

The Impact of Competition on Management Quality: Evidence from Public Hospitals

Nicholas Bloom

Stanford University, NBER and Centre for Economic Performance

Carol Propper

Imperial College, CMPO University of Bristol and CEPR

Stephan Seiler

London School of Economics, Centre for Economic Performance

John Van Reenen

London School of Economics, Centre for Economic Performance, NBER and CEPR

This draft: March 17th 2010

Abstract:

In this paper we examine the causal impact of competition on management quality. We analyze the hospital sector where geographic proximity is a key determinant of competition, and English public hospitals where political competition can be used to instrument for market structure. Since almost all major English hospitals are run by the government, closing hospitals in areas where the governing party has only a small majority is rare due to fear of electoral punishment. We develop a new survey tool to measure management quality and implement this in 61% of all acute hospitals. Our measure of management quality is strongly correlated with financial and clinical outcomes (e.g. survival rates from heart-attacks). More importantly, we find that higher competition (as indicated by a greater number of neighboring hospitals) is positively correlated with increased management quality, and this relationship strengthens when we instrument with local political competition. Adding three more rival hospitals increases the index of management quality by over a standard deviation, which is associated with a 6% reduction in heart-attack mortality rates.

JEL classification: J45, F12, I18, J31

Keywords: management, hospitals, competition, productivity

Acknowledgements: We would like to thank Pedro Castro, John Dowdy, Stephen Dorgan, Robert Huckman, John McConnell, Ben Richardson, Kathy Shaw, Carolyn Whitnall, and participants in seminars at the AEA, LSE, the NBER, Stanford, the Health and Econometrics Conference, and the Department of Health for discussions. We thank the managers and physicians who took part in the survey. Particular thanks are due to our team of interviewers. Financial support is from the HP/EDS Innovation Centre and the ESRC through the Centre for Economic Performance and CMPO.

Corresponding author: j.vanreenen@lse.ac.uk; Centre for Economic Performance, LSE, Houghton Street, London, WC1E 2AE, UK.

In the US and almost every other nation, healthcare costs have been rapidly rising as a proportion of GDP (e.g. Hall and Jones, 2007). Since a large share of these costs are subsidized by the taxpayer, and this share could increase in the US under planned healthcare reforms, there is a great emphasis on improving cost efficiency. Given the large differences in hospital performance across a wide range of indicators (Kessler and McClellan, 2000; Hall et al, 2008; Cutler, Huckman and Kolstad, 2009 and Skinner and Staiger, 2009)¹, one route is through improving the management practices of hospitals. Policy makers have experimented with various ways of increasing effective competition in healthcare to increase productivity. For example, in the UK public National Health Service (NHS) there have been significant reforms in the early 1990s and the early 2000s to introduce greater elements of patient choice and incentives for hospitals to compete for patients.

Economists have long believed that competition is an effective way to improve management and therefore productivity. Adam Smith remarked “monopoly is a great enemy to good management”². Analyzing this relationship is hampered by two factors: first, the endogeneity of market structure and second, credibly measuring management. Identifying the causal effect of competition is challenging, but the fact that exit and entry are strongly influenced by politics in a publicly run healthcare system, like the UK NHS, offers a potential instrumental variable - the degree of political competition. Closing down a hospital is deeply unpopular and since the governing party is deemed to ultimately run the NHS, voters tend to punish this party at the next election if all or part of their local hospital closes down. A vivid example of this was in the 2001 General Election when a government minister was overthrown by a political independent, Dr. Richard Taylor, who campaigned on the single issue of “saving” the local Kidderminster Hospital (where he was a physician) which the government planned to scale down³.

In the UK hospital opening and closures of public hospitals are centrally determined by the Department of Health (the Health Ministry).⁴ Over time and especially since the mid 1990s there has

¹ This variation is not surprising – there is a huge variability in productivity in many other areas of the private and public sector (e.g. Foster, Haltiwanger and Syverson, 2008 and Syverson 2010).

² *The Wealth of Nations*, Book 1, Chapter XI Part 1, p.148

³ See http://news.bbc.co.uk/1/hi/uk_politics/2177310.stm. There is other anecdotal evidence. For example, the Times Online from September 15th, 2006 reported that “A secret meeting has been held by ministers and Labour Party officials to work out ways of closing hospitals without jeopardising key marginal seats....”

⁴ The vast majority of hospital care in the UK is provided in public hospitals. Private hospitals operate in niche markets, specialising in the provision of elective services for which there are long waiting lists in the NHS. Private healthcare funding (including all out of pocket payments) account for only 16.7% percent of UK health care expenditure (Office for National Statistics, 2008).

been rationalization of the number of public hospitals into a smaller number of larger sites, as part of a drive to improve efficiency (Hensher and Edwards, 1999)⁵. If hospitals are less likely to close down in areas because they are politically marginal, there will be a larger number of hospitals in marginal areas than either Labour or Conservative strongholds. Similarly, new hospitals may be more likely to be opened in marginal areas. In either case, in equilibrium, politically marginal areas should benefit from a higher than expected number of hospitals. Clear evidence for this can be seen from Figure 1 which plots out the number of hospitals per person in a political constituency against the winning margin of the governing party (Labour). When Labour's winning margin is small (under 5%) there are about 10% more hospitals than when it or the opposition has a large majority.

Using the share of government-controlled (Labour) marginal constituencies⁶ as an instrumental variable for hospital numbers we find a significant positive impact of greater local hospital competition on management practices. We are careful to condition on a wide range of confounding influences to ensure that our results are not driven by other factors (e.g. financial resources, different local demographics, the severity of patients treated at the hospital, etc.).

A second problem with examining the impact of competition on management is how to measure managerial quality. In recent work we have developed a methodology for quantifying management practices (Bloom and Van Reenen, 2007; Bloom et al, 2009). The measures, covering incentives management, monitoring, target-setting and lean operations were strongly correlated with firm performance. In this paper we apply the same basic methodology to measuring management in the healthcare sector. We implement our methods in interviews across 100 English acute public hospitals (known as hospital trusts) interviewing a mixture of 161 clinicians and managers in two specialities: cardiology and orthopaedics. We cover 61% of all National Health Service (NHS) providers of acute care in England. We show that our management practice scores are correlated with lower mortality rates from AMI⁷ and other surgical procedures, shorter waiting lists and better financial performance. While not causal, this suggests that the management measure has informational content.

⁵ In the hospitals in our data set, 28.5% were part of a merger in the period 1997 – 2006. Mergers in the NHS combine public hospitals with the intention of rationalising capacity. A merger will be followed by closure of hospital sites and or/whole hospitals.

⁶ Each constituency returns a single member of parliament (MP) to the British House of Commons under a first past the post system. The party with a majority of MPs forms the government headed by the Prime Minister.

⁷ Acute myocardial infarction, commonly known as a "heart attack".

Our identification strategy fits into the growing literature on the effect of the political environment on economic outcomes. One strand of the literature compares the incentives of politicians under different political rules. In a majoritarian system, such as the British one, politicians will pay greater attention to regions where there is more uncertainty about the electoral outcome, attempting to capture undecided voters in such “swing states” by devoting greater effort to these states.⁸ In our paper, we exploit political concern over one particular policy, healthcare provision. List and Sturm (2006) also look at a single issue – in their case, environmental policy at state level in the US – and show that when election outcomes are more uncertain politicians use this policy tool to attract undecided voters.

This paper also relates closely to the literature on competition in healthcare. Reforms to the English healthcare system have introduced more competition between hospitals (see Gaynor et al, 2010, for a detailed discussion). There is extensive publicly available information and patients can choose the hospital they wish to receive treatment from. There is no consensus in the literature, however, on the effects of competition on hospital performance⁹, so our paper contributes to a more positive assessment of the role of competitive forces (as in Kessler and McClellan, 2000 for the US or Gaynor et al, 2010 for England¹⁰). Finally, it is linked to the literature on productivity and competition in private sector manufacturing, including papers by Nickell (1996), Syversson (2004) and Schmitz (2005).

The structure of the paper is as follows. The next section discusses the data, Section II describes the relationship between hospital performance and management quality and Section III analyzes the effect of competition on hospital management. Finally Section IV concludes.

I. DATA

The data used for the analysis is drawn from several sources. The first is the management survey conducted by the Centre for Economic Performance (CEP) at the London School of Economics,

⁸ Theoretical models showing this include Lindbeck and Weibull (1987), Persson and Tabellini (1999), Lizzeri and Persico (2001) and Miles-Ferretti et al. (2002). The latter three papers compare majoritarian with proportional elections and find that they lead to different size and compositions of public expenditure (which is due to different groups being targeted). Empirical evidence to support this includes Persson and Tabellini (1999) and Miles-Ferretti et al. (2002). Nagler and Leighley (1992) and Stromberg (2008) establish empirically that candidates allocate relatively more of their election campaign resources to swing states. Clark and Milcent (2008) show the importance of political competition in France for healthcare employment.

⁹ For example, Dranove and Satherthwaite (2000) or Gaynor and Haas-Wilson (1999).

¹⁰ See also Cooper et al (2010)

which includes 18 questions from which the overall management source is computed plus additional information about the process of the interview and features of the hospitals. This is complemented by external data from the UK Department of Health and other health regulators, which provides information on measures of quality and access to treatment as well as hospital characteristics such as patient intake and resources. Finally we use data on election outcomes at the constituency level from the British Election Study. The descriptive statistics for all the relevant variables that are used in our analysis are in Table 1.

I.A. Management Survey Data

The core of this dataset is made up of 18 questions which can be grouped in the following four subcategories: operations (3 questions), monitoring (3 questions), targets (5 questions) and incentives management (7 questions). For each one of the questions the interviewer reports a score between 1 and 5, a higher score indicating a better performance in the particular category. A detailed description of the individual questions and the scoring method is provided in Appendix A.¹¹

To try to obtain unbiased responses we use a *double-blind* survey methodology. The first part of this was that the interview was conducted by telephone without telling the respondents in advance that they were being scored. This enabled scoring to be based on the interviewer's evaluation of the hospital's actual practices, rather than their aspirations, the respondent's perceptions or the interviewer's impressions. To run this "blind" scoring we used open questions (i.e. "can you tell me how you promote your employees"), rather than closed questions (i.e. "do you promote your employees on tenure [yes/no]?"). Furthermore, these questions target actual practices and examples, with the discussion continuing until the interviewer can make an accurate assessment of the hospital's typical practices based on these examples. For each practice, the first question is broad with detailed follow-up questions to fine-tune the scoring. For example, in dimension (1) *Layout of patient flow* the initial question is "Can you briefly describe the patient journey or flow for a typical episode?" is followed up by questions like "How closely located are wards, theatres and diagnostics centres?"

The second part of the double-blind scoring methodology was that the interviewers were not told anything about the hospital's performance in advance of the interview. The interviewers were

¹¹ The questions in appendix A correspond in the following way to these categories. Operations: questions 1-3, Monitoring: questions 4-6, Targets: questions 8-12, Incentives management: questions 7 and 13-18.

specially trained graduate students from top European and U.S. business schools. Since each interviewer also ran 46 interviews on average we can also remove interviewer fixed effects in the regression analysis.

Obtaining interviews with managers was facilitated by a supporting letter from the Department of Health, and the name of the London School of Economics, which is well known in the UK as an independent research university. We interviewed respondents for an average of just under an hour. We approached up to four individuals in every hospital – a manager and physician in the cardiology service and a manager and physician in the orthopaedic service. There were 164 acute hospital trusts with orthopaedics or cardiology departments in England when the survey was conducted in 2006 and 61% of hospitals (100) responded. We obtained 161 interviews, 79% of which were with managers (it was harder to obtain interviews with physicians) and about half in each speciality. Furthermore, the response probability was uncorrelated with observables such as performance outcomes and other hospital characteristics (see Appendix B).¹²

Finally, we also collected a set of variables that describe the process of the interview, which can be used as “noise controls” in the econometric analysis. These included: interviewer fixed effects, the position of the interviewee (clinician or manager), and his tenure in the post. Including these controls helps reduce residual variation.

I.B. Hospital competition and Political Competition

Since there are costs from treating patients far from where they live, healthcare competition always has a strong geographical element. Our main competition measure is simply the number of other public hospitals within a given “catchment area” for each hospital. We show experiments with a Herfindahl index as well which takes activity in the hospital into account, but the market shares are more likely to be endogenous, so we prefer the simpler measure. Our baseline results use a 30km radius (about one hour’s drive) around the hospital, but we also report robust results when using wider market definitions such as 20km or 40km radius instead¹³.

¹² The only variable for which we find an effect in the sample selection regressions is expenditure per patient. As this variable is only used as a sensitivity check we are not too concerned about this.

¹³ We use the number of public hospitals, as private hospitals generally offer a very limited range of services (e.g. they do not have Emergency Rooms).

We use data on outcomes of the national elections at the constituency level from the British Election Study. We observe the vote shares for all parties and use these to compute the winning margin. We define a constituency to be marginal if the winning margin is below 5% (we also show robustness to other thresholds such as 3% or 7%). As hospitals usually have a catchment area that comprises several constituencies we use the share of marginal constituencies in a 30 km radius of the hospital as our main measure of political competition to match the hospital competition measures.

Note that the typical hospital in the UK treats about 72,000 patients a year while the typical political constituency has about 70,000 voters. So the closure of a hospital in a marginal constituency has an important effect on potential voters, increasing the likelihood of the Government losing that constituency in the next election. In other constituencies where the Government has a large lead over (or lag far behind) opposition parties there are lower incentives to avoid hospital closures, as changes of a few percentage points in voting will not alter parliamentary outcomes given the “first past the post” electoral system.¹⁴ We exploit this combination of public hospitals and central controlled hospital closures to generate a quasi-experiment for hospital numbers.

There are three main parties in the UK (Labour, Conservative and Liberal Democrat). We distinguish between marginal constituencies which are controlled by the governing party (Labour) and Opposition parties. We would expect the strongest effects to be on the Labour controlled marginal seats and test for this in the results.¹⁵

Our key instrumental variable is therefore the share of Labour marginal constituencies (constituencies where Labour won, but by less than 5%) from the 1997 election. We use Labour marginals in 1997 since seats which were marginal in the 1997 election were typically perceived as marginal from the mid-1990s until after the early 2000s. The reason is Labour’s victory in 1997 was forecasted from the mid-1990s onwards and their polling margin from 1997 was relatively constant

¹⁴ Britain’s “first past the post” system means that the party with the highest vote share in each constituency wins that constituency. In a proportional representation political system this incentive to keep hospitals open in marginal constituencies does not operate as Governments care about total votes.

¹⁵ There are two reasons for this. First, Labour was the part in power so hospital closures were politically more associated with their Members of Parliament. Second, this period coincided with Tony Blair’s ‘honeymoon’ period in power in which Labour’s popularity were at an all time high, so its marginals were more at risk than opposition marginals as Labour’s vote share trended downwards as its early popularity eroded.

until the early 2000s.¹⁶ Similar results occur if we use a definition of marginality from later elections.

In some regressions we also condition on a flexible polynomial in the Labour vote share and identity of the winning party as this could reflect some unobservables correlated with health status (and therefore number of hospitals).

I.C. Hospital Performance Data

Productivity is difficult to measure in hospitals, so regulators and researchers typically use a wide range of measures¹⁷. The clinical outcomes we use are the mortality rate for emergency admissions for (i) AMI (acute myocardial infarction) and (ii) surgery.¹⁸ We choose these for four reasons. First, regulators in both the USA and the UK use selected death rates as part of a broader set of measures of hospital quality. Second, using emergency admissions helps to reduce selection bias because elective cases may be non-randomly sorted towards hospitals. Third, death rates are well recorded and cannot be easily “gamed” by administrators trying to hit targets. Fourth, heart attacks and overall emergency surgery are the two most common reasons for admissions that lead to deaths.

As a measure of access to care we use the size of the waiting list for all operations (long waits have been an endemic problem of the UK NHS and of considerable concern to the general public, Propper et al, 2008b). As another quality marker associated with hospital cleanliness we use MRSA infection rates (“superbugs”).¹⁹ We use the hospitals’ operating margin as a measure for their financial efficiency and the average intention of staff intending to leave in the next year as an indication of worker job satisfaction. All of these measures have been used by the UK government to rate NHS hospitals in England. Finally, we use the UK Government’s Health Care Commission ratings which represent a composite performance measure across a wide number of indicators. The Health Care

¹⁶ The Conservatives won a narrow (21 seat) victory in 1992, but became increasingly unpopular from the mid 1990s onwards, particularly after the election of Tony Blair as leader of the Labour party in 1994. So by 1995 Labour was far ahead of the Conservatives in the polls and maintained this lead into the 1997 General Election. Note in the 2001 election Labour had a very similar margin of victory (167 seats) as the 1997 election (179 seats). They still won the 2005 election but with a reduced margin (66 seats), with their popularity declining slowly between 2001 and 2005 (Crewe, 2005). There is a high correlation between the share of marginal Labour constituencies in 1997 and 2001 (0.73).

¹⁷ See for example <http://2008ratings.cqc.org.uk/findcareservices/informationabouthealthcareservices.cfm>

¹⁸ Examples of the use of AMI death rates to proxy hospital quality include Kessler and McClellan (2000), Gaynor (2004) and, for the UK, Propper et al (2008a). Death rates following emergency admission were used by the UK regulator responsible for health quality in 2001/2. The AMI mortality rate used here are for all deaths within 28 days after admission. Mortality from emergency surgery are for all deaths within 30 days of admission.

http://www.performance.doh.gov.uk/performance/2002/tech_index_trusts.html

¹⁹ MRSA is Methicillin-Resistant Staphylococcus Aureus.

Commission rates hospitals along two dimensions of “resource use” and “quality of service” (measured on a scale from 1 to 4).²⁰

I.D. Other Controls

First, we control for patient mix (case-mix) by controlling for the age/gender profile of total admissions at the hospital level in all of the regressions.²¹ We also control for the total number of admissions to allow for the fact that better hospitals may attract more patients and raise their quality if there are economies of scope or scale. Second, we control for the health of the population in the hospital’s catchment area by using the age distribution (11 bins) and the overall mortality rate. We also control for population density.

Third, we control for resources of the hospitals, which are all derived from general taxation. The purchasers of health care cover a defined geographical area and are allocated resources on the basis of formula that measures need for healthcare (essentially, the demographics of the area the hospital is located in). The purchasers use these resources to buy healthcare from hospitals, at fixed national prices, for their local population. Purchasers do not own hospitals nor are vertically integrated with hospitals. This system is intended to ensure resources are neither used to ‘prop up’ poorly performing local hospitals nor are subject to local political influence. However, we are concerned to ensure that we control for the impact of resources as they may affect both performance and quality of management and to ensure that our instrumental variables results are not driven by resources that may be associated with political marginality. We therefore considered a large set of measures of the quantity and quality of resources available to the hospital level as controls. These include whether hospitals have been given greater autonomy from central control (“Foundation Trusts”), the number of sites, the age of the hospital’s building as a proxy for the quality of capital and expenditure per patient as a direct measure of funding.²²

²⁰ We use the 2006 values as these are coincident with the timing of the survey and average across the two measures. We also report experiments where we disaggregate the index and construct our own (re-aggregated) index. See Appendix B for more details on the construction of this Pseudo-HCC index.

²¹ Specifically we have 11 age categories for each gender (0-15, 16-45, 46-50, 51-55, 56-60, 61-65, 66-70, 71-75, 76-80, 81-85, >85), so up to 22 controls. These are specific to the condition in the case of AMI and general surgery. For all other performance indicators we use the same variables at the hospital level. Hall et al (2008) show that in the English context the age-gender profile of patients does a good job of controlling for case-mix.

²² We also tried to include further control variables (results are not reported): a dummy for whether a hospital is a teaching or a specialist hospital, total hospital employment, the number of acute beds, the number of medical staff and doctor vacancy rates. The results are not sensitive to including these additional variables.

I.E. Preliminary Data Analysis

The management questions are all highly correlated (see Bloom and Van Reenen, 2007) so we will usually aggregate the questions together either by taking the simple average (as in the figures) or by z-scoring each individual question and then taking the z-score of the average across all questions (in the regressions).²³

Figure 2 divides the HCC score into quintiles and shows the average management score in each bin. There is a clear upward sloping relationship with hospitals that have higher management scores also enjoying higher HCC rankings. Figure 3 plots the entire distribution of management scores for our respondents (in the upper Panel A). There is a large variance with some well managed firms, and other very poorly managed. In Panel B we present a comparison between hospitals and UK manufacturing firms²⁴. Hospitals clearly have lower management scores than manufacturing firms, particularly for incentives management as they have weaker links between performance and pay, promotion, hiring and firing.

II HOSPITAL PERFORMANCE AND MANAGEMENT PRACTICES

Before examining the impact of competition on management practices we will first check that the management score is robustly correlated with external performance measures. This is not supposed to imply any kind of causality. Instead, it merely serves as an “external validity” check to see whether a higher management score is correlated with a better performance.

We estimate regressions of the form:

$$y_i^k = \alpha M_{ij} + \beta' x_{ij} + u_{ij}$$

where y_i^k is performance outcome k (e.g. AMI mortality) in hospital i , M_{ij} is the average management score of respondent j in hospital i , x_{ij} is a vector of controls and u_{ij} the error term.

Since errors are correlated across respondents within hospitals we cluster our standard errors at the

²³ Factor analysis confirms that there is one dominant factor that loads heavily and positively on all questions. As with the earlier work, there is a second factor that loads positively on the incentives management questions, but negatively on the monitoring/operations questions. This suggests that there is some specialization across hospitals in different forms of management.

²⁴ To make the samples somewhat comparable we keep only establishments who have between 50 and 5,000 employees and who are domestically owned (i.e. we drop multinationals from the manufacturing sample). Furthermore, in both panels we are using the average management score from only 16 comparable questions, because two questions on lean manufacturing are difficult to compare across sectors (questions 1 and 2 in Appendix A)

hospital level (they are also robust to heteroscedacity).²⁵ We present the performance and management measures in z-scores so the tables can be read as the association of a one standard deviation of management with the outcome. We consider disaggregating the 18 questions below, but our standard results simply z-score each individual question, average these into a composite and then z-score this average. In terms of timing, we use the 2005/6 average outcomes in the year to be consistent with the management survey.²⁶

Table 2 shows results for regressions of each of the performance measured on the standardized management score. The management score in the top row (A) is calculated as the average of the eighteen questions in the survey. The other rows show results based on the four different categories of questions. Looking across the first row higher management scores are associated with better hospital outcomes across all the measures and this relationship is significant in every case. This immediately suggests our measure of management has informational content.

Looking in more detail, in the first column of Table 2 we present the AMI mortality rate regressed on the management score controlling for a wide number of confounding influences.²⁷ High management scores are associated with significantly lower mortality rates from AMI: a one standard deviation increase in the management score is associated with just under 0.044 of a standard deviation fall in AMI mortality (with the strongest association being for incentives management). In absolute terms this corresponds to a decrease in the mortality rate of a third of a percentage point, which is about 130 less deaths a year. Column (2) examines death rates from all emergency surgery again showing a significant correlation with management quality.²⁸ Column (3) shows better managed hospitals tend to have significantly lower waiting lists and significantly lower MRSA infection rates (column (4)). The financial performance measured by the hospital's operating margin is higher when hospitals have higher management scores in column (5).²⁹ Column (6) indicates that

²⁵ Furthermore we weight the observations with the inverse of the number of interviews conducted at each hospital. This gives equal weight to each hospital in the regressions.

²⁶ We have quarterly data from 2001 until 2006. We also tried using longer time averages going back to 2001 in an effort to assess the importance of transitory measurement error. The qualitative results were similar, but actually tended to weaken as we used years further away from the date of the management survey.

²⁷ As is standard we drop observations where the number of cases admitted for AMI is low because this leads to large swings in observed mortality rates. Following Hall, Propper and Van Reenen (2008) we drop hospitals with under 150 cases of AMI per year, but the results are not sensitive to the exact threshold used.

²⁸ We exclude two specialist hospitals from this regression as they are difficult to compare to the rest in terms of all emergency admissions.

²⁹ We find qualitatively similar results when using other measures of cost efficiency such as income per spell or income per bed. The operating margin is influenced by both revenue and costs per spell. As the income side is fixed (hospitals

higher management scores are also associated with job satisfaction (a lower probability of the average employee wanting to leave the hospital). In the final two columns we use composite measures from the Health Care Commission (HCC) and compute a “pseudo HCC rating” by attempting to reverse engineer the process by which the original rating was calculated (see Appendix B). The management practice score is significantly and positively correlated with both of these measures.

The lower panel of Table 2 repeats the exercise using the different categories of management practice questions, where each row is an individual regression. The results are very similar although the coefficients are less precisely estimated.³⁰ This suggests that averaging over different questions helps to reduce noise. Different categories are more strongly correlated with different performance measures in an intuitive way. For example “Lean Operations” has the most explanatory power for MRSA infection rates and a high “Incentives Management” score significantly lowers the staff’s intention to leave the job.

Overall, the Table 2 is reassuring in that our measure of management practices is positively associated with superior hospital outcomes across a wide range of performance indicators.³¹

III MANAGEMENT PRACTICES AND HOSPITAL COMPETITION

III.A Basic Results

Is one reason for the large variation in management practices across hospitals seen in Figure 3 variation in levels of competition? There is an extensive discussion surrounding competition and performance in healthcare and other sectors suggesting higher levels of competition is associated with superior management practices (e.g. Fabrizio, Rose and Wolfram 2009; Kessler and McClellan, 2000; Nickell, 1996 and Schmitz 2005). To investigate this we start in column (1) of Table 3 by considering the number of other hospitals in the catchment area as a measure of competition. There is a positive and significant conditional correlation between management quality and the number of

receive a fixed national payment per type of case, known as Payment by Results and similar to the US fixed payment per DRG system), the operating margin (as well as other similar measure such as income per spell etc.) is effectively a measure of costs.

³⁰ We also examined decomposing the management score even further. When regressing the scores for individual questions on the HCC rating, 7 out of 18 questions are significant at the 10% level (all seven are also significant at the 5% level and 2 are significant at the 1% level).

³¹ Our results are also consistent with McConnell et al. (2009) who use the Bloom and Van Reenen (2007) methodology to collect management data on 147 US addiction treatment programs, finding a positive management performance relationship.

rival hospitals. Increasing the number of rival hospitals from zero to the mean (7) is associated with an increase in management quality of 0.84 of a standard deviation.

To address the endogeneity concern we use the political instrumental variable described above - the degree to which a hospital is located in a politically marginal area held by the governing party (Labour). Column (2) contains the first stage and column (3) the second stage. In Column (2) the share of Labour-controlled local marginal constituencies is extremely significant in explaining increased total hospital numbers. There are on average over five more hospitals than expected around a hospital in an area that is politically marginal and controlled by the governing party compared to one that is not. In Column (3) we look at the second stage effect of the number of competing local hospitals on management practices and find a positive effect that is significant at the 10% level. According to column (3) adding an extra competitor increases the index of management quality by over one third of a standard deviation (0.361). An alternative measure of competition is to use the Herfindahl-Index, an inverse measure of competition which ranges from 0 (very competitive) to 1 (monopoly). Including this instead of the number of hospitals we obtain a coefficient (standard error) of -1.148 (0.365), which is consistent with column (4).

The specification in columns (1) through (3) only contains very basic controls (population density and age, interviewer dummies and hospital type), so a concern is that the relationship is driven by omitted variables. Consequently we include a richer set of covariates such as area mortality rates, the age and gender mix of hospital patients, etc. in columns (4) to (6).³² Furthermore, we include linear terms in the share of Labour votes and the identity of the winning party in case this also reflects unobserved demographics (these variables are insignificant). The full set of results are in Table B3, but the coefficients on our key variables are little changed by these additional covariates, in fact the second stage coefficient in column (6) is 0.543, slightly stronger than in column (3)³³

Column (7) shows an alternative first stage where we also include marginal constituencies controlled by the opposition parties. Although the coefficient on this variable is positive suggesting that these are also likely to have more hospitals, it is smaller and insignificant at conventional levels. This is

³² The set of control variables used in this specification is identical to the ones used in table 2, except for the additional controls for area demographics and population density.

³³ Dropping the 20 London hospitals does not alter these findings. The instrument remains significant at the 5% level in the first stage and the coefficient on competition is 0.756 with a standard error of 0.369 in the second stage. The number of observations is 141 (instead of 161) in this regression.

consistent with our interpretation that marginals controlled by the governing party are the ones with most political saliency. If we just used marginality regardless of the controlling party, we obtain a coefficient of 2.283 with a standard error of 0.897 in the first stage. In any case, when we use both instruments from column (7) in the second stage the results are very similar to just using Labour marginals (see column (8)).

Finally, although our focus here is on the impact of competition on management quality, we could also consider the impact on more direct measures of hospital performance. As discussed earlier the key indicator of hospital quality used in the health economics literature is the mortality rates from emergency AMI. We present OLS results in column (7) which indicates that hospitals facing more competition have significantly fewer deaths³⁴. Column (8) uses our IV strategy and indicates that there appears to be a causal effect whereby increasing the number of rival hospitals from zero to the mean of seven reduces death rates by 0.84 of a standard deviation.

III.B Robustness

As noted earlier, none of the qualitative results depend on the precise thresholds used for catchment area or definition of political marginal. Using a 40km catchment area instead of the baseline 30km shows slightly stronger results (a coefficient on competition of 0.849 with a standard error of 0.337). Using a 20km catchment area generates a coefficient (standard error) on competition of 0.548(0.294) in IV. Using a 3% (instead of 5%) threshold for marginality reduced the coefficient (standard error) on competition to 0.321(0.158) in the IV estimates and increasing it to 7% magnified the coefficient (standard error) on competition to 1.068 (0.482). We also considered adding higher order controls for Labour's vote share or dropping Labour's vote share completely with robust results.³⁵

³⁴ Running the same OLS regressions but using each of the other seven performance outcomes in Table 2 as a dependent variable reveals that competition is associated with better performance in every case. However, competition is only significant for AMI survival rates.

³⁵ Using a squared and a cubic term for Labour's vote share in addition to the linear one leads to a coefficient (standard error) on competition of 0.551 (0.208). Dropping the Labour vote share completely yields a coefficient of 0.534 (0.220). We also run the first stage of our IV specification using the number of private hospitals as dependent variable. We find that marginality is insignificant in this case. This constitutes another piece of evidence that our marginality measure is not picking up unobserved area health status.

It is possible that marginality is associated with higher funding. As we noted above, funding for healthcare is allocated on the basis of need and is a separate and more transparent process than hospital exit and entry, so there is no automatic association between funding and marginality. However, it is possible that lobbying by a marginal politician could lead to greater funding. We therefore added controls for funding (expenditure per patient) into the regression. The coefficient for this variable is insignificant in both stages and does not alter the coefficient on competition³⁶. We also controlled for the age of the hospitals' buildings to test whether marginal constituencies received more resources in terms of newer capital equipment. In fact we find the contrary to be true: in marginal constituencies hospital buildings tend if anything to be older, presumably because hospital closures are rarer.³⁷

III.C Discussion

Why is the IV estimate of competition so much larger than the OLS estimate? Some of this might be due to attenuation bias or a LATE interpretation. But more plausibly note that the reverse causality problem is likely to bias OLS towards zero as it is likely that hospitals in the neighbourhood of a high performance hospital are more likely to be closed down. The closure is economically and politically easier to justify if patients have a good substitute due to the presence of the high performance hospital. Because of this, a higher management score could generate a lower number of competing hospitals, just as in the standard case where a very efficient firm will tend to drive weaker firms from the market. This would lead OLS estimates to be biased towards zero as we observe in Table 3. Of course, there may be other potential biases going in the opposite direction due to omitted variables.

There are several routes by which competition could improve management practices. The first is simply through competition for patients. When a General Practitioner (the local "gatekeeper" for patients) refers a patient to a hospital for treatment she has the flexibility to refer the patient to any local hospital. Having more local hospitals gives greater choice for General Practitioners and so greater competition for hospitals. Since funding follows patients in the NHS, hospitals are keen to win patient referrals. Hospitals which do not attract enough patients potentially face closure (more likely in non-politically marginal constituencies) or having their CEO and senior management

³⁶ The coefficient (standard error) on the number of hospitals is 0.545 (0.256) and the first stage coefficient (standard error) on the marginality variable is 4.237 (1.346).

³⁷ Including building age, the coefficient (standard error) on the number of hospitals is 0.513 (0.350) and the first stage coefficient on the marginality variable 3.673 (1.502).

replaced (more likely in political marginals). The second mechanism is yardstick competition: with more local hospitals CEO performance is easier to evaluate because yardstick competition is stronger.

Finally, it might be that a greater number of hospitals improve management quality not through competition, as we have assumed, but rather via input markets. Good managers will find markets with a higher density of hospitals to be a more attractive labor market. Hospitals with more rivals will therefore be able to hire better managers, who will help to increase the quality of management practices. We think this is less likely as managers are relatively mobile across England.

VI CONCLUSIONS

In this paper we have described a new methodology for quantifying the quality of management practices in the hospital sector. We have implemented this survey tool on almost two thirds of acute hospitals in England. We found that our measure of management quality was robustly associated with better hospital outcomes across mortality rates and other indicators of hospital performance. This is consistent with Bloom and Van Reenen's (2007) work in the manufacturing sector.

We then exploit the UK's centralized public hospital system to provide an instrumental variable for hospital competition. We use the share of marginal political constituencies around each hospital as an instrument for the number of nearby competing hospitals. This works well because in the UK politicians rarely allow hospitals in politically marginal constituencies to close down, leading to higher levels of hospital competition in areas with more marginal constituencies. We find in both OLS and 2SLS (using our political instrument) that more hospital competition leads to improved hospital management. This suggests public sector competition is useful for improving management practices in the public sector.

In terms of future work, a drawback of our paper is that it is cross sectional since the management data is only available for a single year. We are collecting a second wave of the panel, however, which will enable us to investigate whether recent policy changes encouraging more competition have had an effect on hospital performance. Second, it would be extremely interesting to expand our sample to look at healthcare management in other countries. We have piloted some work along these lines and plan to implement this in the US and other nations. Finally, examining how hospitals of

different management quality and ownership respond differentially to shocks could be very revealing (Duggan, 2000).

REFERENCES

- Besley, Timothy, Torsten Persson and Daniel M. Sturm (2005) “Political Competition and Economic Performance: Theory and Evidence from the United States”, NBER Working Paper No. W11484.
- Bloom, Nicholas and Van Reenen, John (2007) “Measuring and Explaining Management practices across firms and nations”, *Quarterly Journal of Economics*, Vol. 122, No. 4: 1351–1408.
- Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen (2009) “Does Management Matter? New Empirics and old theories”, LSE/Stanford mimeo
- Clark, Andrew E. and Carine Milcent (2008) “Keynesian Hospitals? Public Employment and Political Pressure”, Paris School of Economics Working Paper No. 2008 – 18.
- Cooper, Zack, Stephen Gibbons, Simon Jones and Alistair McGuire (2010) “Does Hospital Competition Save Lives? Evidence from the English Patient Choice Reforms”, LSE mimeo
- Crewe, Martin (2005) “The opinion polls: the election they got (almost) right”, *Parliamentary Affairs*, 58 (4) pp 684-698.
- Cutler, David M. and Jill R. Horwitz (1999) “Converting Hospitals from Not-For-Profit to For-Profit Status: Why and What Effects?” in D. Cutler (ed.) *The Changing Hospital Industry: Comparing Not-For-Profit and For-Profit Institutions*, Chicago: University of Chicago Press and NBER.
- Cutler, David; Huckman, Robert and Kolstad, Jonathan, (2009) “Input constraints and the efficiency of entry: lessons from cardiac surgery”, forthcoming the *American Economic Journal: Economic Policy*.
- Department of Health (2008) *High Quality for All: NHS Next Steps Review Final Report*.
- Dranove, David and Mark Satterthwaite (2000) “The industrial organization of healthcare markets” in Culyer, A. and Newhouse, J. (eds) *The Handbook of Health Economics*, Amsterdam: North Holland
- Duggan, Mark (2000) “Hospital Ownership and Public Medical Spending” *Quarterly Journal of Economics*, November 1243-1374.
- Duggan, Mark (2002) Hospital Market Structure and the Behavior of Not-For-Profit Hospitals *The RAND Journal of Economics*, Vol. 33, No. 3 (Autumn, 2002), pp. 433-446.
- Farrar, S., Sussex, J., Yi, D., Sutton, M., Chalkley, M., Scott, A. and Ma, A (2007) “A National Evaluation of Payment by Results” Report to the Department of Health, November 2007.
- Foster, Lucia, John Haltiwanger, and Chad Syverson (2008) “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, 98(1), 394-425.

- Gaynor, Martin (2004) "Competition and Quality in Health Care Markets. What Do We Know? What Don't We Know?" *Economie Publique* 15: 3-40.
- Gaynor, Martin and Haas-Wilson, Deborah (1999) "Change, Consolidation and Competition in Health Care Markets", *Journal of Economic Perspectives*, 13, 141-164
- Gaynor, Martin, Propper, Carol and Serra, Rodrigo Morrenes (2010) "Death by Market Power: Reform, Competition and Patient Outcomes in the NHS. Mimeo, CMPO, Bristol University.
- Hall, Emma, Propper, Carol and Van Reenen, John (2008) "Can Pay Regulation Kill? Panel Data evidence on the effect of Labor markets on hospital performance" NBER Working Paper No. 13776.
- Hall, Robert and Jones, Chad (2007) "The Value of Life and the Rise in Health Spending" *Quarterly Journal of Economics*, 122(1), 39-72.
- Healthcare Commission (2006) "The annual Health Check in 2006/2007". http://www.healthcarecommission.org.uk/_db/_documents/The_annual_health_check_in_2006_2007_assessing_and_rating_the_NHS_200609225143.pdf.
- Hensher, Martin and Edwards, Nigel (1999) "Hospital provision, activity and productivity in England and Wales since the 1980s" *British Medical Journal*, vol 319, pp. 911-914.
- Jacobs, R., Martin, S., Goddard, M., Gravelle, H., and Smith, P. (2006) "Exploring the determinants of NHS performance ratings: lessons for performance assessment systems" *Journal of Health Services Research and Policy*, 7(4), 211-217.
- Kessler, Daniel P. and Mark B. McClellan (2000) "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics* 115 (May): 577-615.
- Lindbeck, A., J. W. Weibull (1987) "Balanced-Budget Redistribution as the Outcome of Political Competition", *Public Choice*, 52, 273-297.
- List, John A. and Daniel M. Sturm (2006) "How Elections Matter: Theory and Evidence from Environmental Policy" *Quarterly Journal of Economics*, 121(November): 1249-1281.
- Lizzeri, A. and N. Persico (2001) "The Provision of Public Goods Under Alternative Electoral Incentives", *American Economic Review*, Vol. 91(1), 225-239.
- McConnell, John, Hoffman, Kim, Quanbeck, Andrew and McCartney, Dennis, (2009) "Management practices in substance abuse treatment programs" forthcoming *Journal of Substance Abuse Treatment*.
- Milesi-Ferretti, G. M., R. Perotti and M. Rostagno (2002) "Electoral Systems and Public spending" *Quarterly Journal of Economics*, Vol.117(2), 609-657.
- Nagler, J., J. Leighley (1992) "Presidential Campaign Expenditure: Evidence on Allocations and Effects", *Public Choice*, 73, 319-333.

Nickell, Steve (1996) "Competition and Corporate Performance" *Journal of Political Economy*, CIV (4), 724-746.

Office for National Statistics (2008), UK Centre for the Measurement of Government Activity, *Expenditure on Healthcare in the UK*, April 2008, <http://www.statistics.gov.uk/articles/nojournal/ExpenditureonHealth08.pdf>.

Persson, T., G. Tabellini (1999), "The Size and Scope of Government: Comparative Politics with Rational Politicians", *European Economic Review*, 43, 699-735.

Propper, Carol, Simon Burgess and Denise Gossage (2008a) "Competition and Quality: Evidence from the NHS Internal Market 1991-99" *Economic Journal* 118, 138-170.

Propper, Carol, Matt Sutton, Carolyn Whitnall, and Frank Windmeijer (2008b) "Did targets and terror reduce waiting times for hospital care in England" *The BE Journal of Economic Analysis and Policy*, Issue 8, 2, Paper 5.

Schmitz, J. (2005). "What Determines Productivity? Lessons from the Dramatic Recovery of the U.S. and Canadian Iron Ore Industries following Their Early 1980s Crisis." *Journal of Political Economy*, 113(3), 582-625.

Silverman, E. and Skinner, J., (2001) "Are For-Profit Hospitals Really Different? Medicare 'Upcoding' and Market Structure" *NBER Working paper No. W8133*.

Sloan, F (2000) "Not-for-profit ownership and hospital behaviour" in Culyer, A J and Newhouse, J (eds) *Handbook of Health Economics* pp 1141-1174. Elsevier: Amsterdam.

Smith, P. (2002) "Performance management in British health care: will it deliver?" *Health Affairs* 21(3), 103-115.

Stromberg, D. (2008), "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florid", *American Economic Review*, 98(3), 769-807.

Syverson, C. (2004). "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy*, 112(6): 1181-1222.

Syverson, C. (2010), "What determines productivity", Chicago Mimeo.

Table 1: Means and Standard Deviations of Variables

Variable	Mean	Standard Deviation	Observati
<i>Management Quality (non-z-scored)</i>			
Average Management Score	2.57	0.66	161
Average Operations Score	2.83	0.95	161
Average Monitoring Score	3.00	0.75	161
Average Targets Score	2.47	0.78	161
Average Incentives Management Score	2.35	0.70	161
<i>Performance Measures</i>			
Mortality rate from emergency AMI after 28 days (quarterly average, %)	17.08	7.56	156
Mortality rate from emergency surgery after 30 days (quarterly average, %)	2.21	0.84	160
Numbers on waiting list	4,893	2,667	160
Infection rate of MRSA per 10,000 bed days (half yearly, %)	1.61	0.64	160
Operating margin (%)	1.27	2.81	161
Likelihood of leaving in next 12 months (1=very unlikely, 5=very likely)	2.70	0.13	160
Average Health Care Commission rating (1-4 scale)	2.25	0.68	161
Pseudo HCC rating (standardized)	0.00	0.98	161
<i>Competition Measures</i>			
Number of competing hospitals (in 30km radius)	6.89	9.68	161
Herfindahl index (between 0 and 10000, 30km radius)	3,803	3,097	161
<i>Political Variables</i>			
Proportion of marginal Labour constituencies (in 30km radius, %)	4.82	10.91	161
Proportion of marginal non- Labour constituencies (in 30km radius, %)	14.02	21.54	
Proportion of (all) marginal constituencies (in 30km radius. %)	18.84	23.23	161
Labour share of votes (average of constituencies in 30km radius, %)	42.29	13.98	161
<i>Noise Controls</i>			
Respondent a physician (i.e. not a manager,%)	21.1	40.9	161
Respondent's tenure in the post (years)	3.50	3.79	161
<i>Covariates</i>			
Foundation Trust (hospitals with greater autonomy, %)	34.16	47.57	161
Teaching hospital (%)	11.80	32.36	161
Specialist hospital (%)	1.86	13.56	161
Managers with a clinical degree (%)	50.38	31.7	120
Building age	25.98	8.37	152
Expenditure per patient (£ 1000)	9.69	4.51	152
Area mortality (average in 30km radius, per 100,000 population)	930	138	161
<i>Size Variables</i>			
Number of total admissions (quarterly)	18,137	9,525	161
Number of emergency AMI admissions (quarterly)	90.18	5,026	161
Number of emergency surgery admissions (quarterly)	1,498	800	161
Number of sites	2.65	2.01	161

Table 2: Hospital Performance and management practices

Dependent Variable:	(1) Mortality rate from emergency AMI	(2) Mortality rate from all emergency surgery	(3) Total waiting list	(4) MRSA infection rate	(5) Operating Margin	(6) Average intention of staff to leave in next 12 months	(7) Health Care Commission (HCC) overall rating	(8) “Pseudo” HCC rating
A. Overall Management Practices Score	-0.044* (0.026)	-0.015** (0.007)	-0.070** (0.034)	-0.175* (0.096)	0.168 (0.108)	-0.209* (0.119)	0.167** (0.070)	0.268*** (0.098)
B. Operations	-0.002 (0.023)	-0.008* (0.004)	-0.013 (0.030)	-0.158** (0.078)	0.189 (0.128)	-0.056 (0.098)	0.051 (0.062)	0.180* (0.093)
C. Monitoring	-0.025 (0.026)	-0.011** (0.006)	-0.020 (0.036)	-0.050 (0.093)	-0.023 (0.112)	-0.042 (0.087)	0.090 (0.066)	0.127 (0.090)
D. Targets	-0.048** (0.024)	-0.013** (0.007)	-0.079** (0.033)	-0.155* (0.089)	0.136 (0.105)	-0.244** (0.117)	0.174*** (0.065)	0.165* (0.090)
E. Incentives	-0.054** (0.026)	-0.010 (0.007)	-0.076** (0.034)	-0.112 (0.096)	0.143 (0.097)	-0.239** (0.118)	0.257*** (0.067)	0.273*** (0.095)
Observations	140	157	160	160	161	160	161	161

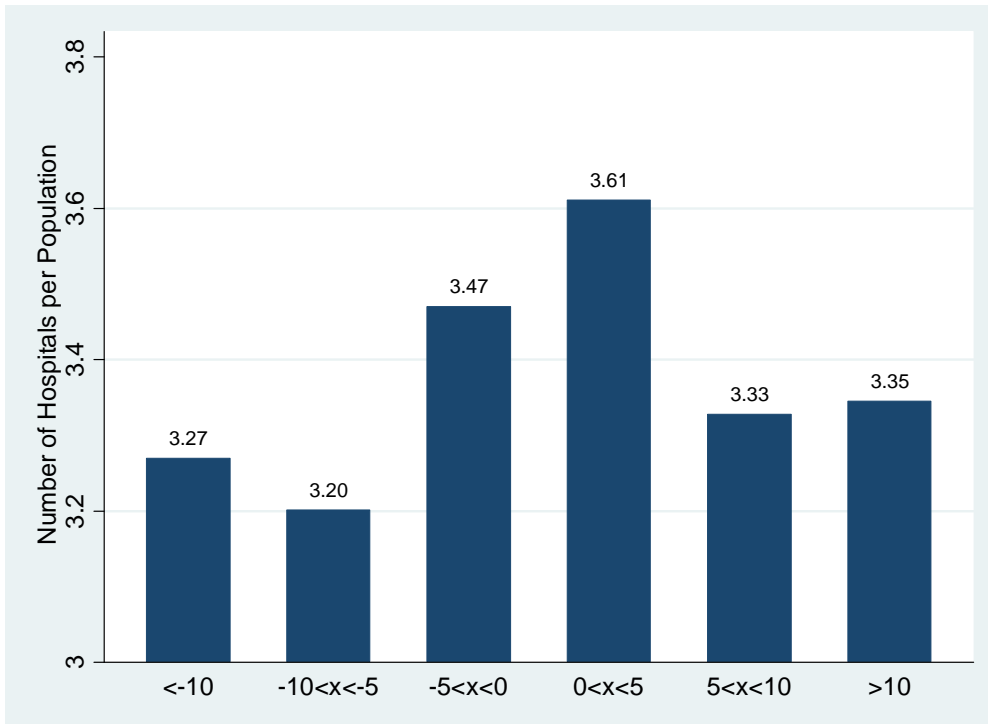
Notes: Every row constitutes a separate regression. All dependent variables are standardized to be mean zero and standard deviation 1. The dependent variables in columns (1) through (4) and (6) are generally considered to be “bad” whereas those in (5), (7) and (8) are “good” – see text for more details. Management scores are also standardized across the questions in Appendix A. These are OLS regressions with standard errors that are clustered at a hospital level (the unit of observation is a management interview with a service line in cardiology or orthopaedics across 100 public acute hospitals). *** indicates significance at the 1% level; ** significance at 5%, * significance at 10%. All columns include controls for whether the respondent was a manager or clinician, 9 regional dummies and the number of sites. Controls for case mix and total admissions are also included, but vary across columns (see text for discussion). All columns also include “noise controls” comprising interviewer dummies and tenure of the interviewee, foundation trust status, share of managers with clinical degree and joint decision making dummy. The observations are weighted by the inverse of the number of interviews with the same hospital. In Column (1) we drop hospitals with less than 150 AMI cases per year. Column (8) is average of HCC’s rating on resource use and quality of service. Column (9) is our self-constructed HCC rating based on several indicators.

Table 3: The effect of Competition on Management Practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Type of Regression	OLS	IV: First Stage	IV: Second Stage	OLS	IV: First Stage	IV: Second Stage	IV: First Stage	IV: Second Stage	OLS	IV: Second Stage
Dependent variable	Management	Number of Competing Hospitals	Management	Management	Number of Competing Hospitals	Management	Number of Competing Hospitals	Management	Mortality rate from emergency AMI	Mortality rate from emergency AMI
Number of Competing Public Hospitals	0.121** (0.058)		0.361* (0.215)	0.120* (0.068)		0.543** (0.220)		0.475** (0.202)	-0.073** (0.023)	-0.122* (0.069)
Proportion of Labor Marginal Constituencies		5.850*** (1.553)			5.156*** (1.398)		5.296*** (1.416)			
Proportion of Non-Labor Marginal Constituencies							1.245 (1.015)			
F-statistic of excluded instrument in corresponding first stage		14.18			13.60		7.36			8.94
General Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	No	No
AMI-specific controls	No	No	No	No	No	No	No	No	Yes	Yes
Observations	161	161	161	161	161	161	161	161	161	140

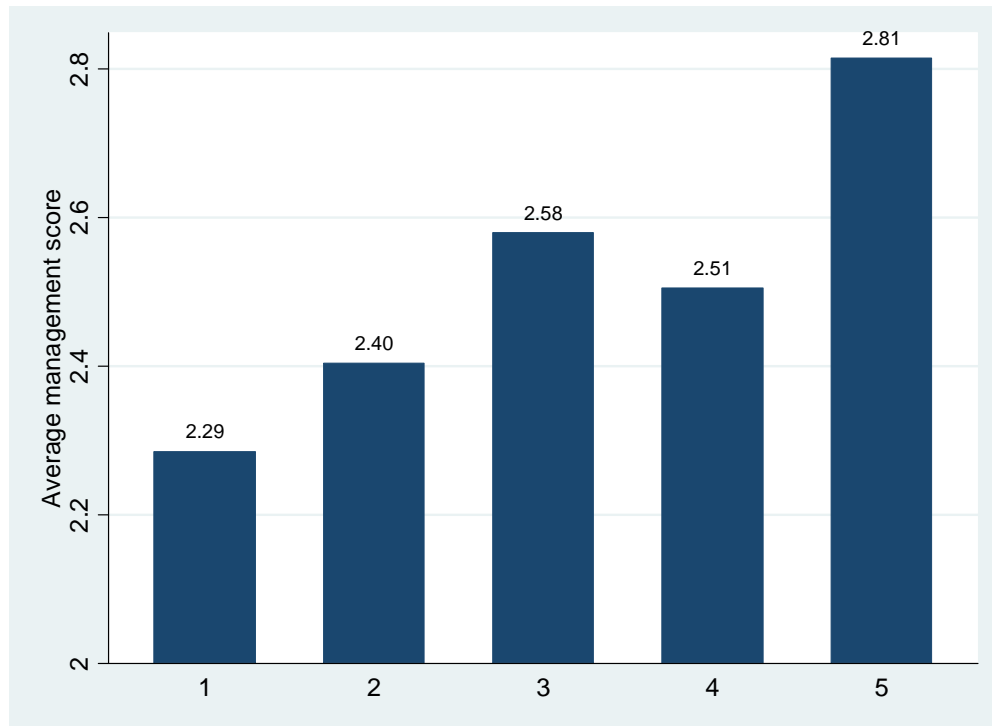
Notes: Competition is measured as the number of hospitals in a 30km radius around the hospital (the “catchment area”) unless otherwise stated. A political constituency is defined as marginal if it was won by less than 5% in the 1997 General Election (proportion of marginal constituencies is based on the catchment area). The Labour share of votes is the absolute share obtained by the Governing party in the 1997 UK General Election averaged over all constituencies in the catchment area. Standard errors are clustered at a hospital level (the unit of observations is a service line in cardiology or orthopaedics). All columns include controls for the total population and age profile (11 categories) in the catchment area, whether the hospital was a Foundation Trust and interviewer dummies (4). “General controls” include share of Labour votes, the number of political constituencies and average mortality in the catchment area; the tenure of the respondent and whether she was a manager or clinician, 9 regional dummies, the number of hospital sites, number of total admissions and the “case-mix” (age/gender profile of admissions), share of managers with a clinical degree and a dummy for whether there was joint decision making at the hospital level. “AMI specific controls” are those in Table 2 column (1). Regressions are weighted by the inverse of the number of interviews within the same hospital.

Figure 1: Governing Party's (Labour) Winning Margin and the number of hospitals in a political constituency



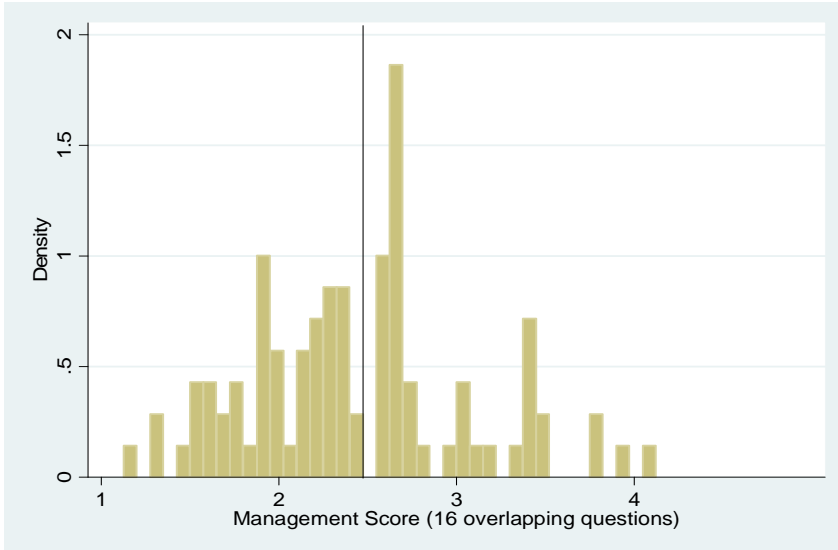
Notes: This figure plots the mean number of hospitals per 1 million people within a 30km radius of a constituency (we measure the distance of the hospital to the constituency centroid) against the “winning margin” of the governing party (Labour). When Labour is not the winning party, the margin is the negative of the difference between the winning party (usually Conservative) and the next closest party. There are 528 political constituencies in England.

Figure 2: Management Score by quintiles of average HCC rating

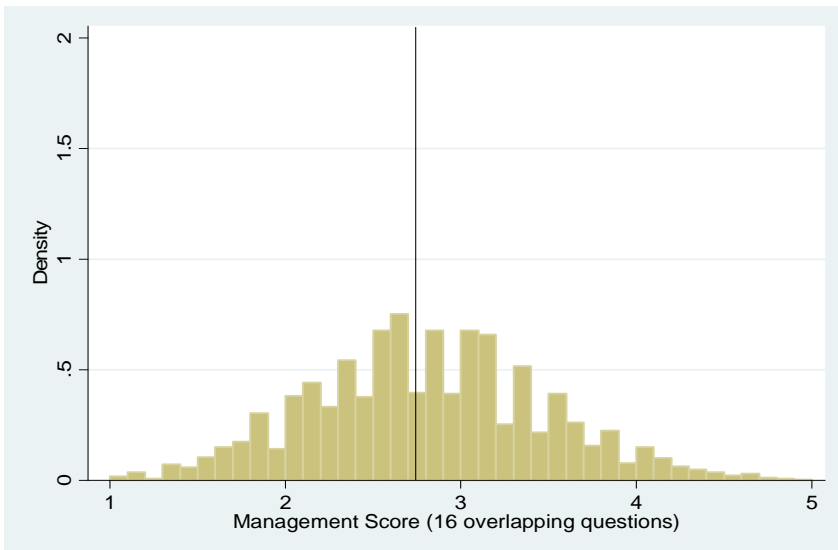


Notes: We divide the HCC average score into quintiles from lowest score (first) to highest score (fifth). We show the average management score (over all 18 questions) in each of the quintiles. The better performing hospitals have higher management scores.

Figure 3: Comparison of Management Scores in Hospitals and Manufacturing Firms



Panel A
Management Scores
in Hospitals



Panel B:
Management Scores
In manufacturing firms

Notes: These are the distributions of the management score for hospitals and manufacturing firms. Only establishments who have between 50 and 5000 employees and who are domestically owned (i.e. multinationals were dropped from the manufacturing sample) were used here. Also observations with a low reliability score (below 3) were dropped. The vertical line represents the average management score in each sample. Only the 16 questions for which manufacturing and healthcare are comparable (questions 3 to 18) were used.

APPENDIX A: MANAGEMENT PRACTICE INTERVIEW GUIDE FOR THE HEALTHCARE SECTOR

Any score from 1 to 5 can be given, but the scoring guide and examples are only provided for scores of 1, 3 and 5. Multiple questions are used for each dimension to improve scoring accuracy.

(1) Lay out of patient flow

Tests how well the patient pathway is configured at the infrastructure level and whether staff pro-actively improve their own work-place organisation

- a) Can you briefly describe the patient journey or flow for a typical episode?
- b) How closely located are wards, theatres, diagnostics centres and consumables?
- c) Has the patient flow and the layout of the hospital changed in recent years? How frequently do these changes occur and what are they driven by?

	Score 1	Score 3	Score 5
Scoring grid:	Lay out of hospital and organisation of workplace is not conducive to patient flow, e.g., ward is on different level from theatre, or consumables are often not available in the right place at the right time	Lay out of hospital has been thought-through and optimised as far as possible; work place organisation is not regularly challenged/changed (or vice versa)	Hospital layout has been configured to optimize patient flow; workplace organization is challenged regularly and changed whenever needed

(2) Rationale for introducing standardisation/ pathway management

Test the motivation and impetus behind changes to operations and what change story was communicated

- a) Can you take me through the rationale for making operational improvements to the management of patient pathway? Can you describe a recent example?
- b) What factors led to the adoption of these practices?
- c) Who typically drives these changes?

	Score 1	Score 3	Score 5
Scoring grid:	Changes were imposed top down or because other departments were making (similar) changes, rationale was not communicated or understood	Changes were made because of financial pressure and the need to save money or as a (short-term) measure to achieve government targets	Changes were made to improve overall performance, both clinical and financial, with buy-in from all affected staff groups. The changes were communicated in a coherent 'change story'

(3) Continuous improvement

Tests process for and attitudes to continuous improvement and whether things learned are captured/documentated

- a) How do problems typically get exposed and fixed?
- b) Talk me through the process for a recent problem that you faced
- c) How do the different staff groups get involved in this process? Can you give examples?

	Score 1	Score 3	Score 5
Scoring grid:	No, process improvements are made when problems occur, or only involve one staff group	Improvements are made irregular meetings involving all staff groups, to improve performance in their area of work (e.g., ward or theatre)	Exposing problems in a structured way is integral to individuals' responsibilities and resolution involves all staff groups, along the entire patient pathway as a part of regular business processes rather than by extraordinary effort/teams

(4) Performance tracking

Tests whether performance is tracked using meaningful metrics and with appropriate regularity

- a) What kind of performance indicators would you use for performance tracking?
- b) How frequently are these measured? Who gets to see these data?
- c) If I were to walk through your hospital wards and theatres, could I tell how you were doing against your performance goals?

	Score 1	Score 3	Score 5
Scoring grid:	Measures tracked do not indicate directly if overall objectives are being met, e.g., only government targets tracked. Tracking is an ad-hoc process (certain processes aren't tracked at all).	Most important performance indicators are tracked formally; tracking is overseen by senior staff.	Performance is continuously tracked and communicated against most critical measures, both formally and informally, to all staff using a range of visual management tools

(5) Performance review

Tests whether performance is reviewed with appropriate frequency and communicated with staff

- a) How do you review your KPI's?
- b) Tell me about a recent meeting
- c) Who is involved in these meetings? Who gets to see the results of this review?
- d) What is the follow-up plan?

	Score 1	Score 3	Score 5
Scoring grid:	Performance is reviewed infrequently or in an un-meaningful way e.g. only success or failure is noted	Performance is reviewed periodically with both successes and failures identified. Results are communicated to senior staff. No clear follow up plan is adopted.	Performance is continually reviewed, based on the indicators tracked. All aspects are followed up to ensure continuous improvement. Results are communicated to all staff.

(6) Performance dialogue

Tests the **quality** of review conversations

- a) How are these meetings structured?
- b) During these meetings do you find that you generally have enough data?
- c) What type of feedback occurs in these meetings?

	Score 1	Score 3	Score 5
Scoring grid:	The right information for a constructive discussion is often not present or the quality is too low; conversations focus overly on data that is not meaningful. Clear agenda is not known and purpose is not explicitly. Next steps are not clearly defined	Review conversations are held with the appropriate data present. Objectives of meetings are clear to all participating and a clear agenda is present. Conversations do not, drive to the root causes of the problems, next steps are not well defined	Regular review/performance conversations focus on problem solving and addressing root causes. Purpose, agenda and follow-up steps are clear to all. Meetings are an opportunity for constructive feedback and coaching

(7) Consequence management

Tests whether differing levels of (personal) performance lead to different consequences (good or bad)

- a) Let's say you've agreed to a follow up plan at one of your meetings, what would happen if the plan weren't enacted?
- b) How long is it between when a problem is identified to when it is solved? Can you give me a recent example?
- c) How do you deal with repeated failures in a specific sub-specialty or cost area?

	Score 1	Score 3	Score 5
Scoring grid:	Failure to achieve agreed objectives does not carry any consequences	Failure to achieve agreed results is tolerated for a period before action is taken	A failure to achieve agreed targets drives retraining in identified areas of weakness or moving individuals to where their skills are appropriate

(8) Target balance

Test whether targets cover a sufficiently broad set of metrics

- a) What types of targets are set for the hospital? What are the goals for your specialty?
- b) Tell me about goals that are not set externally (e.g. by the government, regulators).

	Score 1	Score 3	Score 5
Scoring grid:	Goals focussed only on government targets and achieving the budget	Goals are a balanced set of targets (including quality, waiting times, operational efficiency, and financial balance). Goals form part of the appraisal for senior staff only or do not extend to all staff groups. Real interdependency is not well understood	Goals are a balanced set of targets covering all four dimensions (see left). Interplay of all four dimensions is understood by senior and junior staff (clinicians as well as nurses and managers)

(9) Target inter-connection

Tests whether targets are tied to hospital/Trust objectives and how well they cascade down the organisation

- a) What is the motivation behind your goals?
- b) How are these goals cascaded down to the different staff groups or to individual staff members?
- c) How are your targets linked to hospital performance and its goals?

	Score 1	Score 3	Score 5
Scoring grid:	Goals do not cascade down the organisation	Goals do cascade, but only to some staff groups, e.g., nurses only	Goals increase in specificity as they cascade, ultimately defining individual expectations, for all staff groups

(10) Time horizon of targets

Tests whether hospital/Trust has a '3 horizons' approach to planning and targets

- a) What kind of time scale are you looking at with your targets?
- b) Which goals receive the most emphasis?
- c) Are the long term and short term goals set independently?
- d) Could you meet all your short-run goals but miss your long-run goals?

	Score 1	Score 3	Score 5
Scoring grid:	Top staff's main focus is on short term targets	There are short and long term goals for all levels of the organisation. As they are set independently, they are not necessarily linked to each other	Long term goals are translated into specific short term targets so that short term targets become a 'staircase' to reach long term goals

(11) Target stretch

Tests whether targets are appropriately difficult to achieve

- a) How tough are your targets? Do you feel pushed by them?
- b) On average, how often would you say that you meet your targets?
- c) Do you feel that on targets all specialties, departments or staff groups receive the same degree of difficulty? Do some groups get easy targets?
- d) How are the targets set? Who is involved?

	Score 1	Score 3	Score 5
Scoring grid:	Goals are either too easy or impossible to achieve, at least in part because they are set with little clinician involvement, e.g., simply off historical performance	In most areas, senior staff push for aggressive goals based, e.g., on external benchmarks, but with little buy-in from clinical staff. There are a few sacred cows that are not held to the same standard	Goals are genuinely demanding for all parts of the organisation and developed in consultation with senior staff, e.g., to adjust external benchmarks appropriately

(12) Clarity and comparability of targets

Tests how easily understandable performance measures are and whether performance is openly communicated

- a) If I asked your staff directly about individual targets, what would they tell me?
- b) Does anyone complain that the targets are too complex?
- c) How do people know about their own performance compared to other people's performance?

	Score 1	Score 3	Score 5
Scoring grid:	Performance measures are complex and not clearly understood, or only relate to government targets. Individual performance is not made public	Performance measures are well defined and communicated; performance is public at all levels but comparisons are discouraged	Performance measures are well defined, strongly communicated and reinforced at all reviews; performance and rankings are made public to induce competition

(13) Managing talent

Tests what emphasis is put on talent management

- a) How do senior staff show that attracting and developing talent is a top priority?
- b) Do senior managers, clinicians or nurses get any rewards for bringing in and keeping talented people in the hospital?

	Score 1	Score 3	Score 5
Scoring grid:	Senior staff do not communicate that attracting, retaining and developing talent throughout the organisation is a top priority	Senior management believe and communicate that having top talent throughout the organisation is key to good performance	Senior staff are evaluated and held accountable on the strength of the talent pool they actively build

(14) Rewarding high performers

Tests whether good performance is rewarded proportionately

- a) How does your appraisal system work? Tell me about your most recent round.
- b) Are there any non-financial or financial (bonuses) rewards for the best performers across all staff groups?
- c) How does the bonus system work?
- d) How does your reward system compare to that at other comparable hospitals?

	Score 1	Score 3	Score 5
Scoring grid:	People are rewarded equally irrespective of performance level	There is an evaluation system for the awarding of performance related rewards that are non-financial (beyond progression through nursing grades or clinical excellence awards for doctors) at the individual level (but rewards are always or never achieved)	There is an evaluation system for the awarding of performance related rewards, including personal financial rewards

(15) Removing poor performers

Tests whether hospital is able to deal with underperformers

- a) If you had a clinician or a nurse who could not do his job, what would you do? Could you give me a recent example?
- b) How long would underperformance be tolerated?
- c) Do you find staff members who lead a sort of charmed life? Do some individuals always just manage to avoid being fixed/fired?

	Score 1	Score 3	Score 5
Scoring grid:	Poor performers are rarely removed from their positions	Suspected poor performers stay in a position for a few years before action is taken	We move poor performers out of the hospital/department or to less critical roles as soon as a weakness is identified

(16) Promoting high performers

Tests whether promotion is performance based

- a) Tell me about your promotion system?
- b) What about poor performers? What happens with them? Are there any examples you can think of?
- c) How would you identify and develop your star performers?
- d) Are better performers likely to promote faster or are promotions given on the basis of tenure/seniority?

	Score 1	Score 3	Score 5
Scoring grid:	People are promoted primarily on the basis of tenure	People are promoted upon the basis of performance (across more than one dimension, e.g., isn't related only to research or clinical excellence)	We actively identify, develop and promote our top performers

(17) Attracting talent

Tests how strong the employee value proposition is

- a) What makes it distinctive to work at your hospital, as opposed to your other similar hospitals?
- b) If I were a top nurse or clinician and you wanted to persuade me to work at your hospital, how would you do this?
- c) What don't people like about working at your hospital?

	Score 1	Score 3	Score 5
Scoring grid:	Our competitors offer stronger reasons for talented people to join their hospitals	Our value proposition to those joining our department is comparable to those offered by others hospitals	We provide a unique value proposition to encourage talented people join our department above our competitors

(18) Retaining talent

Tests whether hospital/Trust will go out of its way to keep its top talent

- a) If you had a top performing manager, nurse or clinician that wanted to leave, what would the hospital do?
- b) Could you give me an example of a star performer being persuaded to stay after wanting to leave?
- c) Could you give me an example of a star performer who left the hospital without anyone trying to keep them?

	Score 1	Score 3	Score 5
Scoring grid:	We do little to try and keep our top talent	We usually work hard to keep our top talent	We do whatever it takes to retain our top talent across all three staff groups

APPENDIX B: DATA

Sample

The main sampling frame was all acute public sector hospitals (NHS “trusts”) in England.³⁸ There were 174 such units in 2006, but we dropped hospitals without orthopaedics or cardiology departments (e.g. specialist eye hospitals) so this left us with a sample of 164 possible hospital trusts. We obtained 161 usable responses from 100 hospital trusts which represented 61% of the frame. We sought responses from up to four senior employees in each hospital: a manager and a clinician from two service lines (cardiology and orthopaedics). Table 1 shows that we are split evenly between the specialities (52% cardiology and 48% orthopaedics), but that it was harder to obtain interviews with the physicians than managers (80% of the respondents were managers). We interviewed one respondent in 53 hospitals, two respondents in 34 hospitals, three respondents in 12 hospitals and four respondents in one hospital. The correlation of the average management score across responders within the same hospital was high (0.54).

We examined evidence for selection bias by estimating probit models of whether a trust responded on the observable characteristics. These characteristics are drawn from Department of Health datasets (such as Hospital Episode Statistics). Table B2 contains the results of this exercise. There is no significant correlation at the 5% level between sample response and any of the performance measures or covariates and only one (from 16) of the indicators are significant at the 10% level. This suggests that there was little systematic response bias.

In the regressions all interviews with a short duration (less than 25 minutes) or many unanswered questions (three or more) are excluded as the information obtained is unlikely to be reliable. We weight regressions by the inverse of the number of interviews so that hospitals with multiple responses are weighted less (we also cluster standard errors at the hospital level).

Construction of the Pseudo HCC Rating

In column (8) of Table 2 we reported our best effort to reconstruct the HCC’s rating. Although the exact method of creating the HCC ratings is not publicly known the Appendix of the HCC’s “Annual Health Check 2006/2007” brochure mentions seven “domains” in which the hospitals need to achieve certain standards in order to achieve a high score.

These domains are: safety, clinical and cost effectiveness, governance, patient focus, accessible and responsive care, public health, and care environment and amenities. From the datasets described above we choose eight variables which capture the requirements of these different domains. Infection rates and re-admission risk are chosen to represent the “safety” aspect; operational margin and income per medical FTE capture the financial side; patient satisfaction covers the “patient focus” domain. Waiting times and average length of stay fall into the category “accessible and responsive care” and information on job satisfaction from the NHS staff survey is used to represent the “care environment and amenities” domain.

³⁸ A trust can consist of more than one site (as a firm can consist of more than one plant). The median number of sites was 2 with a range from 1 to 10.

Table B1: Data Sources for hospital performance data

Variable	Notes	Source
Mortality within 28 days of emergency admission for AMI (in hospital and out of hospital)	<ul style="list-style-type: none"> • During financial quarter • Defined according to NHS mortality rate Performance indicators (PIs) for 2001/02 	ONS death records linked with Hospital Episode Statistics (HES), The NHS Information Centre for health and social care.
Mortality within 30 days of surgery for selected emergency procedures.	<ul style="list-style-type: none"> • During financial quarter • Defined according to NHS mortality rate PIs for 2001/02 	ONS death records linked with Hospital Episode Statistics (HES), The NHS Information Centre for health and social care.
Waiting list size	<ul style="list-style-type: none"> • At start of quarter (as proxied by end of previous quarter) 	Department of Health: Provider based waiting times/list statistics ^a
MRSA rates	<ul style="list-style-type: none"> • During 6-month period • 2001/02 (q1) to 2004/05 (q3) 	Health Protection Agency: Half-yearly reporting results for clostridium difficile infections and MRSA bacteraemia
Financial Indicators: <ul style="list-style-type: none"> • Income per spell • Income per bed • Operating Margin 		Trust Financial Returns
Probability of leaving in next 12 months	Respondents are asked to rate chances of leaving on a 1 to 5 scale.	NHS Staff Survey ^c (2006). 128,328 NHS staff responded and results are reported as average of scale by each trust
Healthcare Commission rating ^d	All trusts are scored on a scale of 1 to 4 on “resource use” and quality of “care”	Our main indicator averages over the two measures and standardizes. We also construct our own “pseudo” HCC rating from the underlying indicators (see Appendix B for full description)
Local authority all cause mortality rates	<ul style="list-style-type: none"> • Calendar year 	Office of National Statistics 1995-2004
Casemix of admissions: These are specific to the conditions (AMI, surgery, etc.) considered. For the general performance indicators (like HCC rating) we use case mix for all admitted patients.	<ul style="list-style-type: none"> • Proportion of admitted patients in each sex-specific age band. 11 age categories: 0-15, 16-45, 46-50, 51-55, 56-60, 61-65, 66-70, 71-75, 76-80, 81-85, >85 and two genders, so up to 22 controls. 	Hospital Episode Statistics (HES), The NHS Information Centre for health and social care.

Notes: MRSA is Methicillin-Resistant Staphylococcus Aureus

^a http://www.performance.doh.gov.uk/nhsperformanceindicators/2002/trdca_t.doc. ^b <http://www.performance.doh.gov.uk/waitingtimes/index.htm>

^c <http://www.cqc.org.uk/usingcareservices/healthcare/nhsstaffsurveys.cfm>

^d http://www.cqc.org.uk/db/documents/0607_annual_health_check_performance_rating_scoring_rules_200702284632.pdf

Variable	Notes	Source
Number of admissions <ul style="list-style-type: none"> Total admissions AMI Emergency Surgery 	<ul style="list-style-type: none"> During financial quarter 	Hospital Episode Statistics (HES), The NHS Information Centre for health and social care.
Number of Sites		Hospital Estates and Facilities Statistics ^a , The NHS Information Centre for health and social care.
Foundation Trust Status		Monitor (Foundation Trust Regulator) ^b
Specialist Hospital	<ul style="list-style-type: none"> Self-coded from individual hospital web pages Only 2 in the sample: one specialist cardiology centre and a children hospital 	Self-coded
Building Age	Data is provided at the site level and aggregated up to hospital level using the surface area as weights	Hospital Estates and Facilities Statistics ^a , The NHS Information Centre for health and social care.
Expenditure per patient	<ul style="list-style-type: none"> Cost divided by the number of total admissions 	Cost data from Trusts' Annual Reports and Accounts, available from Trusts' webpages or Monitor ^b (in the case of Foundation Trusts)
Political Variables <ul style="list-style-type: none"> Marginal Constituencies Labour Vote Share Winning Party 	<ul style="list-style-type: none"> 4 elections from 1992 until 2005 	British Election Study

^a <http://www.hefs.ic.nhs.uk/ReportFilter.asp>

^b <http://www.monitor-nhsft.gov.uk/>

^c <http://ratings2005.healthcarecommission.org.uk/Downloads/MoreInformationPageDocs/DetailedResultsAS.xls>

^d<http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/Perfomancedataandstatistics/Beds/index.htm>

Table B2: Tests of sample Selection for public hospitals

Variable	Marginal effect (Standard error)	Observations
<i>Performance Measures</i>		
Mortality rate from emergency AMI after 28 days (quarterly average)	0.129 (0.161)	133
Mortality rate from emergency surgery after 30 days (quarterly average)	0.313 (0.365)	163
Numbers on waiting list	0.025 (0.0454)	163
Infection rate of MRSA per 10,000 bed days (half yearly)	-0.025 (0.041)	163
Operating margin (percent)	0.040 (0.032)	164
Likelihood of leaving in next 12 months (1=very unlikely, 5=very likely)	-0.063 (0.04)	161
Average Health Care Commission rating (1-4 scale)	-0.011 (0.043)	164
Pseudo HCC rating (standardized)	0.027 (0.038)	164
<i>Size Variables</i>		
Number of total admissions (per 100,000 population)	0.213 (0.417)	164
Number of emergency AMI admissions (per 100,000 population)	53.896 (70.863)	164
Number of emergency surgery admissions (per 100,000 population)	0.612 (4.739)	164
Number of sites	0.016 (0.196)	164
<i>Covariates</i>		
Foundation Trust (hospitals with greater autonomy)	0.091 (0.082)	164
Building age	-0.013 (0.013)	154
Expenditure per patient (£ 1000)	-0.015 (0.008)*	156
Area mortality (average of local authorities in 30km radius, per 100,000,000 population)	0.275 (0.277)	163

Notes: These are the results from separate probit ML regression of whether a public hospital had any response to the survey on the relevant variable (e.g. AMI mortality rate in the first row). There is a population of 164 potential acute hospitals in England and we had 100 hospitals with at least one respondent. For the first 2 rows we use the same restrictions as in table 2: we use

only hospitals with more than 150 yearly cases in the AMI regression and exclude specialist hospitals from the regression in the second row. *** indicates significance at 1% level; ** significance at 5%, * for significance at 10%.

Table B3: Full Results for Regressions

Type of Regression Dependent Variable	(1) OLS Management score	(2) 1 st Stage Number Hospitals	(3) IV Management score
Number of Competing Hospitals	0.120* (0.068)		0.543** (0.220)
Proportion of Labor Marginal Constituencies		5.156*** (1.398)	
<u>General Controls</u>			
Number of Constituencies	0.038 (0.051)	0.152** (0.066)	-0.008 (0.062)
Labor Share of Votes	-0.017 (0.019)	0.013 (0.033)	-0.025 (0.025)
Catchment Area Overall Mortality Rate	0.002 (0.004)	-0.008 (0.005)	0.007 (0.005)
Foundation Trust Dummy	0.707*** (0.169)	0.237 (0.314)	0.673*** (0.234)
Proportion of Managers with Clinical Degree	0.628 (0.384)	-0.501 (0.353)	0.813* (0.435)
Dummy Clinicians and Managers Take Decision Jointly	0.275* (0.141)	-0.214 (0.207)	0.407** (0.175)
<u>Interview (Noise) Controls</u>			
Clinician Dummy	-0.680*** (0.166)	-0.134 (0.166)	-0.563*** (0.184)

Interviewer Dummy 1	0.239 (0.654)	0.433 (0.348)	0.105 (0.641)
Interviewer Dummy 2	-0.491 (0.632)	0.404 (0.330)	-0.620 (0.613)
Interviewer Dummy 3	0.249 (0.631)	0.534 (0.347)	0.201 (0.610)
Tenure in the Post	-0.064*** (0.018)	-0.008 (0.023)	-0.061*** (0.022)
<u>Population Controls</u>			
Total Population in Catchment Area (unit: 1,000,000)	-0.529 (0.535)	1.298* (0.718)	-1.230 (0.826)
Age-/ Gender-Controls (F-stat for 11 Variables)	3.36***	4.59***	2.83***
<u>Case-Mix Controls</u>			
Total Admissions (unit: 10,000)	0.339*** (0.103)	0.201 (0.146)	0.306*** (0.112)
Age-/ Gender Controls (F-stat for 21 Variables)	4.00***	2.10***	2.81***
Observations	161	161	161

Notes: Competition is measured as the number of hospitals in a 30km radius around the hospital (the “catchment area”). A Political Constituency is defined as marginal if it was won by less than 5% in the 1997 UK General Election and the proportion of marginal constituencies is based on the 30km catchment area. The Labour share of votes is the absolute share obtained by the Labour party in the 1997 UK General Election averaged over all constituencies in the catchment area. Standard errors are clustered at a hospital level (the unit of observations is a service line in cardiology or orthopaedics). All variables in the regressions are reported in the table. The observations are weighted by the inverse of the number of interviews within the same hospital.