.09				
Readmissions	Treated	172		736	0.345	0.054	0.080	0.083	0.433	0	0.88	0.12	0.08	0.09	0.20
	Controls	172	132	722	0.380	0.053	0.080	0.080	0.440	0	0.84				
	Controls	503	287	726	0.367	0.051	0.079	0.081	0.436	0	0.96				
MRSA rate	Treated	197		729	0.354	0.053	0.081	0.083	0.435	0	1.27	0.12	0.10	0.10	0.28
	Controls	197	156	721	0.396	0.050	0.084	0.082	0.444	0.01	0.82				
	Controls	572	337	728	0.385	0.049	0.082	0.083	0.439	0.01	0.96				
Surplus	Treated	205		722	0.351	0.053	0.081	0.083	0.435	0	1.22	0.12	0.09	0.10	0.27
	Controls	205	160	714	0.389	0.051	0.083	0.082	0.443	0.01	0.82				
	Controls	596	347	718	0.376	0.049	0.081	0.084	0.438	0.01	0.98				

References

Abadie, A.: 2005, Semiparametric difference-in-difference estimators, *Review of Economic Studies* **72**, 1–19.

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

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

Ballantine, J., Forker, J. and 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. and 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. and 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. and Wolfenzon, D.: 2006, Do CEOs matter?, *NYU Working Paper FIN-06-032*, NYU.

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

Besley, T. and 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. and Van Reenen, J.: 2015, Does management matter in schools?, *The Economic Journal* **125**(584), 647–674.

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

Böhlmark, A., Grönquist, E. and 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. and 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.

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

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

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

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

Dhuey, E. and Smith, J.: 2014, How important are school principals in the production of student achievement?, *Canadian Journal of Economics/Revue canadienne d'économique* **47**(2), 634–663.

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

Gaynor, M., Laudicella, M. and 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.

Grissom, J. A., Kalogrides, D. and 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 and Walshe, K.: 2011, *The Future of Leadership and Management in the NHS: No more heroes*, The King's Fund, London.

Heckman, J., Ichimura, H., Smith, J. and Todd, P.: 1998, Characterizing selection bias using experimental data, *Econometrica* **66**(5), 1017–1098.

Heckman, J. T., Ichimura, H. and Tod, P. E.: 1997, Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme, *Review of Economic Studies* **64**(4), 605–654.

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

Lavy, V. and 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.: 1991, Quasi-markets and social policy, *The Economic Journal* **101**(408), 1256–1267.

LeGrand, J.: 2003, *Motivation, agency, and public policy: Of knights and knaves, pawns and queens*, Oxford University Press, Oxford.

Monitor: 2014, The NHS foundation trust code of governance.
URL: *https://www.gov.uk/government/publications/nhs-foundation-trusts-code-of-governance*

Pollitt, C. and Bouckaert, G.: 2000, *Public management reform: A comparative analysis*, Oxford University Press, Oxford.

Powell, A. and Davies, H.: 2016, Managing doctors, doctors managing, *Research report*, Nuffield Trust.

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

Tsai, T. C., Jha, A. K., Gawande, A. A., Huckman, R. S., Bloom, N. and Sadun, R.: 2015, Hospital board and management practices are strongly related to hospital performance on clinical quality metrics, *Health Affairs* **34**, 1304–1311.

Web Appendix

W-1 Predictors of the number of CEOs per hospital over the sample period

Table W-1: 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%

W-2 Estimates of CEO effects using actual and simulated data for remaining input, throughput and performance measures

Tables W-2, W-3 and W-4 present the estimates of Equation 1 using the actual CEO-hospital matches as well as the random CEO-hospital matches for the input, throughput and performance measures that we omit from main body of the paper because of the qualitatively similar results. As for the measures reported in Section 5, when using the actual CEO-hospital matches the F-tests suggest that the estimated CEO effects are jointly statistically significantly different from zero and the proportion of individually statistically significant CEO effects ranges from 26.5% for contracted out to 41.9% for bed growth. And as for the measures reported in Section 5, using the random CEO-hospital matches the F-statistics across the 100 replications are as large as they are for the actual CEO-hospital matches and the proportion of CEO effects that are individually statistically significant is similar to the proportion for the actual CEO-hospital matches.

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 measures with valid variance proportions, the mean variance proportion explained by the CEO effects using the random CEO-hospital matches tends to be close to the variance proportion explained by the CEO effects using the actual CEO-hospital matches.

Table W-2: Estimates of CEO effects for remaining input measures using actual CEO-hospital matches as well as for 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.
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)	27.3 (29.1%)	0.86	2,396	4.5	79.8	5.4	10.3	842.3
(Std. dev.)	(54.1) (n.a., 1.09, 0)	(4.80, 5.12)	(0.002)		(2.03)	(2.95)	(2.15)	(1.32)	(12.2)
Contracted out									
Actual matches	71.9 (<0.001, 68, 176)	18 (26.5%)	0.85	1,645	1.6	79.1	7.7	11.5	535
Random matches:									
Means	35.5 (100%, 67.4, 176)	16.9 (25.1%)	0.85	1,645	1.4	78.2	7.0	13.5	550.0
(Std. dev.)	(35.3) (n.a., 1.52, 0)	(3.30, 4.86)	(0.004)		(1.25)	(3.59)	(3.42)	(1.84)	(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)	31.4 (33.4%)	0.0.95	2,398	4.4	86.9	3.8	4.8	843.6
(Std. dev.)	(71.0) (n.a., 1.09, 0)	(4.91, 5.17)	(0.001)		(1.59)	(1.86)	(1.62)	(0.78)	(12.1)
Beds									
Actual matches	93.6 (<0.001, 95, 224)	31 (32.6%)	0.98	2,398	*-3.7*	*104.7*	*-2.0*	*1.0*	830
Random matches:									
Means	79.9 (100%, 93.8, 224)	28.3 (30.2%)	0.97	2,398	*-1.5*	*99.9*	*0.05*	*1.5*	843.6
(Std. dev.)	(57.6) (n.a., 1.09, 0)	(4.51, 4.78)	(0.001)		*(0.88)*	*(2.21)*	*(1.59)*	*(0.62)*	(12.1)

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 pay 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 (except for (doctors + nurse)/beds), technology index and case mix variables. The results for random CEO-hospital matches are means and standard deviations across 100 replications.

Table W-3: Estimates of CEO effects for remaining input and throughput measures 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)	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.
Beds growth									
Actual matches	24.5 (<0.001, 86, 203)	28 (32.6%)	0.34	2,165	1.7	21.5	4.7	72.1	724
Random matches:									
Means	42.5 (100%, 87.4, 203)	24.8 (28.3%)	0.34	2,165	14.5	7.8	8.2	69.5	765.0
(Std. dev.)	(24.3) (n.a., 1.41, 0)	(3.95, 4.47)	(0.01)		(4.86)	(5.48)	(1.79)	(2.39)	(13.9)
Senior doctors growth									
Actual matches	21.5 (<0.001, 86, 204)	26 (30.2%)	0.26	2,171	16.9	2.4	3.7	76.9	726
Random matches:									
Means	38.9 (100%, 87.4, 204)	21.2 (24.3%)	0.26	2,171	21.4	0.2	4.9	73.4	766.0
(Std. dev.)	(27.5) (n.a., 1.27, 0)	(4.23, 4.82)	(0.004)		(2.48)	(1.08)	(1.38)	(2.85)	(13.3)
Nurses growth									
Actual matches	51.3 (<0.001, 86, 204)	36 (41.9%)	0.16	2,171	1.8	3.0	14.4	80.8	726
Random matches:									
Means	39.4 (100%, 87.4, 204)	22.1 (25.2%)	0.14	2,171	3.1	2.6	8.3	86.1	766.0
(Std. dev.)	(28.7) (n.a., 1.27, 0)	(4.35, 4.96)	(0.01)		(1.54)	(1.72)	(3.01)	(3.51)	(13.3)
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)	26.8 (28.6%)	0.98	2,392	20.4	74.6	2.7	2.3	839.3
(Std. dev.)	(43.1) (n.a., 1.16, 0)	(4.51, 4.81)	(0.0005)		(2.02)	(2.32)	(1.23)	(0.54)	(12.5)
Admissions growth									
Actual matches	27.2 (<0.001, 92, 224)	28 (30.4%)	0.21	2,351	6.4	3.2	8.8	81.5	794
Random matches:									
Means	47.3 (100%, 92.3, 224)	24.7 (26.7%)	0.20	2,351	7.5	2.6	9.2	80.6	821.7
(Std. dev.)	(31.1) (n.a., 1.64, 0)	(4.75, 5.00)	(0.01)		(2.68)	(2.29)	(2.03)	(2.63)	(16.3)

See notes for table W-2

Table W-4: Estimates of CEO effects for remaining throughput and performance measures 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)	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.
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)	27.0 (28.8%)	0.85	2,383	26.1	55.4	7.8	10.8	837.1
(Std. dev.)	(36.9) (n.a., 1.21, 0)	(4.36, 4.60)	(0.003)		(5.94)	(6.80)	(3.18)	(1.73)	(12.7)
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)	24.4 (26.9%)	0.73	2,332	0.11	68.6	9.2	22.0	813.0
(Std. dev.)	(69.5) (n.a., 1.53, 0)	(4.68, 5.17)	(0.005)		(4.30)	(6.09)	(3.90)	(2.01)	(15.7)
Staff 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)	22.3 (30.6%)	0.77	1,838	42.2	29.0	7.0	21.7	
(Std. dev.)	(33.9) (n.a., 1.12, 0)	(4.64, 6.40)	(0.004)		(2.32)	(3.36)	(2.49)	(1.67)	(11.0)
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)	19.7 (30.5%)	0.68	1,965	40.7	20.2	7.3	31.8	552.2
(Std. dev.)	(26.0) (n.a., 2.30, 0)	(3.83, 5.85)	(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)	19.3 (30.1%)	0.49	1,920	21.3	17.8	11.2	49.7	544.3
(Std. dev.)	(20.8) (n.a., 2.33, 0)	(3.54, 5.56)	(0.005)		(1.60)	(2.49)	(3.20)	(2.96)	(21.2)

See notes for table W-2

W-3 Validity of two-step procedure

We assess the validity of the two-step procedure by estimating Equations 2 and 3 for both actual CEO-hospital matches and random CEO-hospital matches. Table W-5 presents the results for the same subset of our input, throughput and performance measures as in Tables 3 and 4.

For most variables, we do not find evidence of any impact of individual CEOs on hospital performance. The results for MRSA rates hint at larger than expected MRSA rates in the first hospital being associated with larger than expected MRSA rates in the second hospital, with the coefficient estimate $\widehat{\delta_2}$ taking the value 0.10. This estimate, however, is not statistically significantly different from zero with a p-value of 0.33. Also, the explanatory power of the average deviations from the expected MRSA rates in the first hospital is very low with an R^2 of 0.01. For AMI deaths the coefficient estimate $\widehat{\delta_2}$ takes the value -0.17, suggesting larger than expected AMI death rates in the first hospital are associated with smaller than expected AMI death rates in the second hospital and vice versa. This coefficient is more precisely estimated with a p-value of 0.12 and the explanatory power of the average deviations from the expected AMI death rates in the first hospital is slightly larger with an R^2 of 0.04. However, the negative coefficient suggests that there is no impact of individual CEOs on AMI death rates.

Turning to the results for the regressions using random CEO-hospital matches, we see that regardless of the input, throughput or performance measure the coefficient estimates $\widehat{\delta_2}$ are very small, with the mean coefficient estimates across the 100 replications ranging from -0.01 to 0.004. The next column shows the proportion of t-tests across our 100 replications that reject the hypothesis that $\widehat{\delta_2}$ is equal to zero when using a significance level of 10%. This rejection frequency is around 10% and therefore close to the nominal level of the test. The explanatory power of the average deviations in input, throughput or performance at the first hospital is very low, with the mean R^2 ranging from 0.01 to 0.02. Overall, applying the two-step procedure to the random CEO-hospital matches generates results that are clearly different from the results for the actual CEO-hospital matches. The results for the random CEO-hospital matches show no impact of CEOs, exactly what we would expect for random matches, suggesting the two-step procedure is valid.

Table W-5: Association between mean of residuals for CEO's spell in first hospital and mean of residuals for CEO's spell in second hospital using actual CEO-hospital matches as well as random CEO-hospital matches

		Coefficient (standard error)	p-value/rejection freq. using 10% signif. level	R^2	Obs.
(Doctors + nurses)/beds	Actual matches	-0.01 (0.15)	0.96	0	94
	Random matches:				
	Means	0.002 (0.15)	9%	0.01	(92.0)
	(Standard dev.)	(0.14, 0.03)		(0.01)	(1.23)
Senior doctors/ staff	Actual matches	0.03 (0.12)	0.80	0	95
	Random matches:				
	Means	0.01 (0.14)	9%	0.01	93.7
	(Standard dev.)	(0.13, 0.02)		(0.02)	(1.09)
Waiting times	Actual matches	-0.01 (0.08)	0.93	0	93
	Random matches:				
	Means	0.004 (0.10)	9%	0.01	91.7
	(Standard dev.)	(0.10, 0.01)		(0.02)	(1.64)
Length of stay	Actual matches	0.05 (0.06)	0.47	0.01	94
	Random matches:				
	Means	-0.001 (0.09)	7%	0.01	92.9
	(Standard dev.)	(0.09, 0.02)		(0.01)	(1.32)
AMI deaths	Actual matches	-0.17 (0.11)	0.12	0.04	61
	Random matches:				
	Means	-0.01 (0.14)	10%	0.02	53.4
	(Standard dev.)	(0.13, 0.03)		(0.03)	(3.25)
Readmissions	Actual matches	0.07 (0.10)	0.47	0.01	78
	Random matches:				
	Means	0.006 (0.13)	6%	0.01	71.0
	(Standard deviations)	(0.11, 0.03)		(0.01)	(1.44)
MRSA rate	Actual matches	0.10 (0.10)	0.33	0.01	80
	Random matches:				
	Means	-0.003 (0.11)	11%	0.01	85.5
	(Standard deviations)	(0.11, 0.02)		(0.02)	(1.64)
Surplus	Actual matches	-0.05 (0.30)	0.87	0	95
	Random matches:				
	Means	0.003 (0.14)	10%	0.01	93.8
	(Standard deviations)	(0.14, 0.04)		(0.02)	1.09

The residuals are from a regression of the input, throughput or performance measure on hospital characteristics, financial year effects and hospital effects. The results for random CEO-hospital matches are means and standard deviations across 100 replications.

W-4 Additional non-parametric estimates of CEOs' impact on hospital behaviour and performance

Table W-12 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.

Table W-12 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-12 shows that the parallel trend assumption for hospital spending on CEO pay is unlikely to 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-6: Changes in input and throughput measures *before* the CEO turnover events analysed in Section 5

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Doctors + nurses/beds	Treated	183	0.23 (0.02)		
	Controls	184	0.19 (0.02)	0.04 (0.02)	0.12
	Controls	536	0.20 (0.01)	0.03 (0.02)	0.12
Senior doctors/staff	Treated	183	0.82 (0.14)		
	Controls	184	0.64 (0.07)	0.19 (0.15)	0.23
	Controls	536	0.70 (0.05)	0.12 (0.11)	0.30
Nurses/staff	Treated	183	-0.32 (0.14)		
	Controls	184	-0.82 (0.13)	0.50 (0.19)	0.01
	Controls	536	-0.67 (0.07)	0.35 (0.15)	0.02
Contracted out	Treated	103	0.023 (1.86)		
	Controls	95	0.602 (1.61)	-0.579 (2.48)	0.82
	Controls	287	1.38 (1.02)	-1.35 (2.03)	0.50
Technology	Treated	183	0.022 (0.004)		
	Controls	184	0.019 (0.004)	0.003 (0.006)	0.60
	Controls	536	0.023 (0.003)	-0.001 (0.005)	0.89
Beds	Treated	183	-34.8 (5.45)		
	Controls	184	-31.4 (5.29)	-3.36 (7.60)	0.66
	Controls	536	-25.1 (2.75)	-9.64 (5.69)	0.09
Admissions	Treated	183	4313 (398)		
	Controls	183	4291 (398)	21.8 (563)	0.97
	Controls	535	4278 (227)	36.4 (452)	0.94
Length of stay	Treated	183	-0.35 (0.05)		
	Controls	183	-0.47 (0.05)	0.12 (0.08)	0.13
	Controls	535	-0.38 (0.04)	0.03 (0.08)	0.72
Day cases	Treated	179	0.66 (0.27)		
	Controls	182	-0.12 (0.33)	0.79 (0.43)	0.07
	Controls	531	0.53 (0.19)	0.13 (0.37)	0.72

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 7 because for some treated observations we do not observe the lagged change in the outcome variable. For details on selection of treated and control observation refer to notes in Table 7. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-7: Changes in throughput and performance measures *before* the CEO turnover events analysed in Section 5

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Waiting times	Treated	177	-9.49 (1.29)		
	Controls	178	-10.0 (1.29)	0.54 (1.82)	0.77
	Controls	525	-9.37 (0.76)	-0.12 (1.50)	0.93
Cancelled operations	Treated	180	-25.0 (17.1)		
	Controls	181	-42.5 (16.2)	17.5 (23.5)	0.46
	Controls	530	-40.2 (9.52)	15.2 (19.1)	0.43
Staff satisfaction	Treated	123	0.009 (0.009)		
	Controls	123	0.004 (0.010)	0.005 (0.013)	0.70
	Controls	348	0.004 (0.005)	0.005 (0.011)	0.64
AMI deaths	Treated	122	-0.65 (0.29)		
	Controls	120	-1.12 (0.33)	0.46 (0.44)	0.29
	Controls	360	-1.03 (0.17)	0.38 (0.34)	0.27
Stroke deaths	Treated	147	-1.97 (0.42)		
	Controls	148	-1.70 (0.38)	-0.27 (0.56)	0.63
	Controls	448	-1.59 (0.21)	-0.38 (0.44)	0.39
FPF deaths	Treated	144	-0.49 (0.25)		
	Controls	145	-0.23 (0.26)	-0.25 (0.36)	0.48
	Controls	438	-0.33 (0.16)	-0.16 (0.31)	0.60
Readmissions	Treated	150	0.71 (0.12)		
	Controls	151	0.67 (0.09)	0.04 (0.15)	0.80
	Controls	445	0.61 (0.05)	0.09 (0.11)	0.41
MRSA rate	Treated	156	-2.69 (0.47)		
	Controls	157	-2.92 (0.46)	0.23 (0.66)	0.73
	Controls	451	-3.17 (0.27)	0.47 (0.53)	0.37
Surplus	Treated	183	-2607 (1359)		
	Controls	184	-2057 (993)	-549 (1682)	0.74
	Controls	536	-1001 (676)	-1606 (1403)	0.25

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 8 because for some treated observations we do not observe the lagged change in the outcome variable. For details on selection of treated and control observation refer to notes in Table 8. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-8: Non-parametric estimates of CEOs' impact on input and throughput measures with potential treated observations limited to the 95 CEOs observed in two hospitals for at least two years each

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Doctors + nurses/beds	Treated	106	0.19 (0.02)		
	Controls	106	0.14 (0.02)	0.05 (0.04)	0.14
	Controls	308	0.18 (0.02)	0.01 (0.03)	0.85
Senior doctors/staff	Treated	106	0.77 (0.13)		
	Controls	106	0.67 (0.16)	0.11 (0.20)	0.60
	Controls	308	0.77 (0.11)	0.00 (0.19)	0.99
Nurses/staff	Treated	106	-0.12 (0.14)		
	Controls	106	-0.21 (0.14)	0.09 (0.20)	0.65
	Controls	308	-0.24 (0.08)	0.11 (0.15)	0.46
Contracted out	Treated	74	1.26 (1.81)		
	Controls	74	0.64 (1.82)	0.61 (2.57)	0.81
	Controls	211	1.06 (1.09)	0.19 (2.14)	0.93
Technology	Treated	106	0.021 (0.006)		
	Controls	106	0.017 (0.005)	0.004 (0.008)	0.63
	Controls	308	0.016 (0.003)	0.047 (0.006)	0.45
Beds	Treated	106	-36.9 (7.53)		
	Controls	106	-15.4 (6.66)	-21.5 (10.1)	0.03
	Controls	308	-25.5 (3.68)	-11.4 (7.66)	0.14
Admissions	Treated	106	4800 (598)		
	Controls	106	4826 (594)	-25.3 (843)	0.98
	Controls	308	4892 (448)	-91.1 (841)	0.91
Length of stay	Treated	106	-0.50 (0.10)		
	Controls	106	-0.38 (0.06)	-0.13 (0.11)	0.27
	Controls	308	-0.32 (0.04)	-0.19 (0.09)	0.04
Day cases	Treated	105	1.21 (0.36)		
	Controls	105	0.16 (0.37)	1.04 (0.52)	0.04
	Controls	305	1.20 (0.26)	0.01 (0.49)	0.99

The maximum number of treated observations is less than 95×2 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 $t-2$. One or 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. 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. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-9: Non-parametric estimates of CEOs' impact on throughput and performance measures with potential treated observations limited to the 95 CEOs observed in two hospitals for at least two years each

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Waiting times	Treated	105	-10.3 (1.78)		
	Controls	105	-10.0 (1.55)	-0.30 (2.36)	0.90
	Controls	305	-10.7 (0.93)	0.40 (1.89)	0.83
Cancelled operations	Treated	105	-26.5 (19.2)		
	Controls	105	4.30 (17.2)	-30.8 (25.8)	0.23
	Controls	303	-14.1 (11.4)	-12.4 (22.4)	0.58
Staff satisfaction	Treated	84	0.010 (0.012)		
	Controls	84	0.040 (0.011)	-0.030 (0.016)	0.06
	Controls	242	0.027 (0.006)	-0.017 (0.013)	0.19
AMI deaths	Treated	79	-0.52 (0.41)		
	Controls	79	-0.90 (0.41)	0.39 (0.58)	0.51
	Controls	233	-0.66 (0.21)	0.14 (0.43)	0.75
Stroke deaths	Treated	90	-2.04 (0.40)		
	Controls	90	-1.09 (0.43)	-0.96 (0.59)	0.11
	Controls	269	-1.54 (0.24)	-0.51 (0.48)	0.29
FPF deaths	Treated	89	-0.08 (0.31)		
	Controls	89	-0.78 (0.29)	0.70 (0.42)	0.10
	Controls	267	-0.46 (0.16)	0.38 (0.33)	0.25
Readmissions	Treated	90	0.54 (0.12)		
	Controls	90	0.50 (0.10)	0.03 (0.16)	0.83
	Controls	264	0.48 (0.06)	0.05 (0.12)	0.65
MRSA rate	Treated	102	-2.67 (0.64)		
	Controls	102	-1.94 (0.56)	-0.73 (0.85)	0.39
	Controls	296	-2.26 (0.35)	-0.41 (0.71)	0.56
Surplus	Treated	106	2235 (2001)		
	Controls	106	3522 (1394)	-1287 (2439)	0.60
	Controls	308	1683 (649)	552 (1611)	0.73

The maximum number of treated observations is less than 95×2 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 $t-2$. One or 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. 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. The change in outcome variable is $y_{j(t+1)} - y_{j(t-1)}$. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-10: Non-parametric estimates of CEOs' impact on input and throughput measures over a period of 3 years instead of 2 years

		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Doctors + nurses/beds	Treated	151	0.30 (0.02)		
	Controls	151	0.29 (0.02)	0.01 (0.03)	0.65
	Controls	433	0.32 (0.02)	-0.02 (0.03)	0.66
Senior doctors/staff	Treated	151	1.27 (0.18)		
	Controls	151	1.03 (0.12)	0.24 (0.22)	0.27
	Controls	433	1.05 (0.08)	0.23 (0.17)	0.19
Nurses/staff	Treated	151	-0.17 (0.17)		
	Controls	151	-0.48 (0.16)	0.31 (0.23)	0.18
	Controls	433	-0.65 (0.10)	0.47 (0.19)	0.01
Contracted out	Treated	98	1.50 (1.48)		
	Controls	98	-0.61 (1.72)	2.11 (2.27)	0.35
	Controls	274	0.48 (1.09)	1.02 (2.03)	0.62
Technology	Treated	151	0.032 (0.006)		
	Controls	151	0.026 (0.006)	0.006 (0.008)	0.46
	Controls	433	0.024 (0.003)	0.008 (0.006)	0.18
Beds	Treated	151	-48.6 (7.06)		
	Controls	151	-36.8 (5.63)	-11.8 (9.02)	0.19
	Controls	433	-32.6 (4.06)	-16.0 (8.04)	0.05
Admissions	Treated	151	6208 (597)		
	Controls	151	6262 (592)	-54 (841)	0.95
	Controls	433	7722 (499)	-1513 (916)	0.10
Length of stay	Treated	151	-0.69 (0.078)		
	Controls	151	-0.56 (0.057)	-0.14 (0.097)	0.14
	Controls	433	-0.53 (0.043)	-0.17 (0.086)	0.05
Day cases	Treated	150	1.23 (0.39)		
	Controls	150	1.83 (0.48)	-0.60 (0.62)	0.34
	Controls	429	1.69 (0.27)	-0.46 (0.51)	0.38

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 $t-2$. One or up to three 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. The change in outcome variable is $y_{j(t+2)} - y_{j(t-1)}$. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-11: Non-parametric estimates of CEOs' impact on throughput and performance measures over a period of 3 years instead of 2 years

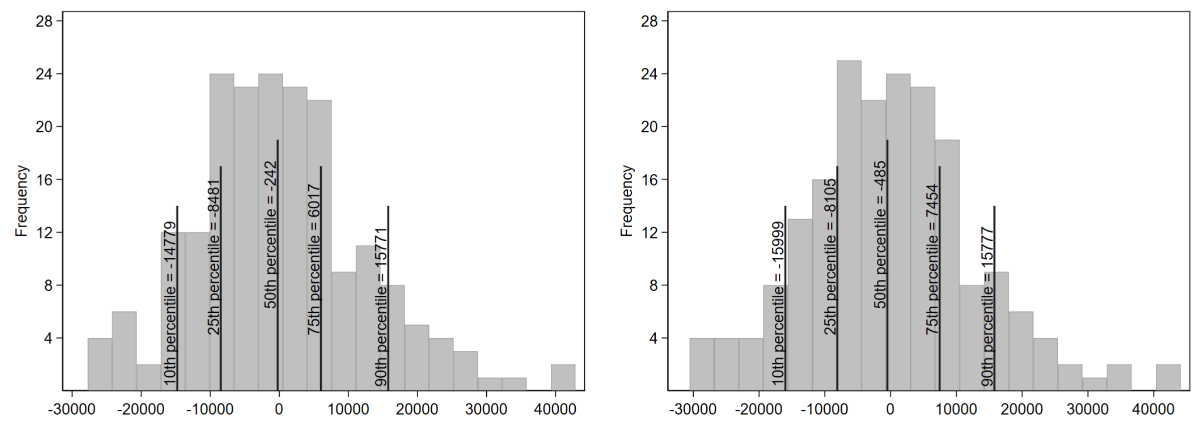
		Obs.	Mean change in variable (std. error)	Difference in mean changes (std. error)	p-value
Waiting times	Treated	144	-18.1 (1.69)		
	Controls	144	-15.8 (1.66)	-2.33 (2.37)	0.33
	Controls	415	-15.8 (0.95)	-2.33 (1.89)	0.22
Cancelled operations	Treated	148	-40.8 (20.3)		
	Controls	148	-3.43 (16.1)	-37.4 (25.9)	0.15
	Controls	424	-21.5 (10.7)	-19.4 (21.8)	0.37
Staff satisfaction	Treated	114	0.023 (0.012)		
	Controls	114	0.043 (0.011)	-0.020 (0.016)	0.22
	Controls	322	0.040 (0.006)	-0.017 (0.012)	0.18
AMI deaths	Treated	111	-0.93 (0.31)		
	Controls	111	-1.22 (0.33)	0.29 (0.46)	0.53
	Controls	320	-1.33 (0.18)	0.39 (0.35)	0.26
Stroke deaths	Treated	129	-2.94 (0.36)		
	Controls	129	-1.49 (0.46)	-1.45 (0.59)	0.01
	Controls	378	-1.85 (0.24)	-1.09 (0.46)	0.02
FPF deaths	Treated	127	-0.85 (0.23)		
	Controls	127	-1.14 (0.23)	0.29 (0.33)	0.37
	Controls	372	-0.79 (0.14)	-0.06 (0.27)	0.83
Readmissions	Treated	120	0.80 (0.12)		
	Controls	120	0.86 (0.09)	-0.06 (0.15)	0.67
	Controls	345	0.99 (0.05)	-0.19 (0.12)	0.10
MRSA rate	Treated	145	-4.20 (0.52)		
	Controls	145	-3.45 (0.58)	-0.75 (0.78)	0.33
	Controls	411	-3.61 (0.32)	-0.59 (0.63)	0.35
Surplus	Treated	151	-45 (1546)		
	Controls	151	-302 (1053)	-257 (1870)	0.89
	Controls	433	-1480 (871)	1435 (1735)	0.41

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 $t-2$. One or up to three 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. The change in outcome variable is $y_{j(t+2)} - y_{j(t-1)}$. Standard error and p-value for difference in means from two-sample t-tests with equal variance.

Table W-12: 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.



(a) Basic pay (b) Total pay

Figure W-1: Deviations of estimated hospital effects in pay from mean of all estimated hospital effects in pay (196 observations)

W-5 Properties of the hospital pay effects

Figure W-1 presents histograms of the estimated hospital effects $\widehat{\psi_{j(i,t)}}$ for both basic pay and total pay. Since the $\psi_{j(i,t)}$ are estimated relative to an arbitrary omitted hospital, we transform the estimates into deviations from the mean of all $\psi_{j(i,t)}$. At 25% of hospitals the executive directors are paid an extra £6,500 or more in basic pay and an extra £7,600 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 £8,000 or more below the average pay package.

We explore 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 are in Table W-13. An important time-invariant hospital characteristic is the region where the hospital is based. We expect the hospital effects to reflect differences in the cost of living across the different regions in England. The first specification in Table W-13 includes only region dummies, with the omitted region being the North West. Thus, the constant of −£6,117 for basic pay is the North West average of the deviations from the mean of all estimated hospital effects for basic 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

Table W-13: Association between estimated hospital effects in pay and time-invariant hospital characteristics

	Basic pay		Total pay		Obs. in each category
North West (omitted category)					
North East	10,443**	7,984**	12,661***	10,144**	9
	(4,418)	(3,894)	(4,649)	(4,120)	
Yorkshire and Humber	8,342**	5,881**	8,430**	5,903*	20
	(3,321)	(2,943)	(3,495)	(3,114)	
West Midlands	3,723	4,140	3,800	4,262	22
	(3,225)	(2,886)	(3,394)	(3,054)	
East Midlands	6,604*	1,099	6,596	908	10
	(4,239)	(3,802)	(4,461)	(4,023)	
East of England	3,046	2,338	4,173	3,443	18
	(3,436)	(3023)	(3,615)	(3,199)	
London	13,681***	11,501***	13,889***	11,653***	38
	(2,782)	(2,464)	(2,928)	(2,607)	
South West	1,594	908	2,388	1,692	19
	(3,376)	(2,964)	(3,553)	(3,136)	
South East	7,118**	7,226***	7,736**	7,858***	26
	(3,071)	(2,731)	(3,231)	(2,890)	
Major teaching hospital		13,638***		14,134***	26
		(2,262)		(2,393)	
Minor teaching hospital		7,307***		7,463***	25
		(2,344)		(2,479)	
Specialist acute		-9,596***		-9,767***	12
		(3,209)		(3,395)	
Specialist orthopaedic		-16,426***		-17,298***	4
		(5,333)		(5,642)	
Constant	-6,117***	-6,797***	-6,538***	-7,250***	
	(2,021)	(1,863)	(2,127)	(1,971)	
R^2	0.15	0.36	0.14	0.35	

The hospital effects are extracted from the regressions reported in Table 9 and transformed into deviations from the mean of all estimated hospital effects. 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%

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 $-£\,6{,}797 + £\,11{,}501 + £\,13{,}638 = £\,18{,}342$ above the sample average of the hospital effects in basic pay. This amount is above the 90^{th} percentile of the distribution of deviations from the mean of all hospital effects displayed in Figure W-1a, suggesting that major teaching hospitals in London have hospital effects above the 90^{th} percentile. Overall, the region dummies, the teaching status dummies and the specialist status dummies jointly explain 35% of the variation in the hospital effects.