

HEALTHY BUSINESS? MANAGERIAL EDUCATION AND MANAGEMENT IN HEALTH CARE

Nicholas Bloom, Renata Lemos, Raffaella Sadun, and John Van Reenen*

Abstract—We investigate the link between hospital performance and managerial education by collecting a large database of management practices and skills in hospitals across nine countries. We find that hospitals closer to universities offering both medical education and business education have lower mortality rates from acute myocardial infarction (heart attacks), better management practices, and more MBA-trained managers. This is true compared to the distance to universities that offer only business or medical education (or neither). We argue that supplying bundled medical and business education may be a channel through which universities improve management practices in hospitals and raise clinical performance.

I. Introduction

ACROSS the world, health care systems are under severe stress due to shocks like the COVID-19 pandemic, aging populations, the rising costs of medical technologies, tight public budgets, and increasing expectations. Given the evidence of enormous variations in efficiency levels across different hospitals and health care systems, these pressures could be mitigated by improving hospital productivity. For example, high-spending areas in the United States incur costs that are 50% higher than low-spending ones (Fisher et al., 2003, in the Dartmouth Atlas).¹ Some commentators focus on technologies (such as information and communication technologies) as a key reason for such differences, but others have focused on divergent preferences and human capital among medical professionals (Phelps & Mooney, 1993; Eisenberg, 2002; Sirovich et al., 2008). One aspect of the latter are management practices such as checklists (Gawande, 2009).

In this paper we measure management practices across hospitals in the United States and eight other countries using a survey tool originally applied by Bloom and Van Reenen (2007) for the manufacturing sector. The underlying concepts of the survey tool are very general and provide a metric to measure the adoption of best practices over operations, monitoring, targets, and people management in hospitals. We document considerable variation in management practices both

Received for publication September 27, 2017. Revision accepted for publication February 4, 2019. Editor: Shachar Kariv.

*Bloom: Stanford University, SIEPR, NBER; Lemos: World Bank, CEP-LSE; Sadun: Harvard University; Van Reenen: MIT, CEP-LSE.

We thank the European Research Council and Economic and Social Research Centre for financial support through the Centre for Economic Performance. We are grateful to Daniela Scur for ongoing discussion and feedback on the paper. Dennis Layton, Stephen Dorgan, and John Dowdy were invaluable partners in this project although we have received no financial support from McKinsey (or any other company).

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/rest_a_00847.

¹Annual Medicare spending per capita ranges from \$6,264 to \$15,571 across geographic areas (Skinner, 2011), yet health outcomes do not positively co-vary with these spending differentials (Baicker & Chandra, 2004; Chandra, Staiger, & Skinner, 2010). Finkelstein, Gentzkow, and Williams (2016) estimate that at least half of these effects arise from place-based supply factors rather than unobserved patient-specific health and demand factors.

between and within countries. Hospitals with high management scores have high levels of clinical performance, as proxied by outcomes such as survival rates from emergency heart attacks (acute myocardial infarction, AMI). These hospitals also tend to have a higher proportion of managers with greater levels of business skills as measured by whether they have attained MBA-type degrees.

To further investigate the importance of the supply of human capital on managerial and clinical outcomes, we draw on data from the World Higher Education Database (WHED), which provides the location of all universities in our chosen countries (see Valero & Van Reenen, 2019). We calculate geographical closeness measures (the driving time from a hospital to the nearest university) by geocoding the location of all hospitals and universities in our sample. We show that hospitals that are closer to universities offering both medical and business courses within their premises have significantly better clinical outcomes and management practices than those located farther away. This relationship holds even after conditioning on a wide range of location-specific characteristics such as average income, population density, and climate. By contrast, the distance to universities with only a business school, only a medical school, or neither (as in a pure liberal arts college offering only arts, humanities, or religious courses) has no significant relationship with hospital management quality, suggesting that the results are not entirely driven by unobserved heterogeneity in location characteristics correlated with educational institutions.

Proximity to schools offering bundles of medical and managerial courses is positively associated with the fraction of managers with formal business education (MBA-type courses) in hospitals, consistent with the idea that the courses increase the supply of employees with these combined skills. We do not have an instrumental variable for the location of universities and therefore cannot demonstrate that the correlations are causal. Nevertheless, these results are suggestive of a strong, and so far unexplored, relationship between managerial education and hospital performance.

Our paper relates to several literatures. First, the paper is related to the literature documenting the presence of wide productivity differences across hospitals. Chandra et al. (2016) estimate a large heterogeneity in hospital total factor productivity across U.S. hospitals of an order of magnitude similar to that documented in manufacturing and retail. We contribute to this literature by suggesting that management and, indirectly, management education may be a possible factor driving the productivity dispersion via its effect on management practices. Second, our paper contributes to the literature on the importance of human capital (especially managerial human capital) for organizational performance. Examples of

this work include Bertrand and Schoar (2003) for CEOs, Moretti (2004) for ordinary workers, and Gennaioli et al. (2013) at the regional and national levels. More specifically Doyle, Ewer, and Wagner (2010) examine the causal importance of physician human capital on patient outcomes, while Goodall (2011) looks at the importance of physician leadership in hospitals. Finally, this paper is related to the work on measuring management practices across firms, sectors, and countries—for example, Osterman (1994), Huselid (1995), Ichniowski, Shaw, and Prennushi (1997), Black and Lynch (2001), and Bloom et al. (2014).

The structure of the paper is as follows. In section II, we provide an overview of the methodology used to collect the hospital management data, the health outcomes data, the skills data, and other data used in the analysis. Section III describes the basic summary statistics emerging from the data, section IV presents the results, and section V concludes. The online appendixes give much more detail on the data (A), additional results (B), sampling frame (C), and case studies of management practices in individual hospitals (D).

II. Data

A. *Collecting Measures of Management Practices across Countries*

To measure hospital management practices, we adapt the World Management Survey (Bloom & Van Reenen, 2007; Bloom et al., 2014) methodology to health care. This is based on the work of international consultants and the health care management literature. The evaluation tool scores a set of twenty basic management practices on a grid from 1 (“worst practice”) to 5 (“best practice”) in four broad areas: operations (four questions), monitoring (five questions), targets (five questions), and human resource management (six questions). The full list of dimensions is in appendix table A1.

Hospitals with very weak management practices (score of 2 or below) have almost no monitoring, very weak targets (e.g., only an annual hospital-level target), and extremely weak incentives (e.g., tenure-based promotion, no financial or nonfinancial incentives, and no effective action taken over underperforming medical staff). In contrast, hospitals with a score of 3 or above have some reasonable and proactive performance monitoring, processes in place for continuous improvement, a mix of targets covering a broad set of metrics and timescale, performance-based promotion, and systematic ways to address and correct persistent underperformance. To compute the main management practices score used in our regression analysis, we standardize the index to have zero mean and standard deviation of 1 by z-scoring the average of the z-scores of the twenty individual management questions.

The data were collected for Canada, France, Italy, Germany, Sweden, the United States, and the United Kingdom in 2009; India in 2012; and Brazil in 2013. For the United King-

dom, we combine two waves of the survey: 2006 and 2009.² The choice of countries was driven by funding availability, the availability of hospital sampling frames, and research and policy interest.

In every country, the sampling frame for the management survey was randomly drawn from administrative register data and included all hospitals that (a) have an orthopedics or cardiology department, (b) provide acute care, and (c) have overnight beds. Interviewers were each given a random list of hospitals from a sampling frame representative of the population of hospitals with these characteristics in the country.³ Within each department, we targeted the director of nursing, medical superintendent, nurse manager or administrator of the specialty—that is, the clinical service lead at the top of the specialty who was still involved in its management on a daily basis.

We used a variety of procedures to persuade hospital employees to participate in the survey. First, we encouraged our interviewers to be persistent: they ran on average two interviews a day that lasted for an average of an hour each. Second, we never asked hospital managers about the hospital’s overall performance during the interview (these were obtained from external administrative sources). Third, we sent informational letters and, if necessary, copies of country endorsements letters (e.g., U.K. Health Department). Following these procedures helped us obtain a reasonably high response rate of 34%, similar to the response rates for our manufacturing and school surveys. The country-specific response rates ranged from 66%, 53%, and 49% of eligible hospitals in, respectively, Sweden, Germany, and Brazil, down to 21% of eligible hospitals in the United States.⁴ In terms of selection bias, we compare our sample of hospitals for which we secured an interview with the sample of all eligible hospitals in our sampling frame for each country on dimensions such as size, ownership, and geographical location. Looking at the overall pattern of results, we obtain few significant coefficients with marginal effects small in magnitude.⁵ We further construct sampling weights and observe that our main unweighted results hold even when using this alternative sample weighting scheme. We describe our selection analysis as

²The 2006 U.K. data have been used in Bloom et al. (2015).

³During the survey, if the hospital did not have an orthopedics department or if the manager in this department was not available, we then tried to get in touch with the cardiology department. In our sample, there are 937 observations for multispecialty departments, 460 observations for the orthopedics department, 262 for cardiology, 138 for surgery when orthopedics- or cardiology-related procedures were carried out in the surgery department, and 163 for other departments that still carried out orthopedics- or cardiology-related procedures when the departments mentioned above did not exist in the hospital (the rest is surgery/other).

⁴This was mainly due to not completing all the interviews (due to rescheduling) rather than outright refusals. The explicit refusal rate was 11%, ranging from no refusals in hospitals in Sweden to 22% of all eligible hospitals in Germany.

⁵For example, there were higher response rates in India for certain locational characteristics (population density, education, and located farther away from the coast), in the United States for public hospitals, and in Germany and Italy for hospital size.

well as the sampling frame sources and response rates in more detail in appendix C.

To elicit candid responses, we took several steps. First, our interviewers received extensive training in advance on hospital management. Second, we employed a double-blind technique: interviewers were not told in advance about the hospital's performance—they only had the hospital's name and telephone number—and respondents were not told in advance that their answers were scored. Third, we told respondents we were interviewing them about their hospital management, asking open-ended questions like, "Tell me how you track performance?" and "If I walked through your ward, what performance data might I see?" The combined responses to these types of questions are scored against a grid. For example, the following two questions helped to score question 6, on performance monitoring, which went from 1, defined as, "Measures tracked do not indicate directly if overall objectives are being met. Tracking is an ad hoc process (certain processes aren't tracked at all)," to 5, defined as, "Performance is continuously tracked and communicated, both formally and informally, to all staff using a range of visual management tools." Interviewers kept asking questions until they could score each dimension. Three other steps were taken to guarantee data quality. First, each interviewer conducted on average of 39 interviews in order to generate consistent interpretation of responses. They received one week of intensive initial training and four hours of weekly ongoing training.⁶ Second, 70% of interviews had another interviewer silently listening and scoring the responses, which the second interviewer discussed with the lead interviewer after the end of the interview. This provided cross-training, consistency, and quality assurance. Third, we collected a series of "noise controls," such as interviewee and interviewer characteristics. We included these controls in the regressions to reduce potential response bias. We describe the country sampling frames, their sources, and eligibility criteria in appendixes A and C. Some hospitals are part of larger networks, so in our analysis, we clustered standard errors by hospital network to take into account potential similarities across these hospitals.⁷

B. Collecting Hospital Health Outcomes

Given the absence of publicly comparable measures of hospital-level performance across countries, we collected country-specific measures of mortality rates from AMI (acute myocardial infarction, commonly called heart attacks). AMI is a common emergency condition, recorded accurately, and believed to be strongly influenced by the organization of hospital care (Kessler & McClellan, 2000), and used as a standard marker of clinical quality. We tried to create a consistent measure across countries, although there are inevitably some

⁶See, for example, the video of the training for our 2009 wave: http://worldmanagementsurvey.org/?page_id=187.

⁷In the U.K. sample, we have two years (2006 and 2009), so clustering also deals with serial correlation over time in the same network.

differences in construction, so we include country dummies in almost all of our specifications.⁸ We observe substantial differences in the spread of this measure across countries: the country-specific coefficient of variation is 0.51 for Brazil, 0.52 for Canada, 0.21 for Sweden, 0.10 for the United States, and 0.34 (2006) and 0.15 (2009) for the United Kingdom.

C. Classifying Differences across Universities

In the WHED we can distinguish whether universities offer courses in business (management, administration, entrepreneurship, marketing, or advertising), medical (clinical), and humanities (arts, language, and/or religion) and a range of other divisions (see Feng & Valero, 2020; Valero & Van Reenen, 2019). We geocode the location of each school using its published address and compute drive times between hospitals and universities of different types using GoogleMaps. The computation of travel times is restricted to hospitals and universities in the same county (see appendix A for a more detailed explanation).

D. Collecting Location Characteristics Information

Using the geographic coordinates of hospitals in our sample, we also collected a range of other location characteristics. At the subnational regional level (e.g., states in the United States), we use the variables provided in Gennaioli et al. (2013).⁹ For data at the grid level, we construct a data set based on the G-Econ Project at Yale that estimates geographical measures for each grid cell, which represents 1 degree in latitude by 1 degree longitude. Table B1 presents descriptive statistics for the sets of location characteristics used in this analysis.

III. Descriptive Statistics

A. Variation in Management Practices

Table 1 shows some descriptive statistics, and figure 1 shows the differences in management scores across countries (which is the simple average of the questions ranging between 1 and 5). The United States has the highest management

⁸For Brazil, we compute a simple risk-adjusted measure by taking the unweighted average across rates for myocardial infarction specified as acute or with a stated duration of four weeks or less from onset for each race-gender-age cell for each hospital for the years of 2012 and 2013. For Canada, we use risk-adjusted rate for AMI mortality for the years 2004–2005, 2005–2006, and 2006–2007. For Sweden, we use the 28-day case fatality rate from myocardial infarction from 2005 to 2007. For the United States, we use the risk-adjusted 30-day death (mortality) rates from heart attack from July 2005 to June 2008. For the United Kingdom, we use 30-day risk-adjusted mortality rates purchased from the company "Dr Foster," the leading provider of NHS clinical data. (See appendix A for more information and sources.) For each hospital, we consider three years of data (the survey year plus two years preceding, or the closest years to the survey with available data) to smooth over possible large annual fluctuations.

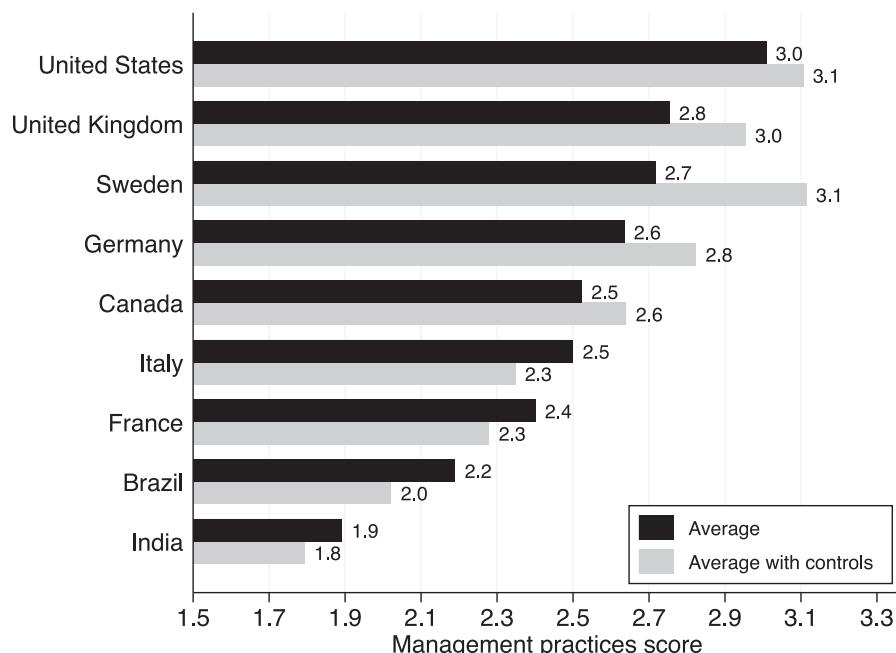
⁹The regional data from Gennaioli et al. (2013) consist of NUTS1, NUTS2, state, or provincial level, depending on the country.

TABLE 1.—DESCRIPTIVE STATISTICS

	Mean	Median	SD	Minimum	Maximum
Hospital characteristics					
AMI mortality rate (<i>z</i> -score)	0.02	-0.08	1.01	-2.2	4.8
Management practice score	2.42	2.4	0.65	1	4.3
Management practice score (<i>z</i> -score)	-0.02	-0.04	1.01	-2.2	3
Hospital beds	270.2	132.5	365.4	6	4,000
Share of managers with MBA-type course	0.26	0.15	0.29	0	1
Number of competitors: 0	0.14	0	0.35	0	1
Number of competitors: 1 to 5	0.61	1	0.49	0	1
Number of competitors: More than 5	0.24	0	0.43	0	1
Dummy public	0.51	1	0.5	0	1
Dummy private for profit	0.3	0	0.46	0	1
Dummy private not for profit	0.19	0	0.39	0	1
Distances to universities					
Driving hours, nearest joint medical-business schools	1.16	0.65	1.84	0	41.8
Driving distance (km) to nearest joint medical-business schools (M-B)	80.32	36.64	135.41	0	2,842.4
Driving hours, nearest business school, no medical school	1.46	0.86	2.16	0	44.4
Driving hours, nearest medical school, no business school	1.47	0.89	2.2	0	44.4
Driving hours, nearest school, no medical or business school	1.24	0.71	2.06	0	44.4
Driving hours, nearest stand-alone humanities school	1.86	1.14	2.42	0	44.4
Driving hours, nearest university in general	0.62	0.32	1.47	0	41.8

These are descriptive statistics of the main variables used in the analysis. The maximum sample size is 1,960. More descriptive statistics are in table B1.

FIGURE 1.—MANAGEMENT PRACTICES ACROSS COUNTRIES



This figure shows the country average management score on a scale of 1 to 5 (all twenty individual questions are averaged within a hospital, and then the unweighted average is taken across all hospitals within a country). The dark bar is this simple average, and the lighter gray bar controls for various characteristics. Controls include log of the number of hospital beds, ownership (for profit, nonprofit, and government), survey noise controls (interviewee seniority, tenure, department and type—nurse, doctor, or nonclinical manager; interview duration and year; an indicator of the reliability of the information (as coded by the interviewer), and 21 interviewer dummies. Number of observations: Brazil = 286, Canada = 174, France = 147, Germany = 124, India = 490, Italy = 154, Sweden = 43, United Kingdom = 235, and United States = 307.

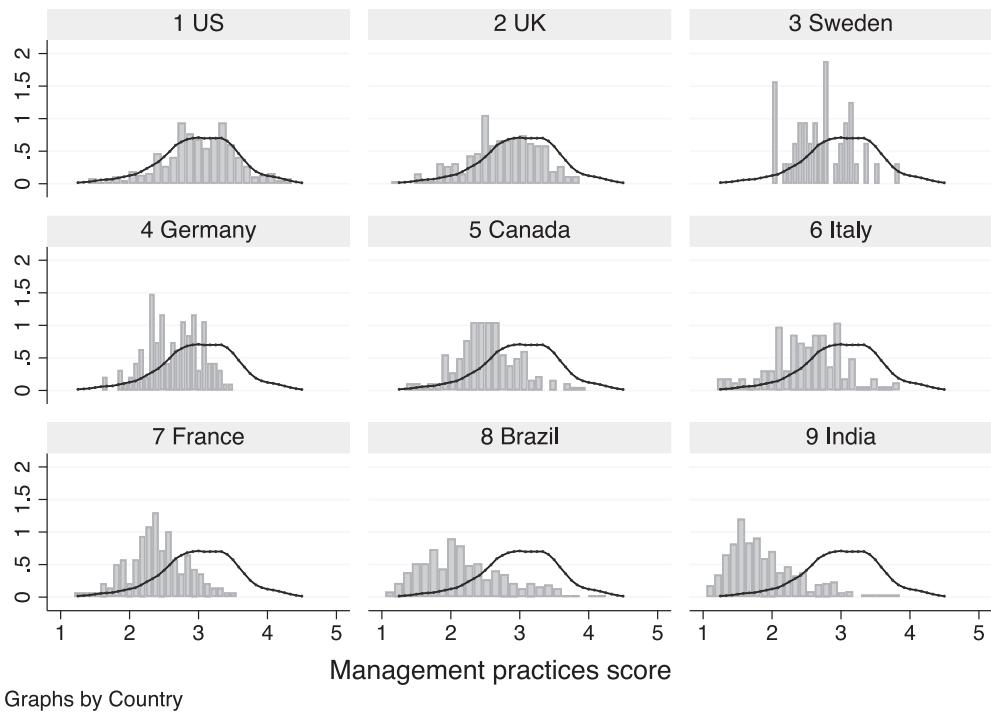
score (3.0), closely followed by the United Kingdom, Sweden, and Germany (all around 2.7) with Canada, Italy, and France slightly lower (at around 2.5). The emerging economies of Brazil (2.2) and India (1.9) have the lowest scores.¹⁰ The rankings do not change substantially (except for Sweden, which rises to the top) when we include controls

for hospital characteristics and interview noise. Country fixed effects are significant (*p*-value on the *F*-test of joint significance is 0.00) and account for 32% of the variance in the hospital-level management scores, which is a greater fraction than for manufacturing firms, where the figure is 25% for the same set of countries.¹¹

¹⁰In the appendix, we provide examples of management practices in the average hospital in the United States (at the top of the ranking) and in India (at the bottom of the ranking).

¹¹One possible explanation is that manufacturing firms often produce an internationally traded good, so firms are more globally exposed while hospitals serve local markets. Table C2 presents hospital characteristics across countries. Although there are many differences in cross-country means (e.g.,

FIGURE 2.—MANAGEMENT PRACTICES WITHIN COUNTRIES



This figure shows the histogram of hospital management scores (the simple average over the twenty questions) within each country. The smoothed kernel of the distribution for the United States is shown in each panel. Number of observations: Brazil = 286, Canada = 174, France = 147, Germany = 124, India = 490, Italy = 154, Sweden = 43, United Kingdom = 235, and United States = 307.

Figure 2 shows the distribution of management scores within each country compared to the smoothed (kernel) fit of the U.S. distribution. Across OECD countries, lower average country-level management scores are associated with an increasing dispersion toward the left tail of the distribution. While the fraction of hospitals with very weak management practices in OECD countries is small (from 5% in the United States to 18% in France), this fraction rises to 45% in Brazil and 68% in India. At the other end of the distribution, the fraction of hospitals with a score of 3 or above ranges from 50% in the United States to 3% in India.

We examined the relationship between the management score and hospital characteristics when country dummies and noise controls are included (coefficients and confidence intervals reported in appendix figure A1). Larger hospitals (where size is proxied by the log of number of beds) tend to have higher management scores, whereas government-run hospitals tend to have lower management scores relative to for private for-profit and private nonprofit hospitals. Bloom, Propper, Seiler, and Van Reenen (2015) show causal evidence of the impact of higher competition on improved managerial quality in English hospitals. Consistent with this earlier research, we find that the self-reported measure of competition we collected during the interview is positively and

significantly correlated with the management score.¹² The magnitude and significance of these correlations are largely unchanged when these variables are jointly included in the regression.

B. AMI Mortality Rates and Management

As an external validation of our management measure across countries, we investigate whether management is related to clinical outcomes. Table 2 shows that management practices are significantly negatively correlated with AMI mortality rates.¹³ In column 1, the management coefficient suggests that a 1 SD increase in a hospital's management score is associated with a fall of -0.188 SD in AMI deaths rates, and this relationship holds even after controlling for a wide variety of factors. Column 2 includes a measure of size (hospital beds), ownership dummies (for profits, non-profits, and government owned), local competition faced by the hospital, and statistical noise controls. Column 3 includes regional geographic controls (e.g., income per capita, education, population density, climate, ethnicity). Column 4 includes regional dummies, and column 5 uses more

¹²Our measure of competition is collected during the survey by asking the interviewee, "How many other hospitals with the same specialty are within a 30-minute drive from your hospital?"

¹³Note that we can do this for only a subset of hospitals (477 from the total of 1,960 observations), as AMI data are not available for all hospitals. The results discussed in this section—and, in particular, the relationship between AMI mortality rates and management—are similar if we focus only on the cardiology subsample.

the median French hospital has 730 beds compared to 45 in Canada), within all countries, nonresponders were not significantly different from participating hospitals. Characteristics are different because the health care systems differ, and our sample reflects this.

TABLE 2.—AMI MORTALITY RATES ARE CORRELATED WITH MANAGEMENT PRACTICES

Dependent Variable: AMI mortality rate (z-score)	(1)	(2)	(3)	(4)	(5)
Z(Mgmt)	-0.185*** (0.055)	-0.201*** (0.065)	-0.189*** (0.064)	-0.189*** (0.070)	-0.195*** (0.065)
ln(Hospital beds)		-0.045 (0.081)	-0.048 (0.084)	-0.099 (0.090)	-0.064 (0.084)
Dummy private for profit		-0.121 (0.206)	-0.119 (0.209)	0.012 (0.268)	-0.047 (0.219)
Dummy private nonprofit		-0.341** (0.147)	-0.275** (0.138)	-0.202 (0.143)	-0.226 (0.144)
Omitted base is government owned					
Noise controls		Yes	Yes	Yes	Yes
Other hospital characteristics		Yes	Yes	Yes	Yes
Geographic controls at the regional level			Yes	Yes	
Geographic controls at the grid level					Yes
Observations	477	477	477	477	477
Number of clusters	397	397	397	397	397
Fixed effects (number)	country(5)	country(5)	country(5)	region(75)	country(5)
R ²	0.02	0.16	0.20	0.34	0.18

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All columns estimated by OLS. Standard errors clustered by hospital network in parentheses. Dependent variable $z(\text{AMI})$ refers to a pooled measure of country-specific AMI mortality rates (measures are standardized by country and year of survey). $Z(\text{Mgmt})$ refers to the hospital's z -score of management (the z -score of the average z -scores of the twenty management questions). Noise controls include interviewee seniority, tenure, department (orthopedics, surgery, cardiology, or other) and type (nurse, doctor, or nonclinical manager), year and duration of the interview, an indicator of the reliability of the information as coded by the interviewer, and 21 interviewer dummies. Hospital characteristics include number of competitors constructed from the response to the survey question on number of competitors, coded as 0 for none (16% of responses), 1 for less than five (59% of responses), and 2 for five or more (25% of responses). Geographic controls at the regional level include log of income per capita, years of education, share of population with high school diploma, share of population with college degree, population, temperature, inverse distance to coast, log of oil production per capita, and log of number of ethnic groups. Geographic controls at the grid level include log of gross product per capita, 2005 USD at market exchange rates, log of gross product per capita, 2005 USD at purchasing power parity exchange rates, distance to major navigable river, distance to ice-free ocean, average precipitation, average temperature, and elevation. Whenever one of these two sets of geographic controls is added, hospital latitude, hospital longitude, and population density within 100 km radius are also added.

disaggregated geographical controls. Although the coefficient on management varies between columns (from -0.185 to -0.201), it is always significant at the 1% level.

In additional analysis (available upon request), we investigated whether the relationship between AMI mortality rates and management was heterogeneous across countries. Overall, the results indicate that the coefficients are in fact similar across countries. Further, to provide a sense of the magnitudes implied by these coefficients, we rerun this regression using raw (i.e., non- z -scored) AMI mortality rates on the U.S. sample, which provides the largest number of hospitals with risk-adjusted AMI data. In this sample, a 1 SD change in the management score is associated with a reduction of 0.320 (standard error 0.173) in the AMI mortality rate. This third of a percentage point fall in AMI death rates compares to a mean of 16% and a standard deviation of 1.75 (implying a share of the standard deviation of $0.18 = 0.32/1.75$, nearly identical to the pooled correlation in column 1 of table 2).¹⁴

Table 2 is broadly consistent with findings from prior quantitative work in this area. For example, Bloom et al. (2015) look at management practices in English hospitals in 2006 and also find a positive link between management and hospital performance, such as survival rates from general surgery, lower staff turnover, lower waiting lists, shorter lengths of stay, and lower infection rates. McConnoll et al. (2013) document a negative and significant relationship between management (measured using the WMS survey instrument) and AMI

mortality rates in the context of 597 cardiac units in the United States. Chandra et al. (2016) look at the WMS management scores and risk-adjusted AMI mortality in U.S. hospitals and also report a negative relationship. The correlations described so far are also in line with existing qualitative studies documenting a positive association between specific aspects of a hospital's organizational culture and AMI mortality rates. For example, in-depth qualitative studies (Bradley et al., 2001) document that hospitals with better performance in terms of adoption of beta-blockers (used to reduce mortality and future cardiac events after AMI) and lower AMI mortality rates tend to have clear and well-communicated goals throughout the organization, make systematic use of problem-solving tools (such as root cause analysis), have greater reliance on data, and have stronger communication and coordination routines relative to low-performing hospitals. These studies also observe that the presence of these different approaches is not fully captured by surveys that simply track adoption of specific clinical protocols or checklists. This is because although these standardized tools are reported to be widely used in both high- and low-performing organizations, there can still be wide variation in the ways in which they are implemented. The results are also consistent with the case study evidence on hospitals like Virginia Mason (Kenney, 2015), ThedaCare (Toussaint, Conway, & Shortell, 2016), and Intermountain (Leonhardt, 2009) that are famous for adopting the types of management practices that we include in the survey and for having better clinical outcomes.

While the causal channels are yet to be fully established—and cannot be discerned in the qualitative research mentioned above or in our sample given the cross-sectional nature of the data—these studies suggest that differences in basic processes and practices such as the ones captured in the WMS

¹⁴For comparison, we also repeated this analysis on the second largest sample with AMI data, Brazil (109 observations), where, however, we could retrieve only non-risk-adjusted AMI rates. In this sample, a standard deviation change in management is associated with a 2.404 decrease in the AMI rate (standard error 0.914), which corresponds to 29% of the standard deviation of the variable (8.23).

instrument may contribute to better clinical performance by focusing attention and resources toward the issue of the quality of care; reducing the likelihood of preventable deaths and medical errors, which are often related to poor communication or imperfect transitions of care; and helping to identify and address the inevitable complexities and risks that arise in patients hospitalized with AMI.

IV. The Role of Managerial Education

In this section, we explore a possible factor behind the variation in management across hospitals and the relationship between the management score and AMI mortality rates: differences in managerial education opportunities among clinical managers.

Exposure to basic managerial training among individuals involved in health care provision is generally low in the United States (Myers & Pronovost, 2017). Although comparable international information on managerial training received by health care professionals is not available, data collected within the management interviews allow us to provide some basic information on the presence and heterogeneity of managerial training among clinical managers employed in acute care hospitals. In particular, we asked the interviewee, “What percentage of managers have an MBA?” and prompting the interviewer to include in this calculation management-related courses that extend over at least six months (this would include, for example, executive education courses that do not lead to a formal MBA degree, such as Johns Hopkins’s master of science in health care management and Georgetown’s certificate in business administration at the School of Continuing Studies). On average 26% of managers are reported to have received managerial training, with a standard deviation of 0.29.

Perhaps unsurprisingly, the variable measuring the share of managers in the hospital who have attended an MBA-type course is positively and significantly correlated with the management score. For example, in a regression model including as additional controls country dummies, proxies for interview noise, and the hospital characteristics examined in figure 2 (hospital size, ownership dummies, and local competition), a 10% increase in the managerial skills variable (e.g., the average hospital moves from having 26% to 28.6% of managers with an MBA-type course) is associated with a 0.059 SD increase in the management score.

Since the fraction of managers with an MBA-type degree in the hospital is likely to be endogenous to the quality of management practices adopted in the hospital, in order to better identify the role of managerial training *per se*, we now turn to analyze alternative—and arguably more exogenous—proxies for the supply of managerial human capital in the hospital. More specifically, we focus on the distance between the hospital and universities. We start by considering the role of all universities (many of which we do not expect to have any particular correlation with clinical outcomes) and then focus on universities offering both clinical and managerial

education as the closest proxy for the courses that would result in a higher supply of managerially trained clinical managers and, potentially, with better clinical outcomes. Table 3 starts by exploring the relationship between these distance metrics and AMI mortality.

Column 1 of table 3 regresses AMI mortality rates on driving hours to the nearest university.¹⁵ Although there is a positive coefficient on distance to a university, it is statistically insignificant. In columns 2 and 3, we focus on a much more specific variable: the distance to universities offering both medical and business courses (henceforth, “Joint M-B school”).¹⁶ Since there could be unobserved heterogeneity specific to university locations confounding the relationship between hospital performance and the distance to universities, we also include driving distance to universities specializing solely in arts, humanities, or religious courses (“stand-alone HUM”) and therefore not offering clinical/medical or business-type courses (and expect to find no significant relationship between these universities and hospital performance). To validate the use of this type of school as a placebo, we verified that the nearest stand-alone HUM school and joint M-B school are similar in proximity to the hospitals in our sample: 82% of hospitals have a driving time difference of two hours or less between these two types of universities (this is shown in figure B2 in the appendix). We also observe that the means of a range of location characteristics of the nearest joint M-B school and stand-alone HUM school are not statistically significant (in table B2).¹⁷ Finally, we also include the driving time to universities that do not offer medical, business, or humanities¹⁸ (“no M, B, HUM”). We find that AMI mortality rates are positively and significantly correlated with the driving distance to a joint M-B school: a 10% increase in the drive time to a joint M-B school is associated with an increase in AMI mortality rates by 0.039 SD. Reassuringly, we do not observe a significant relationship between management and the other university types. Column 3 shows that the relationship between AMI mortality rates and driving distance to a joint M-B school is essentially unchanged when we include a range of geographic characteristics in our specification (such as income, education, population, and temperature). The significance of the joint M-B school in the AMI regressions of table 3 may be due to other nearby universities that do not have medical/clinical or business courses but offer other types of quantitative courses (such as engineering). To investigate this issue, we calculated distances to other schools such as (a) the nearest university offering business courses but no medical/clinical courses (“B school, no M”), (b) the nearest

¹⁵The average driving time between hospitals and universities is 37 minutes, with a median of 19 minutes.

¹⁶We calculate driving distances from each hospital to the nearest joint M-B school, which is 70 minutes on average. The results are qualitatively and quantitatively similar if we run this regression on the subsample of hospitals with AMI data.

¹⁷The only measures that are statistically significant are latitude and longitude.

¹⁸For example, a stand-alone law school, polytechnic school, religious school, or art school.

TABLE 3.—AMI MORTALITY RATES AND MANAGERIAL EDUCATION

Dependent Variable: AMI Mortality Rate (<i>z</i> -score)	(1)	(2)	(3)	(4)	(5)
In(Driving hours, nearest school)	0.036 (0.232)				
In(Driving hours, nearest joint M-B schools)		0.392** (0.169)	0.387** (0.163)	0.356** (0.166)	0.330** (0.156)
In(Driving hours, nearest stand-alone HUM)		-0.083 (0.155)	-0.202 (0.171)		
In(Driving hours, nearest school, no M, B, HUM)		0.071 (0.155)	0.048 (0.158)		
In(Driving hours, nearest B school, no M)				0.054 (0.159)	
In(Driving hours, nearest M school, no B)				0.066 (0.164)	
In(Driving hours, nearest school, no M or B)				-0.196 (0.191)	
Geographic controls at the Regional level					
Observations	477	477	477	477	477
Number of clusters	397	397	397	397	397
Test of equality: Joint M-B = HUM		0.08	0.03		
Test of equality: Joint M-B = B, no M				0.19	
Test of equality: Joint M-B = M, no B				0.28	
Test of joint significance: HUM, no M-B-HUM		0.78	0.48		
Test of joint significance: B, M, No B-M				0.72	
<i>R</i> ²	0.15	0.16	0.20	0.20	0.19

p* < 0.1, *p* < 0.05, ****p* < 0.01. All columns estimated by OLS. Standard errors clustered by hospital network in parentheses. Dependent variable *Z*(AMI) refers to a pooled measure of country-specific AMI mortality rates (measures are standardized by country and year of survey). All columns include noise controls, hospital characteristics, and country dummies. Noise controls include interviewee seniority, tenure, department (orthopedics, surgery, cardiology, or other) and type (nurse, doctor, or nonclinical manager), year and duration of the interview, an indicator of the reliability of the information as coded by the interviewer, and 21 interviewer dummies. Hospital characteristics include log of the number of hospital beds, dummies for private for profit and nonprofit, and number of competitors constructed from the response to the survey question on number of competitors, and is coded as 0 for "none" (16% of responses), 1 for less than five (59% of responses), and 2 for five or more (25% of responses). Geographic controls at the regional level include log of income per capita, years of education, share of population with high school diploma, share of population with college degree, population, temperature, inverse distance to coast, log of oil production per capita, and log of number of ethnic groups. Hospital latitude, hospital longitude, and population density within 100 km radius are also added. M = medical school; B = business school; HUM = humanities school.

university offering medical/clinical courses but not business courses ("M school, no B"), and (c) the nearest university offering other courses but no business or medical courses ("nearest school, no M or B"), and verified that the distributions are similar across all types of schools (figure B3 in the appendix). In column 4 of table 3, we include variables measuring driving distances to all four types of schools. The distance to joint M-B schools has explanatory power over and above distances to other school types, and it has a coefficient similar to the previous column in terms of magnitude. Since none of these other school types are individually or jointly significant (see the bottom rows of the relevant columns), we drop them in column 5, which is our preferred specification.¹⁹

Table 4 explores the relationship between distance to universities and the management practices score. The specifications are the same as for table 3, but with a different dependent variable. There is a negative correlation between distance to the nearest university and management practice scores. As with table 3, columns 2 to 4 show that only the category of joint M-B schools has explanatory power over and above distances to other school types. The results in our preferred

specification in column 5 suggest that a 10% increase in drive time to a joint M-B school is associated with a decrease in hospital management quality of 0.014 SD. These results are qualitatively and quantitatively unchanged when we focus on the subsample of hospitals with AMI data.²⁰

A. Robustness Checks

We investigate the robustness of the relationships discussed in tables 3 and 4 to several potential concerns. Some of these robustness checks are shown in table 5 (the first five columns have AMI as the dependent variable and the last two columns have the management practice score as the dependent variable). First, the distance from schools offering a medical/clinical course may reflect unobservable school characteristics other than the supply of managerial education directed at clinicians and correlated with both clinical quality and management. For example, institutions offering both medical and business education may be systematically different from those that do not in terms of their quality. To look into this issue, we investigated whether schools offering medical and business training are associated with proxies for higher school quality. This analysis is shown in appendix table B3. Schools offering medical and business training are indeed older, more likely to be listed in the Quacquarelli Symonds World University Ranking (QSWUR) in 2011, and

¹⁹To get a sense of these magnitudes, we estimated the relationship between AMI mortality rates and the distance from the closest universities offering M-B courses on the U.S. sample, using the raw (i.e., non-*z*-scored) AMI rates as a dependent variable. In this sample, a 1% increase in distance to the closest M-B school is associated with a 1 point increase in AMI rate (57% of a standard deviation). When we repeated the same exercise in Brazil (109 observations) using the raw non-risk-adjusted AMI rates, the coefficient implies that a 1% increase in the distance metric is associated with a 3.675 increase in AMI mortality rates (45% of a standard deviation).

²⁰The relationship between management and the distance metric is -0.208 (standard error 0.102) in the AMI subsample.

TABLE 4.—HOSPITAL MANAGEMENT SCORE AND MANAGERIAL EDUCATION

Dependent Variable: Management score (<i>z</i> -score)	(1)	(2)	(3)	(4)	(5)
In(Driving hours, nearest school)	-0.139*** (0.045)				
In(Driving hours, nearest joint M-B schools)		-0.124*** (0.043)	-0.114*** (0.044)	-0.109** (0.044)	-0.149*** (0.038)
In(Driving hours, nearest stand-alone HUM)		-0.049 (0.037)	-0.019 (0.039)		
In(Driving hours, nearest school, no M, B, HUM)		-0.065 (0.041)	-0.058 (0.042)		
In(Driving hours, nearest B school, no M)				0.000 (0.041)	
In(Driving hours, nearest M school, no B)				-0.035 (0.043)	
In(Driving hours, nearest school, no M or B)				-0.057 (0.044)	
Geographic controls at the regional level			Yes	Yes	Yes
Observations	1,959	1,959	1,959	1,959	1,959
Number of clusters	1,869	1,869	1,869	1,869	1,869
Test of equality: Joint M-B = HUM		0.24	0.15		
Test of equality: Joint M-B = B, no M				0.09	
Test of equality: Joint M-B = M, no B				0.25	
Test of joint significance: HUM, no M-B-HUM		0.03	0.24		
Test of joint significance: B, M, no B-M				0.34	
<i>R</i> ²	0.60	0.61	0.61	0.61	0.61

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All columns estimated by OLS. Standard errors clustered by hospital network in parentheses. Dependent variable $Z(\text{Mgmt})$ refers to the hospital's *z*-score of management (the *z*-score of the average *z*-scores of the twenty management questions). All columns include noise controls, hospital characteristics, and country dummies. Noise controls include interviewee seniority, tenure, department (orthopedics, surgery, cardiology, or other) and type (nurse, doctor, or nonclinical manager), year and duration of the interview, an indicator of the reliability of the information as coded by the interviewer, and 21 interviewer dummies. Hospital characteristics include log of the number of hospital beds, dummies for private for profit and nonprofit, and number of competitors constructed from the response to the survey question on number of competitors, and is coded as 0 for "none" (16% of responses), 1 for less than five (59% of responses), and 2 for five or more (25% of responses). Geographic controls at the regional level include log of income per capita, years of education, share of population with high school diploma, share of population with college degree, population, temperature, inverse distance to coast, log of oil production per capita, and log of number of ethnic groups. Hospital latitude, hospital longitude and population density within a 100 km radius is also added. M = medical school; B = business school; HUM = humanities school.

more likely to offer postgraduate degrees. Columns 1 and 6 include these additional controls for school quality, and although some of them are significant, their inclusion does not affect the magnitude or significance of the coefficient on the distance to joint M-B schools in either the AMI or the management regressions.

A second issue is that geographical areas with universities offering both clinical and managerial education might be systematically different from those that do not provide these schools—for example, unobserved heterogeneity in income levels might drive both better clinical outcomes and higher levels of the management score.²¹ This could bias our results to the extent that the regional controls included in our analysis are not able to capture these finer differences in geographical characteristics. Columns 2 and 7 of table 5 include regional dummies in the specification and shows that the coefficient on distance to a joint M-B school is still statistically significant when these controls are included.²²

Third, we investigated the robustness of the relationship between AMI mortality rates and the distance metric to the inclusion of county-level Census-based controls for differences in the skill composition, employment composition in manufacturing and health care, unemployment rate, employ-

ment growth rate, and per capita income levels. We performed this analysis for the population of U.S. hospitals because of the availability of both AMI data and detailed Census variables (this analysis does not require the availability of the management data—hence, the larger sample).²³ When using the specification of column 5 in table 3 on this U.S. sample, the coefficient (standard error) on distance is 0.454 (0.111). When we include Hospital Referral Regions (HRR) dummies in column 3 of table 5, the coefficient on the distance metric decreases slightly to 0.404, and when we include county-level controls in column 4, the coefficient (standard error) drops to 0.232 (0.125).²⁴

Overall, these results suggest that while regional differences are important, they cannot fully account for the relationship between clinical outcomes and the availability of schools offering managerial and clinical education. Finally, we checked whether the robustness of the relationship between AMI mortality rates and the distance metric captured unobservable characteristics of the parent organization (e.g.,

²¹We use a sample of hospitals in the United States for which AMI measures are reported in 2009, our year of reference for the OECD countries. We approximate the sample used in the United States to the cross-country sample used in this paper by excluding sole community providers and hospitals operated by the Catholic Church.

²²In this specification we use as county-level controls (all measured in 2009) employment in manufacturing (coefficient, -0.440; SE, 0.470); employment in health care (coefficient, -0.521; SE, 0.515); %25+ with bachelor's degree or higher (coefficient, -0.010; SE, 0.004); log per capita income (coefficient, -0.608; SE, 0.177); unemployment rate (coefficient, -0.018; SE, 0.014); employment growth (2000–2009) (coefficient, -1.443; SE, 2.215).

²¹Differences in income per capita across areas may also affect the quality of emergency care infrastructures across hospitals, thus increasing the speed of arrival of patients at the hospital and improving their clinical outcomes.

²²Within-country regional dummies are of a full set of dummies at the NUTS 2 level for France, Germany, Italy, Sweden, and the United Kingdom, and an equivalent state- or provincial-level division for Brazil, Canada, India, and the United States.

TABLE 5.—ROBUSTNESS CHECKS

	(1) Z(AMI)	(2) Z(AMI)	(3) Z(AMI)	(4) Z(AMI)	(5) Z(AMI)	(6) Z(Mgmt)	(7) Z(Mgmt)
ln(D-hours to joint M-B)	0.332** (0.160)	0.454** (0.196)	0.404*** (0.130)	0.232* (0.125)	0.287* (0.161)	-0.145*** (0.038)	-0.165*** (0.045)
Measures of university quality							
ln(Age of joint M-B)	0.043 (0.092)					0.044** (0.020)	
Global QS Rank Dummy	0.497 (0.508)					0.240* (0.126)	
ln(Reversed Global QS Rank)	-0.056 (0.091)					-0.040* (0.023)	
Offers postgraduate degree dummy	0.203 (0.129)					0.004 (0.059)	
Noise controls	Yes	Yes				Yes	Yes
Hospital characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls at the regional level	Yes						
Geographic controls at the grid level		Yes					Yes
Observations	477	477	2,011	2,011	1,178	1,959	1,959
Number of clusters	397	397	732	732	213	1,869	1,869
Fixed effects	Country	Region	HRR		Network	Country	Region
Sample	WMS	WMS	U.S. AHA	U.S. AHA	U.S. AHA	WMS	WMS
<i>R</i> ²	0.20	0.37	0.24	0.10	0.36	0.62	0.66

p* < 0.1, *p* < 0.05, ****p* < 0.01. All columns estimated by OLS. Standard errors clustered by hospital network in parentheses. Dependent variable Z(AMI) refers to a pooled measure of country-specific AMI mortality rates (measures are standardized by country and year of survey). Dependent variable Z(Mgmt) refers to the hospital's *z*-score of management (the *z*-score of the average *z*-scores of the twenty management questions). Noise controls include interviewee seniority, tenure, department (orthopedics, surgery, cardiology, or other) and type (nurse, doctor, or nonclinical manager), year and duration of the interview, an indicator of the reliability of the information as coded by the interviewer, and 21 interviewer dummies. Hospital characteristics include log of the number of hospital beds, dummies for private and nonprofit, and number of competitors constructed from the response to the survey question on number of competitors, and is coded as 0 for "none" (16% of responses), 1 for less than five (59% of responses), and 2 for five or more (25% of responses). Geographic controls at the regional level include log of income per capita, years of education, share of population with a high school diploma, share of population with a college degree, population, temperature, inverse distance to coast, log of oil production per capita, and log of number of ethnic groups. Geographic controls at the grid level include log of gross product per capita, 2005 USD at market exchange rates, log of gross product per capita, 2005 USD at purchasing power parity exchange rates, 2005, distance to major navigable river, distance to ice-free ocean, average precipitation, average temperature, and elevation. Whenever one of these two sets of geographic controls is added, hospital latitude, hospital longitude, and population density within 100 km radius are also added.

better-managed chains of hospitals may proactively locate their hospitals in areas providing a greater supply of clinicians with managerial training). To do so, we focused again on the U.S. sample, where we could obtain close to population information of network affiliations using the AHA register and, within the United States, on the hospitals in the sample that belong to networks. Within this sample, we added to the specification network fixed effects in column (5) of table 5. This exploits within-network variation in AMI mortality rates and distance to schools, thus controlling for possible network-level confounders (the sample is smaller, as we require at least two hospitals in the chain for which performance data were available).²⁵ These results confirm that greater distance to joint M-B schools is associated with higher AMI mortality rates.²⁶ Overall, these basic robustness checks provide reassurance that the relationship between the distance metrics and our variables of interest does not proxy for basic differences in university quality, regional characteristics, and network-level heterogeneity.

²⁵This is analogous to a manufacturing context where one could use plant-specific variation within a firm (i.e., firm fixed effects with plant-level data).

²⁶We also repeat the specification in column 8 but add HRR fixed effects to check if our results are robust to market characteristics and find similar results. Using a larger U.K. sample, we explore another dimension of hospital performance: the average probability of staff intending to leave in the next year as a measure of worker job satisfaction for the United Kingdom reported by the NHS staff surveys and used on Bloom et al. (2015). Reassuringly, we find similar patterns to those described in table 3, indicating a significant, positive correlation between distance to the nearest joint M-B school and the likelihood of the average employee wanting to leave the hospital.

B. Business Education

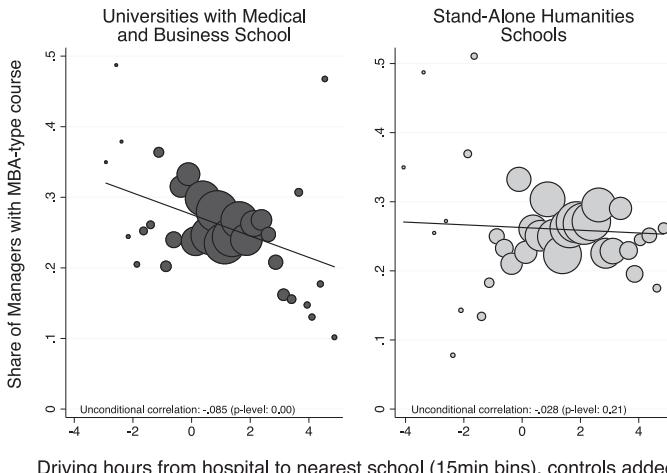
What could be the reason for the relationship between distance from universities providing medical and business education and better hospital outcomes (in terms of AMI survival rates and management practices)? One obvious mechanism is that there is a greater supply of workers with managerial skills when a hospital is close to a joint M-B school.

In figure 3 we investigate the relationship between the share of managers with an MBA-type degree and the hospital's closeness to a joint M-B school (left-hand side).²⁷ There is a clear downward slope: being closer to these types of schools is associated with a higher fraction of managers with MBAs. By contrast, the right-hand-side panel of figure 3 shows that there is no relationship between the share of MBAs and the distance to stand-alone HUM schools. We formalize figure 3 in appendix table B5. Consistent with the two earlier tables, closeness to a joint M-B school (but not other types of school) is associated with significantly more hospital managers with business education.²⁸

²⁷All variables in figure 3 are orthogonalized off geographical controls through a first-stage regression.

²⁸One way to bring these ideas together is by instrumenting the share of MBA with the distance to a joint M-B school, reflecting the idea that proximity increases the managerial skill supply, which in turn benefits hospital performance. If the only way that university proximity matters is through skill supply, this should identify the causal impact of managerial education on hospital performance. With the important caveats that the exclusion restriction may not be valid (universities could in principle affect hospitals through routes other than the supply of human capital) and that the instrument is not strong, we observe that results are consistent with a large causal effect (see appendix table B4).

FIGURE 3.—SHARE OF MANAGERS WITH MBA-TYPE COURSE AND DRIVING HOURS TO NEAREST SCHOOL



Each panel shows the mean share of managers with MBA-type courses in a hospital (vertical axis) as a function of the drive time to the nearest type of school. Mean of share of managers with MBA-type courses and travel time in 15-minute bins. Controls include noise controls: interviewee seniority, tenure, department (orthopedics, surgery, cardiology, or other), and type (nurse, doctor, or nonclinical manager), year and duration of the interview, an indicator of the reliability of the information as coded by the interviewer, and 21 interviewer dummies, and geographic controls at the regional level: log of income per capita, years of education, share of population with high school diploma, share of population with college degree, population, temperature, inverse distance to coast, log of oil production per capita, and log of number of ethnic groups. Excludes 31 hospitals with driving hours longer than five hours. Weighted markers represent the number of hospitals in each bin. Unadjusted correlation with the full sample of 1,960 observations is at the bottom of each panel.

V. Conclusion

We have collected data on management practices in 1,960 hospitals in nine countries. We document a large variation of these management practices within each country and find that our management index is positively associated with improved clinical outcomes as measured by survival rates from AMI. We show evidence that a hospital's proximity to a university that supplies joint business and clinical education is associated with a higher management practice score (and better clinical outcomes). Proximity to universities that do not have medical schools or do not have business schools does not statistically matter for hospital management scores, suggesting that the bundle of managerial and clinical skills has an impact on hospital management quality. We find that hospitals that are closer to the combined clinical and business schools also have a higher fraction of managers with MBAs, which is consistent with this interpretation.

Our work suggests that management matters for hospital performance and that the supply of managerial human capital may be a way of improving hospital productivity. Given the enormous pressure health systems are under, this may be a complementary way of dealing with health demands in addition to the usual recipe of greater medical inputs. The cross-sectional nature of our data does not allow us to rule out sophisticated sources of endogeneity, including the possibility that universities may create managerial programs catering to clinicians in response to the presence of a high-quality hospital in the area. Panel or experimental evidence would help to track out causal impacts. Such evidence from either randomized control trials or natural experiments is an obvi-

ous next step in this agenda. Furthermore, the current data consist primarily of one observation per hospital, under the assumption that different departments and hierarchical levels within a specific hospital should share broader organizational characteristics. Future research should explore this assumption empirically and investigate in further detail the scope for managerial differences not only across but also within hospitals. Finally, it would be valuable to study in much more detail the relationship between basic management practices and the implementation of specific clinical protocols (e.g., surgery checklists) to develop a better understanding of the way in which management affects the day-by-day routines of clinicians. We leave these exciting topics for further research.

REFERENCES

Baicker, Katherine, and Amitabh Chandra, "Medicare Spending, the Physician Workforce, and the Quality of Healthcare Received by Medicare Beneficiaries," *Health Affairs* 24 (2004), 184–197.

Bertrand, Marianne, and Antoinette Schoar, "Managing with Style: The Effect of Managers on Firm Policies," *Quarterly Journal of Economics* 118 (2003), 1169–1208.

Black, Sandra, and Lisa Lynch, "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity," *this REVIEW* 83 (2001), 434–445.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen, "The New Empirical Economics of Management," *Journal of the European Economic Association* 12 (2014), 835–876.

Bloom, Nicholas, Carol Propper, Stephan Seiler, and John Van Reenen, "The Impact of Competition on Management Quality: Evidence from Public Hospitals," *Review of Economic Studies* 82 (2015), 457–489.

Bloom, Nicholas, and John Van Reenen, "Measuring and Explaining Management Practices across Firms and Nations," *Quarterly Journal of Economics* 122 (2007), 1351–1408.

Bradley Elizabeth H., Eric S. Holmboe, and Jennifer Mattera, "A Qualitative Study of Increasing β -Blocker Use after Myocardial Infarction," *Journal of the American Medical Association* 285 (2001), 2604–2611.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson, "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector," *American Economic Review* 106 (2016), 2110–2144.

Chandra, Amitabh, Douglas O. Staiger, and Jonathan Skinner, "Saving Money and Lives," in Pierre L. Yong, Robert S. Saunders, and LeighAnne Olsen, eds., *The Healthcare Imperative: Lowering Costs and Improving Outcomes*, Institute of Medicine (Washington, DC: National Academies Press, 2010).

Doyle, Joe, Steven Ewer, and Todd Wagner, "Returns to Physician Human Capital: Evidence from Patients Randomized to Physician Teams," *Journal of Health Economics* 29:6 (2010), 866–882.

Eisenberg, John M., "Physician Utilization: The State of Research about Physicians' Practice Patterns," *Medical Care* 40 (2002), 1016–1035.

Feng, Andy, and Anna Valero, "Skill Biased Management: Evidence from Manufacturing Firms," *Economic Journal* ueaa005 (2020), <https://doi.org/10.1093/ej/ueaa005>.

Finkelstein, Amy, Matt Gentzkow, and Heidi Williams, "Sources of Geographic Variation in Healthcare: Evidence from Patient Migration," *Quarterly Journal of Economics* 131 (2016), 1681–1726.

Fisher, Elliott S., David Wennberg, Theresa Stukel, Daniel Gottlieb, F. L. Lucas, and Etoile L. Pinder, "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality and Accessibility of Care," *Annals of Internal Medicine* 138 (2003), 273–287.

Gawande, Atul, *The Checklist Manifesto* (New York: Holt, 2009).

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silvanos, and Andrei Shleifer, "Human Capital and Regional Development," *Quarterly Journal of Economics* 128:1 (2013), 105–164.

Goodall, Amanda, "Physician-Leaders and Hospital Performance: Is There an Association?" *Social Science and Medicine* 73 (2011), 535–539.

Huselid, Mark, "The Impact of Human Resource Management Practices on Turnover, Productivity and Corporate Financial Performance," *Academy of Management Journal* 38 (1995), 635–672.

Ichniowski, Casey, Kathryn Shaw, and Prennushi, Giovanni, "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," *American Economic Review* 87 (1997), 291–313.

Kenney, Charles, *A Leadership Journey in Health Care: Virginia Mason's Story* (Boca Raton, FL: CRC Press, 2015).

Kessler, Daniel P., and Mark B. McClellan, "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics* 115 (2000), 577–615.

Leonhardt, David, "Dr. James Will Make It Better," *New York Times*, November 8, 2009.

McConnell K. John, Richard C. Lindrooth, Douglas R. Wholey, Thomas Maddox, and Nick Bloom, "Management Practices and the Quality of Care in Cardiac Units," *JAMA Internal Medicine* 173 (2013), 684–692.

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

Myers, Christopher G., and Peter J. Pronovost, "Making Management Skills a Core Component of Medical Education," *Academic Medicine* 92 (2017), 582–584.

Osterman, Paul, "How Common Is Workplace Transformation and Who Adopts It?" *Industrial and Labor Relations Review* 47 (1994), 173–188.

Phelps, Charles, and Cathleen Mooney, "Variations in Medical Practice Use: Causes and Consequences," in Richard Arnould, Robert Rich, and William White, eds., *Competitive Approaches to Health Care Reform* (Washington, DC: Urban Institute Press, 1993).

Sirovich, Brenda, Patricia M. Gallagher, David E. Wennberg, and Elliott S. Fisher, "Discretionary Decision Making by Primary Care Physicians and the Cost of U.S. Health Care," *Health Affairs* 27 (2008) 813–823.

Skinner, Jonathan, "Causes and Consequences of Regional Variations in Health Care" (pp. 45–93), in Mark V. Pauly, Thomas G. McGuire, and Pedro P. Barros, eds., *Handbook of Health Economics*, vol. 2 (Amsterdam: Elsevier, 2011).

Toussaint, John, Patrick Conway, and Stephen Shortell, "The Toyota Production System: What Does It Mean, and What Does It Mean for Health Care?" *Health Affairs* blog (2016).

Valero, Anna, and John Van Reenen, "The Economic Impact of Universities: Evidence from across the Globe," *Economics of Education Review* 68 (2019), 53–67.