

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Agglomeration Economics

Volume Author/Editor: Edward L. Glaeser, editor

Volume Publisher: The University of Chicago Press

Volume ISBN: 0-226-29789-6

Volume URL: <http://www.nber.org/books/glae08-1>

Conference Dates: November 30-December 1, 2007

Publication Date: February 2010

Chapter Title: Understanding Agglomerations in Health Care

Chapter Author: Katherine Baicker, Amitabh Chandra

Chapter URL: <http://www.nber.org/chapters/c7986>

Chapter pages in book: (211 - 236)

Understanding Agglomerations in Health Care

Katherine Baicker and Amitabh Chandra

7.1 Introduction

Understanding the drivers of productivity differences across areas is crucial to designing effective public policies to promote growth and efficient use of resources. Knowledge spillovers and economies of scale may be key causes of differences in productivity across space. The agglomeration economies literature explores the positive link between productivity and city size or density: cities, by virtue of their density, may facilitate the generation, transmission, and acquisition of new ideas. This is the Marshallian notion of “knowledge spillovers,” where one’s neighbors influence one’s adoption of new technologies so that cities should be more productive places. Most of this research, however, focuses on aggregate productivity measures, such as average income. These measures can be both crude and affected by common local factors, such as price levels, which makes it difficult to discern from them the mechanisms that drive productivity differences.

In this chapter, we explore the drivers of differences in medical sector productivity to understand agglomeration economies better—particularly, the role that information spillovers play in making some places more productive. The medical sector is one of the largest in the U.S. economy (comprising 17 percent of gross domestic product in 2008) and exhibits dramatic differences in productivity across space. The enormous variation in the quantity and

Katherine Baicker is professor of health economics at the Harvard School of Public Health and a research associate of the National Bureau of Economic Research. Amitabh Chandra is professor of public policy at the John F. Kennedy School of Government, Harvard University, and a faculty research fellow of the National Bureau of Economic Research.

Paper prepared for the NBER conference on Cities in February 2008. We have benefited from conversations with Ed Glaeser, Ashish Jha, and Jonathan Skinner and acknowledge research support from the Taubman Center at the Harvard Kennedy School.

quality of care received by patients in different geographic areas is not just due to different health insurance coverage or patient characteristics. Even within Medicare, the public insurance program for the elderly that provides relatively uniform incentives to providers and patients, there is enormous geographic variation in the use of productive and unproductive health care (Dartmouth Atlas Project). Some of this variation occurs at the level of hospital referral regions (HRRs), which are empirically defined markets for health care, but a large portion may occur within these regions. In Baicker and Chandra (2004a), we demonstrate that areas with the highest quality care are often not the areas that spend the most intensively on health care: in states where Medicare spends more on beneficiaries, they are less likely to receive high-quality care. Furthermore, it is not the case that areas that “do more” do more of everything. Rather, U.S. health care is characterized by variation in both the *overuse* of intensive, costly care that is often of dubious clinical benefit (Fisher et al. 2003a, 2003b) and the *underuse* of effective, high-quality care, such as the administration of beta-blockers after heart attacks, mammograms for older women, influenza vaccines, and eye exams for diabetics. These underused procedures are relatively inexpensive, are known to have significant medical benefits, and are rarely contraindicated. Whether cities are better at avoiding the problems of underuse and overuse in health care is not known. The combined evidence on the overuse of dubious care and the underuse of effective care strongly suggests the opportunity for productivity improvements in health care.

While ideally, all patients would receive the highest possible quality of care, in the presence of variation in care, there may be spillover benefits to any particular hospital providing higher-quality care: one hospital’s investment in quality might drive learning across organizations, improving the quality of care provided in neighboring hospitals. By the same token, the use of financially lucrative interventions of dubious therapeutic value may diffuse through a similar process. Prior work in other industries suggests that organizations often learn from each other and that innovations at one company can drive similar innovations in its neighbors. While this potential learning mechanism has been found in many settings, however, it remains largely unexplored in understanding changes in health care quality. Whether that occurs in clinical care and whether high-performing hospitals drive improvements in care among neighbors is largely unknown.

Understanding the primary sources of variation in quality and the role that learning might play in driving improvements is critically important in developing effective public policies: policymakers will be best able to craft interventions that enhance the productivity of American health care spending if they understand the factors that cause some hospitals to lag behind and the channels through which they might catch up. For example, if the variation in care occurs primarily at the hospital level, understanding the degree to which hospital characteristics (such as teaching or profit status)

affect the quality of care can help target interventions toward those organizations where improvements are most needed. If the variation is primarily regional, other approaches will be needed. Finally, understanding the role that learning plays can provide insight into policies that might optimize the quality of care delivered across an entire community: if quality enhancements at one hospital are adopted by others, then this positive externality might suggest that it would be optimal to subsidize the investment in quality. On the other hand, if utilization of more expensive but lower-value care also diffuses through a similar learning process, then policies might aim to discourage hospital investments in technologically intensive care of dubious value. In order to better understand the roles of variation and learning in the hospital setting, we seek to answer three questions. First, how much variation in hospital care (measured by both quality and the use of low-value care) is due to differences across regions versus differences within regions? Second, how much of the variation within and across regions can be explained by readily identifiable factors? And finally, what is the role of agglomeration in generating quality? Is there evidence that the quality of care in an institution is influenced by the care at neighboring institutions?

7.2 Background

There is a vast economic literature on diffusion of innovation (e.g., Jovanovic and Nyarko 1995, 1996); see Hall (2004) for a recent survey and synthesis. Nearly all of these economic models posit that diffusion evolves as a profit-maximizing strategy, such as because some producers are waiting for the price of the innovation to drop, or because they had invested in the previous technology, or because of capital-labor complementarity. For example, Zucker, Darby, and Brewer (1998) link the timing of entry and location of biotechnology firms to the presence of academics that publish in basic science journals, while Jaffe, Trajtenberg, and Henderson (1993) document that inventors tend to cite patents that were developed in the same geographic region. In medical care, there is a large literature on knowledge spillovers, beginning with Coleman, Katz, and Menzel (1957), who found that doctors who were more integrated with colleagues were the first to adopt a new drug.

In a world where the barriers to adoption are financial, individual health providers each make decisions about whether to adopt the new innovation. Their incentives to adopt the innovation depend critically on the nature of their own perceived production function—which in turn reflects how rapidly they themselves absorb new clinical evidence on treatments—and the financial structure and rewards of the innovation. When the innovation is both financially remunerative and perceived to be highly effective, the speed of adoption will be rapid; as, for example, the rapid diffusion of tetracy-

cline in the 1950s, documented in the classic Coleman, Katz, and Mendel study (1957). If there are few economic incentives to adopt, as in the case of beta-blockers for the treatment of heart attacks, then diffusion will depend almost entirely on the extent to which individual perceptions of physicians change over time (Skinner and Staiger 2006). In other words, one needs to consider the convergence of physician-specific production functions $f_j(t)$, where t is time, toward the “true” production function f^* . This is similar to how diffusion is modeled in the sociological literature (Rogers 1995), where the emphasis is not on economic factors or profit maximizing but instead on a cultural or intrinsic inability to perceive the benefits of new technologies, despite evidence of their effectiveness. Parente and Prescott (1994) have focused on the importance of relatively minor differences in barriers to technology adoption in explaining differences in income growth across countries. This view emphasizes the view that cross-country differences in productivity arise from variation in technology adoption (Comin and Hobijn 2004).

There is certainly historical precedent for the inexplicably slow adoption of new medical innovations, and Donald Berwick (2003) has written a remarkable account of why this may be the case. In 1601, Captain James Lancaster administered lemon juice to the crew of just one of his fleet of four ships. Halfway through the trip, 40 percent of the sailors in the “control” ships had died of scurvy compared to no deaths in the “treatment” ship. Despite the strength of the evidence and affordability of the intervention, citrus fruits became required for British Navy sailors only in 1795, a mere 194 years after Lancaster’s trial. Perhaps even more striking is Joseph Lister’s recommendation that surgeons wash their hands, use gloves, and swab wounds with carbonic acid (Lister 1867). Lister acknowledged that his research was influenced by the Austrian physician Ignaz Semmelweis. Semmelweis demonstrated that maternal mortality from puerperal fever (an infection of the genital tract after giving birth) could be reduced from 12.2 percent to 2.4 percent by making physicians wash their hands with chlorinated lime between autopsy and obstetrical rotations. That Semmelweis did not “have a model” (which in medical parlance would correspond to a germ theory of disease to explain his results) is thought to have contributed to the medical establishment’s reluctance to embrace his findings.

It is even more difficult to explain in standard economic models the lethargy with which some pernicious treatments are dropped: Heidenrich and McClellan (2001) found that a substantial fraction of the improvement in survival for heart attacks could be explained by the gradual decline in the use of lidocaine, a treatment that actually tended to increase mortality. More recently, a randomized control trial found that providing information to “opinion leaders” in a hospital resulted in large increases in the use of appropriate medications following heart attacks and decreases in the use of outdated therapies (Soumerai et al. 1998). There may also be heterogeneity

in diffusion because of intrinsic differences in abilities of physicians (as in Jovanovich and Nyarko [1995, 1996]).

In summary, the adoption and use of health care innovations will be influenced by financial rewards, the perceived benefit from the innovation, and organizational structure. Low-value care of dubious clinical benefit may be adopted quickly if economic incentives encourage adoption. High-value treatment may not be adopted without financial incentives or if the efficacy of the treatment is difficult to observe. The sociological factors that moderate these channels may vary based on physician or hospital characteristics.

The degree to which different factors affect the level and diffusion of best and worst practices is an inherently empirical question. We turn next to the data that we will bring to bear in exploring the nature of variation in the utilization of high-quality and low-value care.

7.3 Data

Our analysis requires hospital-level measures of both high-quality and low-value care. Because we will be performing analysis at the hospital rather than patient level, it is important that we construct measures of each that are not contaminated by unobserved patient characteristics such as severity of illness—especially if those patient characteristics are likely to be correlated across hospitals within a geographic area.

To measure high-value health care, we use hospital compliance with a number of technical “process of care” measures. These measures use samples of patient discharge records for the treatment and capture interventions “for which there is strong scientific evidence and professional consensus that the process of care either directly improves outcomes or is a necessary step in a chain of care that does so,” such as the prescription of warfarin for atrial fibrillation or biennial eye examination for diabetics (U.S. Department of Health and Human Services, Centers for Medicare and Medicaid Services; Jha et al. 2005). The advantage of these measures is that detailed risk adjustment (based on how sick a hospital’s patient pool is) is not critical, as few patients are contraindicated for these procedures. The inability to perform high-quality risk adjustment at the hospital level is the principal reason that many other direct measures of quality, such as survival or the use of intensive interventions such as rescue angioplasty after a heart attack, are not utilized in this literature.

To construct a measure of the utilization of low-value care, we follow previous work that suggests that hospital expenditures on Medicare beneficiaries at the end of life can proxy for this high-intensity, low-benefit spending. These measures abstract from confounders such as unobserved illness by focusing on the set of patients who are terminally ill. Fisher et al. (2003a, 2003b) established that this end-of-life (EOL) spending is pervasive in areas that have a lot of beds, specialists, and health care facilities. They

have also demonstrated that higher EOL spending does not improve patient outcomes or satisfaction. Finally, this work has shown a correlation between a hospital's EOL spending and its treatment of acute conditions such as hip fracture and acute myocardial infarction (AMI), suggesting that hospitals treat many different patients with similar intensity. We next describe both measures in greater detail.

7.3.1 Measuring High-Value Care

The Hospital Quality Alliance, a public-private collaboration between the Centers for Medicare and Medicaid Services (CMS) and several hospital organizations, began reporting individual hospitals' performance on select process-of-care measures through an on-line Web site, "*Hospital Compare*," on April 1, 2005; See U.S. Department of Health and Human Services, CMS, and Jha et al. (2005) for details. These measures focus on three major conditions for which evidence-based treatments are supported by a solid body of evidence: acute myocardial infarction, pneumonia, and congestive heart failure (CHF). We analyzed data from 2004 to 2006, as 2004 is the first full year of reported data on these measures. We chose to retain only those measures for which a majority of hospitals reported at least twenty-five observations, a cutoff used in previous work to ensure sufficient statistical precision. Eleven process measures yielded at least twenty-five observations for a majority of hospitals: Aspirin at arrival and at discharge and beta-blocker prescription at arrival and at discharge (for AMI patients); assessment of left ventricular function, the provision of discharge instructions, and angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) prescription for patients with left ventricular systolic dysfunction (LVSD; for CHF patients); blood culture performed before receiving the first antibiotic in the hospital, first dose of antibiotics within four hours of admission, initial antibiotic selected appropriately, and assessment of arterial oxygenation within twenty-four hours of arrival (for pneumonia patients). We computed quality scores for 6,917 hospital-year observations.

To create condition-specific quality scores for AMI, pneumonia, and CHF, we used a common methodology prescribed by the Joint Commission, which suggests a summary score calculated as the sum of the number of times a hospital performed the appropriate action across all measures for that condition (numerator), divided by the number of "opportunities" the hospital had to provide appropriate care (denominator). Composite scores were only calculated if a hospital had at least twenty-five patients for at least one measure in that condition. We also calculated an overall quality measure that used information for all three clinical conditions by taking the mean of summary scores across conditions for each hospital. (As a specification test, we also used factor analysis to combine the three composite measures for each hospital, but the correlation between the factor index and the simple

average was 0.98, so we used the average.) We thus have four measures of the quality of care: the condition-specific scores for AMI, pneumonia, and CHF patients, and the composite score that aggregates these three. The correlations between the condition-specific scores and the composite-quality score were 0.62 for heart attacks (AMI), 0.82 for congestive heart failure (CHF), and 0.77 for pneumonia. We focus primarily on the composite measure in our analysis, but results using the component measure are similar and are summarized next.

7.3.2 Measuring Low-Value Spending

Our measure of the use of high-intensity but low-value care is based on the intensity of inpatient care for Medicare beneficiaries in the six months preceding death (EOL). By focusing on variation in the treatment of patients with identical life expectancies, the EOL spending measure better reflects the portion of spending that is attributable to differences in practice patterns as opposed to differences in severity of illnesses. End-of-life expenditures have been shown to be highly correlated with both total Medicare spending and spending for specific disease cohorts. Another advantage of focusing on EOL spending is that the sample size is large enough to calculate measures at the hospital level.

To construct our measure of EOL spending, we used Medicare part A (hospital) and part B (outpatient) spending and utilization data for hospital and physician services for chronically ill Medicare beneficiaries who died during the years 1999 to 2003. At death, each of these patients was assigned to the hospital in which he or she had received the majority of care in the previous six months. All of the expenditure and utilization data from that patient's claims were then assigned to that hospital. The vast majority of patients' care occurred at the assigned hospital; the average percent of inpatient days spent at the assigned hospital was 89.6 percent. Spending data were adjusted for differences in age, sex, race, and the relative frequency of chronic illness among the beneficiaries studied.

7.3.3 Hospital and Area Characteristics

We linked the hospital-level quality and EOL data to the American Hospital Association Annual Survey Database that has information on each hospital's nurse-to-patient (Census) ratio, profit status (public, not for profit, for profit), membership in the Council of Teaching Hospitals, bed size, location (region, county, health referral region [described next], and urban versus rural), percentage of Medicare and Medicaid patients, and presence of a medical intensive care unit. Nurse/Census ratios were calculated by dividing the number of full-time equivalent nurses on staff by 1,000 patient days.

We used each hospital's county to link these data to area-level population characteristics from the U.S. Department of Health and Human Services, Health Resources and Services Administration Area Resource File. This

data set contains information on residents' characteristics, such as education, race, income, and age. It also contains information on county-level health care providers, such as the number of physicians broken down by specialty practicing in the county.

7.3.4 Defining Geographic Areas

To construct local markets for health care, we adopt the methodology of the "Dartmouth Atlas of Health Care," which divides the United States into 306 hospital referral regions (HRRs) based on attributing zip codes to HRRs by the use of an algorithm reflecting commuting patterns and the location of major referral hospitals. The HRRs may cross state and county borders, because they are determined solely by migration patterns of patients. For example, the Evansville, Indiana HRR encompasses parts of three states, because it draws patients so heavily from Illinois and Kentucky. The HRRs are best viewed as the level at which tertiary services such as cardiac surgeries are received (although they are not necessarily the appropriate geographical level for primary care services).

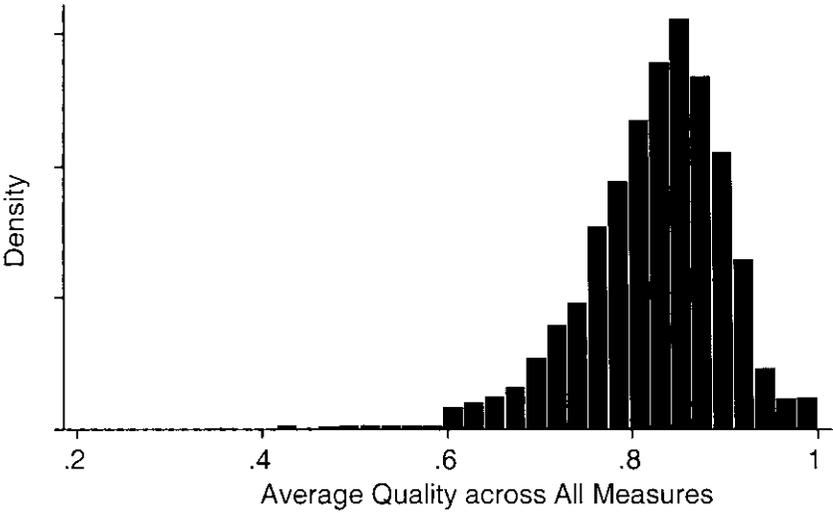
Analysis at the HRR level is preferable to analysis at the city or state level, since it uses the empirical pattern of patient commuting to determine the geographic boundaries of each referral region rather than assuming that the arbitrary political boundaries of states and cities also define the level at which the health care is delivered. Furthermore, for the purpose of studying geographic productivity spillovers, an analysis at the HRR level is superior to one at the level of the individual hospital for two reasons. First, patients can be assigned to an HRR based on their residence rather than on the hospital at which they received treatment (which may be endogenous). In addition, productivity spillovers are likely to operate at a broader level than that of a given hospital; for example, these spillovers are expected to reach beyond the boundary of the firm to affect productivity at all firms in a region. Physicians often have operating privileges in multiple hospitals and interact (socially and professionally) with other doctors who may or may not practice in their hospital, and patients are commonly referred to other hospitals within the HRR for treatment. The HRR encompasses a physician's likely peers and a hospital's natural competitors.

7.4 Agglomerations in High-Value and Low-Value Care

In the two panels of figure 7.1, we illustrate the degree of variation at the hospital level in high-value, high-quality care (panel A) and in the intensity of end-of-life, low-value spending (panel B). In panel A, the quality measure used is the composite measure that puts equal weight on treatments for heart attacks, pneumonia, and congestive heart failure. In principal, this rate should ideally be close to 1 (100 percent) for all hospitals, as very few patients are contraindicated for the measures that comprise the quality index. Panel

A

Distribution of Hospital Quality Average Hospital Score across All Measures



B

Distribution of EOL Spending Average Hospital Spending in Medicare Patients' Last 6 Months of Life

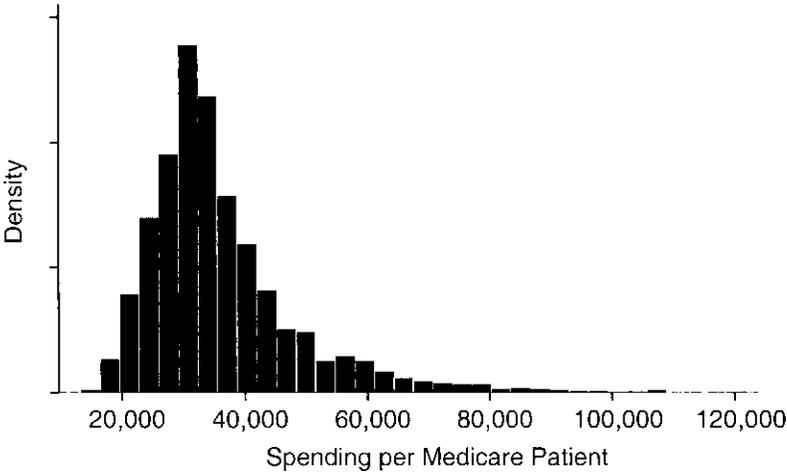


Fig. 7.1 Distribution of quality (panel A) and end-of-life spending (panel B) for Medicare patients in U.S. hospitals

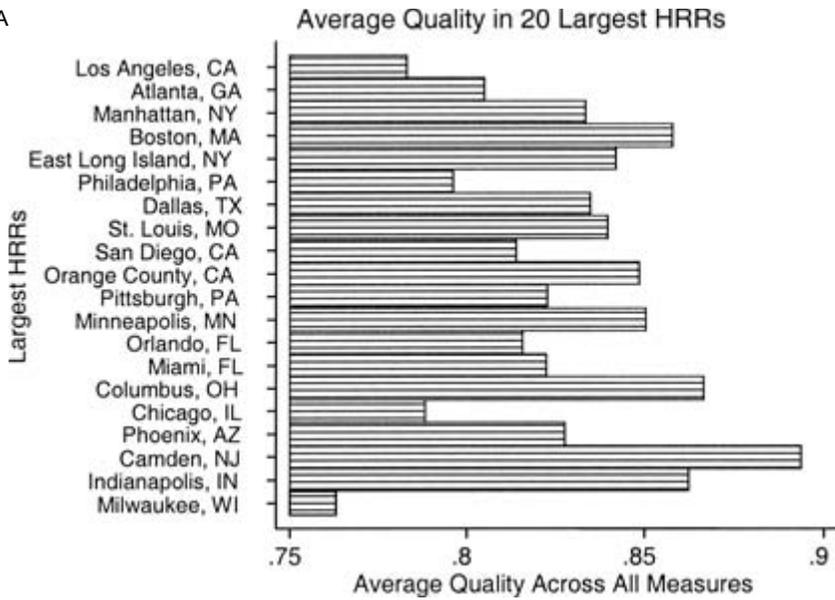
Note: Quality measures refer to a composite measure of quality that puts equal weight on treatments for heart attacks, pneumonia, and congestive heart failure. Spending data is deflated to 2006 dollars.

B shows the roughly log-normal distribution of the use of technologically intensive, low-value EOL care across hospitals.

Figure 7.1 illustrates the wide variability with which hospitals in the United States utilize effective and ineffective care. This variation in hospital utilization patterns is not constrained to particular small areas nor uniformly spread throughout the country: hospitals in some cities perform systematically better than those in others. Figure 7.2 illustrates variation in the use of both high-quality and low-value care across the twenty largest HRRs in the United States. These HRRs comprise the largest cities in America—each has a number of teaching hospitals, and while it is possible that individual hospitals may differ in the severity of their patient caseloads or the managerial expertise of their administrators, it is not obvious why Philadelphia should *ex ante* provide a lower level of care than Chicago or Pittsburgh. As panel A of this figure illustrates, the hospitals comprising the Philadelphia HRR score poorly on performance on standard quality measures, while their northern neighbors in Camden, New Jersey, do substantially better. Similarly, it is not entirely clear why the hospitals in the Los Angeles HRR, which serve high-income patients and include places such as the University of California, Los Angeles, and Cedars Sinai, should perform so poorly relative to those next door in the Orange county HRR. In panel B of figure 7.2, we graph the analogous use of EOL spending across the twenty largest HRRs. There is similarly wide variation in the use of this form of technologically intensive medicine. Once again, there is no *a priori* reason to believe that some HRRs should be using these technologies more than others. Furthermore, casual inspection of the two panels does not support the conclusion that the use of high- and low-value care is correlated (either negatively or positively). This casual observation is confirmed by the estimated (statistically insignificant) correlation coefficient between the two measures across all HRRs of -0.06 .

Table 7.1 summarizes the key insights of figures 7.1 and 7.2. At the hospital level, we report the variation in hospital performance on the quality and EOL spending measures. Of particular interest are the last two columns of this table, which present analysis of variance decompositions of the extent to which the hospital-level variation in figures 7.1 and 7.2 is accounted for by HRR factors. For the quality measures, between 17 and 20 percent of the hospital-level variation is driven by HRR factors, leaving a substantial amount of within variation. With EOL spending, the opposite is true: here, there is substantial between HRR variation in use, with HRR effects accounting for over 70 percent of the hospital variation. This is another way of saying that when it comes to EOL spending, hospitals in the same HRR practice a similar type of medicine, resulting in agglomerations of regions where intensive medicine is practiced. But when it comes to quality, there is much more disagreement between providers, even those who are within the same HRR. In the following analysis, we will estimate regression models to

A



A

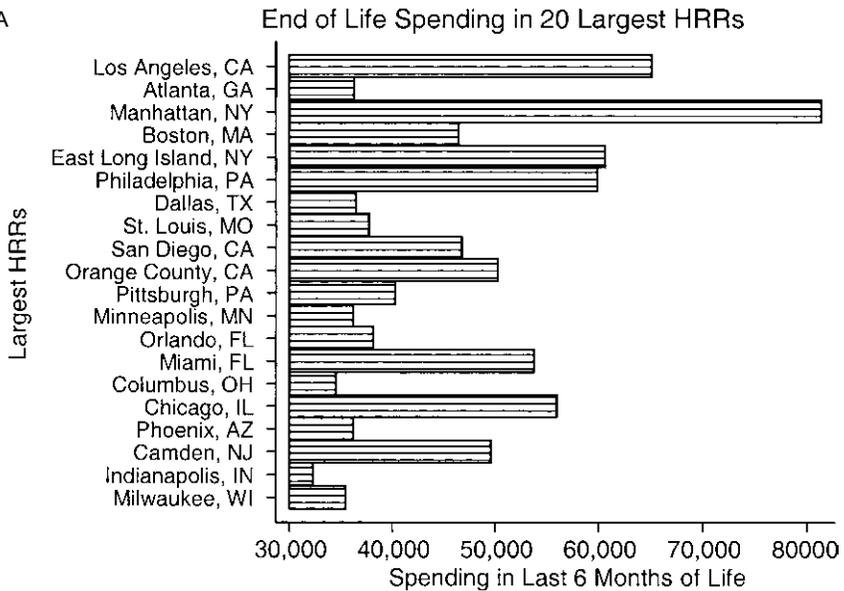


Fig. 7.2 Distribution of quality (panel A) and end-of-life spending (panel B) for Medicare patients in the largest twenty HRRs

Note: Quality measures refer to a composite measure of quality that puts equal weight on treatments for heart attacks, pneumonia, and congestive heart failure. Spending data is deflated to 2006 dollars.

Table 7.1 Hospital- and HRR-level variation in quality and end-of-life spending

	Summary statistics			Variance decomposition	
	Average quality	Standard deviation	Range (5% to 95%)	Variation between HRRs (%)	Variation within HRRs (%)
High-value care					
Heart attack	92.50%	6.60%	80.0–99.3%	17	83
Heart failure	75.30%	16.10%	44.4–95.5%	17	83
Pneumonia	82.10%	8.00%	68.3–94.0%	21	79
Composite score	82.40%	8.90%	66.5–94.3%	19	81
Low-value care					
End-of-life spending	\$36k	\$13k	\$20.6–61.0k	72	28

Notes: Composite (quality) score refers to a composite measure of quality that puts equal weight on treatments for heart attacks, pneumonia, and congestive heart failure. Spending data is deflated to 2006 dollars. The analysis of variance decompositions in the last two columns are weighted by the number of hospital beds.

better understand the drivers of between versus within variation. Means and standard deviations for the variables that we use as regressors are shown in table 7.2. While not reported, we also performed the decompositions in this table without weighting the analysis by the number of beds (thereby allowing smaller hospitals to exert as much influence as larger ones in the estimation of HRR effects). The results from this unweighted analysis are almost identical to those reported in table 7.1, suggesting that the practice style of a particular hospital is orthogonal to its size.

These patterns motivate us to examine the correlates of high- and low-value care across regions (an HRR-level analysis, in table 7.3) and within regions (a hospital-level analysis with HRR fixed effects, in table 7.4). While these regression models should not be interpreted causally, they are a useful framework within which to examine correlations that serve to motivate more formal theories of technology adoption and diffusion. To better understand why some HRRs provide higher-quality care, we use HRR-level data and regress each HRR's quality score on a number of covariates. The first column of table 7.3 describes results for the use of our composite quality measure (a metric of low-cost, high-value care) and the second column for the use of EOL spending (a metric of high-cost, low-value care). Regression coefficients and robust standard errors are presented. The covariates that we have used (including nine indicator variables for the Census regions) explain only 30 percent of the between-HRR variation in quality but over 70 percent of the between-HRR variation in EOL spending. Because we have included Census region fixed effects, our results are not driven by explanations that are unique to states in New England or the South, for example. Holding the composition of the physician workforce (the mix of specialists and generalists) constant, adding one more physician to an HRR raises quality but also

Table 7.2 Summary statistics

	HRR-level data		Hospital-level data	
	Mean	Standard deviation	Mean	Standard deviation
Percent age > 65	0.13	0.03	0.15	0.06
Percent < high school	0.19	0.06	0.21	0.11
In largest 100 cities	0.16	0.27	0.18	0.38
Publicly owned	0.16	0.19	0.17	0.38
Teaching hospital	0.14	0.19	0.09	0.28
Log(beds)	7.48	0.84	5.02	0.92
RNs per capita	1.25	0.27	1.21	0.66
Percent black	0.09	0.10	0.10	0.15
Percent Medicaid	0.20	0.07	0.18	0.14
Per capita MDs	0.00	0.00	0.00	0.00
Percent specialists	0.57	0.14	0.58	0.14
Overall quality	0.84	0.04	0.82	0.09
EOL spending	36.19	9.02	36.10	12.93

Notes: HRR refers to hospital referral regions as defined by the “Dartmouth Atlas of Health Care.” Covariates are from the American Hospital Association and the Area Resource File. Quality of care is from the Hospital Quality Association. End-of-life spending is from Medicare claims data. Data sources are described in more detail in the text.

EOL spending. Interestingly, if we were to hold the size of the workforce constant and simply replace a generalist with a specialist, there would be no change in quality but a substantial increase in the use of EOL spending (consistent with some of our previous research). Also noteworthy is the fact that HRRs with a greater share of teaching hospitals perform substantially better on quality and are no more profligate in their use of EOL care. Large cities are not more likely to offer one type of care vis-à-vis smaller cities. But HRRs with larger shares of minority patients, less-educated patients, and low-income patients (as measured by the percent receiving Medicaid) are all less likely to offer high-quality care and slightly more likely to spend on EOL care. These findings caution against the explanation that richer places spend more on EOL care simply because the marginal utility of health care in the end of life, however small, exceeds the marginal utility of other consumption (as in Hall and Jones [2007]).

In addition to the variation between HRRs in quality and the use of low-value care, the variance decompositions in table 7.1 also hinted at substantial variation within areas, especially in the use of high-quality care. To examine the correlates of this variation, in table 7.4, we estimate models with data at the hospital level that include HRR fixed effects, isolating within-HRR variation in quality and spending. Regression coefficients and standard errors clustered at the HRR level are presented. In these regressions, it is not possible to control for the specialist composition of a hospital’s

Table 7.3 Association between composite quality, EOL spending, and HRR characteristics

	HRR-level regressions	
	Overall quality	Low-value EOL spending
Per capita total active nonfederal MDs	5.96 (2.90)	4,779.78 (862.85)
Share of MDs who are specialists	-0.04 (0.03)	12.05 (4.61)
HRR in one of 100 largest cities	-0.01 (0.01)	1.67 (1.63)
Share of hospitals that are public (nonfederal)	-0.01 (0.01)	-0.38 (1.65)
Share of hospitals that are teaching hospitals	0.03 (0.01)	3.85 (2.69)
Total beds of hospitals in HRR (log)	-0.01 (0.00)	1.40 (0.55)
Full-time employed RNs per patient Census day	0.00 (0.01)	2.78 (1.34)
Percent of patients who are black	-0.04 (0.03)	18.36 (4.80)
Share of inpatient days by Medicaid patients	-0.01 (0.05)	-5.59 (5.25)
Percent of population over age 65	0.00 (0.06)	-2.84 (9.81)
Percent of population with less than high school	-0.13 (0.05)	30.19 (11.25)
Region dummies	Yes	Yes
R^2	0.29	0.72
Observations	262	262

Notes: Regression coefficients from between-HRR regressions are shown, along with robust standard errors; HRR refers to hospital referral regions as defined by the “Dartmouth Atlas of Health Care.” Covariates are from the American Hospital Association and the Area Resource File. Quality of care is from the Hospital Quality Association. Composite quality is based on treatment for heart attack, heart failure, and pneumonia patients. End-of-life spending is from Medicare claims data. Data sources are described in more detail in the text.

area, since that is a characteristic of the area that is absorbed by the HRR fixed effect. Table 7.4 shows that hospitals with a larger share of minority, poor, and less-educated patients offer lower-quality care. Once again, it is important to note that our measures of quality do not require the presence of costly investments in new technology: these are primarily low-cost, high-value interventions. On the other hand, EOL spending, which is expensive and technologically intensive, is more likely to be offered at hospitals that serve minority and less-educated populations. It is surprising that teaching hospitals do not offer higher-quality care—some of this effect is simply a consequence of their urban location, a large component of which we have

Table 7.4 Association between composite quality, EOL spending, and hospital characteristics

	Within-HRR regressions	
	Overall quality	Low-value EOL spending
In one of 100 largest cities	0.01 (0.00)	3.63 (0.69)
Public nonfederal hospital dummy	-0.01 (0.01)	-1.17 (0.42)
Teaching hospital dummy	0.00 (0.00)	2.99 (0.76)
Log(hospital beds)	0.02 (0.00)	3.61 (0.20)
Full-time employed RNs per patient Census day	0.01 (0.00)	0.77 (0.43)
Percent of patients who are black	-0.09 (0.02)	8.96 (2.12)
Share of inpatient days by Medicaid patients	-0.02 (0.01)	-4.68 (1.29)
Percent of zip over age 65	0.00 (0.03)	-4.57 (3.74)
Percent of zip with less than high school	-0.05 (0.02)	4.68 (1.98)
Region dummies	Yes	Yes
R^2	0.247	0.759
Observations	6,917	9,751

Notes: Regression coefficients from within-HRR regressions are shown, along with standard errors clustered at the HRR level; HRR refers to hospital referral regions as defined by the “Dartmouth Atlas of Health Care.” Covariates are from the American Hospital Association and the Area Resource File. Quality of care is from the Hospital Quality Association. Composite quality is based on treatment for heart attack, heart failure, and pneumonia patients. End-of-life spending is from Medicare claims data. Data sources are described in more detail in the text.

adjusted for by controlling with an indicator variable for hospitals located in one of the largest one-hundred cities. This fact is consistent with the work of Fisher et al. (2004), who note substantial variations in the efficiency of academic medical centers. And contrary to many assertions in the literature, a greater presence of nurses does not seem to result in higher-quality care being provided.

At the bottom of table 7.4, we note that these covariates explain only 25 percent of the within-HRR variation in quality but over 75 percent of the within-HRR variation in EOL spending. So, while the provision of quality, even of the low-cost, high-value kind, may be a function of physician beliefs or idiosyncratic factors at the hospital level, the same explanation does not apply to the phenomena of using low-value, high-cost care. Here,

a basic set of hospital and HRR characteristics accounts for a large share of the variation.

7.5 The Role of Learning

These facts point to a world where the use of high-quality care is idiosyncratically dispersed through the health care sector in ways that are uncorrelated with geographic region. Yet, geography at the level of the hospital market is a powerful descriptor of the use of low-value care. But this (within-region) variation in the use of high-quality care also suggests an interesting hypothesis: if high-quality hospitals are located right next to low-quality hospitals, do they learn from their high-quality neighbors? This question is extremely difficult to resolve empirically, as demonstrated in the substantial “peer-effects” literature (Sacerdote 2001). Our goal is less ambitious—we simply ask if the evidence is consistent with learning, and we rely on the use of simple ordinary least squares (OLS) regressions to explore the case for learning. We exploit the panel structure of our data to ask if investments in quality made by a hospital’s neighbors last year predict that hospital’s performance this year. The obvious concern with this approach is that investments in quality made by neighboring hospitals are correlated but do not causally affect each other’s performance. To address this concern, we condition on a hospital’s own performance last year. In other words, we compare two hospitals with similar quality in a given year and ask if the one with higher-quality neighbors improves by more than the one with lower-quality neighbors. We perform the parallel analysis for the use of EOL spending. The fact that quality and EOL spending are measured on separate scales makes it difficult to compare the regression coefficients on neighbors’ performance across these two different outcomes. We address this concern by standardizing both variables (that is, we take each variable, subtract its mean, and divide by its standard deviation). These standardized variables are both mean 0 and variance 1 (z -scores).

We construct each hospital’s neighbors based on a composite of all other hospitals in the HRR. That is, neighbor quality for each hospital is just the average value of that variable for all of the other hospitals in the same HRR. Results using alternate specifications (such as choosing only the largest nearby hospital as the neighbor or focusing on the best performing hospital as the critical neighbor) produce very similar results.

Our exploration of the possibility of learning from neighbors is shown in table 7.5 (high-quality care) and table 7.6 (low-value end-of-life spending). We present regression coefficients and standard errors clustered at the HRR level. In both tables, the first four columns use the dependent variable Y_{it} (composite quality or EOL spending) directly, while the last four columns substitute the standardized (z -score) version of the dependent variable (to present results that are more easily compared across variables). The first

Table 7.5 Hospital-level association between quality and neighbors' quality

	Overall quality				Overall quality: Standardized			
	Base (1)	Add neighbors (2)	Add lagged dependent variable (3)	Instrumental variable for neighbors (4)	Base (5)	Add neighbors (6)	Add lagged dependent variable (7)	Instrumental variable for neighbors (8)
In one of 100 largest cities	0.001 (0.005)	0.003 (0.004)	-0.001 (0.002)	-0.002 (0.002)	0.012 (0.040)	0.025 (0.036)	-0.002 (0.023)	-0.001 (0.023)
Public nonfederal hospital dummy	-0.006 (0.005)	-0.011 (0.005)	-0.008 (0.003)	-0.008 (0.003)	-0.079 (0.046)	-0.084 (0.047)	-0.033 (0.024)	-0.033 (0.024)
Teaching hospital dummy	0.005 (0.005)	0.000 (0.005)	-0.004 (0.003)	-0.004 (0.003)	0.041 (0.040)	0.013 (0.039)	-0.010 (0.021)	-0.010 (0.021)
Ln(total hospital beds)	0.019 (0.002)	0.022 (0.002)	0.009 (0.001)	0.009 (0.001)	0.144 (0.020)	0.158 (0.021)	0.074 (0.011)	0.074 (0.011)
Full-time employed RNs per patient	0.013 (0.004)	0.012 (0.004)	0.004 (0.002)	0.004 (0.002)	0.119 (0.032)	0.106 (0.034)	0.031 (0.018)	0.031 (0.018)
Census day	-0.094 (0.016)	-0.075 (0.016)	-0.026 (0.009)	-0.029 (0.010)	-0.527 (0.131)	-0.429 (0.129)	-0.191 (0.076)	-0.190 (0.076)
Percent black patients	-0.018 (0.013)	-0.022 (0.012)	-0.003 (0.007)	-0.003 (0.007)	-0.215 (0.118)	-0.273 (0.121)	-0.116 (0.064)	-0.116 (0.064)
Percent of inpatient days by Medicaid patients	6.679 (2.940)	2.978 (1.592)	0.933 (0.999)	1.766 (1.172)	58.159 (24.430)	33.566 (16.149)	17.732 (9.404)	17.464 (10.485)
Per capita total active nonfederal MDs	-0.086 (0.028)	-0.065 (0.015)	-0.021 (0.011)	-0.028 (0.011)	-0.707 (0.233)	-0.717 (0.159)	-0.386 (0.101)	-0.385 (0.101)
Percent of MDs who are specialists								

(continued)

Table 7.5 (continued)

	Overall quality				Overall quality: Standardized			
	Base (1)	Add neighbors (2)	Add lagged dependent variable (3)	Instrumental variable for neighbors (4)	Base (5)	Add neighbors (6)	Add lagged dependent variable (7)	Instrumental variable for neighbors (8)
Percent of zip over age 65	-0.030 (0.025)	-0.018 (0.026)	0.002 (0.017)	0.001 (0.018)	-0.465 (0.220)	-0.401 (0.235)	-0.112 (0.154)	-0.112 (0.153)
Percent of zip with less than high school	-0.075 (0.017)	-0.058 (0.016)	-0.017 (0.009)	-0.020 (0.009)	-0.764 (0.141)	-0.632 (0.142)	-0.198 (0.082)	-0.197 (0.082)
Quality in neighboring hospitals		0.556 (0.045)	0.213 (0.031)	0.071 (0.063)	0.462 (0.060)	0.462 (0.060)	0.162 (0.032)	0.167 (0.071)
Lagged dependent variable			0.680 (0.019)	0.696 (0.020)			0.659 (0.033)	0.659 (0.033)
R^2	0.140	0.227	0.597	0.594	0.148	0.194	0.548	0.547
N	6,917	4,353	4,353	4,353	6,917	4,353	4,353	4,353

Notes: Regression coefficients and standard errors clustered at the HRR level are shown; HRR refers to hospital referral region as defined by the “Dartmouth Atlas of Health Care.” Covariates are from the American Hospital Association and the Area Resource File. Quality of care is from the Hospital Quality Association. Composite quality is based on treatment for heart attack, heart failure, and pneumonia patients. Standardized quality is expressed as z-scores. Data sources are described in more detail in the text.

Table 7.6 Hospital-level association between EOL spending and neighbors EOL spending

	End-of-life low-value care				End-of-life low-value care: Standardized			
	Base (1)	Add neighbors (2)	Add lagged dependent variable (3)	Instrumental variable for neighbors (4)	Base (5)	Add neighbors (6)	Add lagged dependent variable (7)	Instrumental variable for neighbors (8)
In one of 100 largest cities	4.009 (0.868)	2.970 (0.601)	0.721 (0.199)	0.722 (0.199)	0.303 (0.066)	0.220 (0.044)	0.051 (0.014)	0.051 (0.014)
Public nonfederal hospital dummy	-1.119 (0.415)	-1.320 (0.391)	-0.475 (0.134)	-0.465 (0.133)	-0.084 (0.031)	-0.099 (0.029)	-0.035 (0.010)	-0.034 (0.010)
Teaching hospital dummy	3.272 (0.869)	2.656 (0.784)	-0.027 (0.273)	-0.047 (0.267)	0.249 (0.066)	0.198 (0.058)	-0.005 (0.021)	-0.007 (0.020)
Ln(total hospital beds)	4.242 (0.270)	3.968 (0.196)	1.057 (0.091)	1.035 (0.090)	0.321 (0.020)	0.295 (0.015)	0.076 (0.007)	0.074 (0.007)
Full-time employed RNs per patient	0.895 (0.550)	1.132 (0.445)	0.304 (0.141)	0.292 (0.140)	0.068 (0.042)	0.085 (0.033)	0.023 (0.011)	0.021 (0.010)
Census day	10.602 (2.581)	6.663 (1.886)	1.744 (0.622)	1.781 (0.623)	0.803 (0.195)	0.494 (0.140)	0.123 (0.046)	0.127 (0.046)
Percent black patients	-7.454 (1.477)	-5.110 (1.265)	-1.244 (0.462)	-1.256 (0.458)	-0.563 (0.112)	-0.379 (0.094)	-0.087 (0.035)	-0.089 (0.034)
Percent of inpatient days by Medicaid patients	4,900.000 (879,786)	787.954 (272,110)	288.773 (119,069)	370.209 (141,113)	369.956 (66,490)	57.030 (20,382)	17.144 (9,014)	26.385 (10,614)
Per capita total active nonfederal MDs	19.641 (5.470)	0.869 (1.865)	0.828 (0.872)	1.226 (0.962)	1.483 (0.414)	0.056 (0.138)	0.043 (0.063)	0.088 (0.072)
Percent of MDs who are specialists								

(continued)

Table 7.6 (continued)

	End-of-life low-value care				End-of-life low-value care: Standardized			
	Base (1)	Add neighbors (2)	Add lagged dependent variable (3)	Instrumental variable for neighbors (4)	Base (5)	Add neighbors (6)	Add lagged dependent variable (7)	Instrumental variable for neighbors (8)
Percent of zip over age 65	-6.829 (3.684)	-4.690 (3.123)	-1.906 (0.922)	-1.928 (0.906)	-0.517 (0.278)	-0.348 (0.232)	-0.138 (0.068)	-0.141 (0.066)
Percent of zip with less than high school	11.872 (3.259)	7.882 (2.014)	3.009 (0.750)	3.061 (0.743)	0.895 (0.247)	0.586 (0.149)	0.219 (0.054)	0.225 (0.054)
EOL spending in neighboring hospitals		0.711 (0.026)	0.170 (0.014)	0.151 (0.020)		0.700 (0.026)	0.172 (0.014)	0.142 (0.019)
Lagged dependent variable			0.762 (0.017)	0.770 (0.017)			0.753 (0.017)	0.764 (0.017)
R^2	0.624	0.718	0.875	0.875	0.628	0.719	0.877	0.877
N	9,751	7,312	7,312	7,312	9,751	7,312	7,312	7,312

Notes: Regression coefficients from within-HRR regressions are shown, along with standard errors clustered at the HRR level; HRR refers to hospital referral regions as defined by the “Dartmouth Atlas of Health Care.” Covariates are from the American Hospital Association and the Area Resource File. End-of-life spending is from Medicare claims data. Standardized spending is expressed as z -scores. Data sources are described in more detail in the text.

column does not include neighbors' performance, constraining the effect to be zero. The second column adds in the performance of neighbors, and the third column adds the lagged dependent variable for a hospital's own performance in the previous year. With X_i denoting the characteristics of hospital i , N_{-it} the performance of hospital i 's neighbors in year t , and $Y_{i(t-1)}$ the lagged performance of hospital i , this last specification is formally expressed as:

$$(1) \quad Y_{it} = \alpha + X_i\beta + \delta N_{-it} + \lambda Y_{i(t-1)} + e_{it}.$$

This specification may understate the effect of neighbors on performance, for it effectively rules out any lagged neighbor effect on current outcomes. Such an effect is absorbed by the coefficient λ , which through back substitution can be shown to capture the role of lagged neighbor performance on contemporaneous outcomes. The identifying assumption here is that $E(e_{it} | X_i, N_{-it}, Y_{i(t-1)}) = 0$ or that unobserved shocks to hospital performance aren't predictable in controlling for hospital characteristics and lagged performance (which includes lagged neighbors' performance). In other words, we are assuming that there is not a common process that contemporaneously improves a hospital's performance and that of its neighbors. This would be violated if, for example, a local quality improvement effort was implemented and affected a hospital and its neighbors or if hospitals and their neighbors were learning about a new technology from a common source. To begin to explore this possibility, in the fourth column, we instrument for neighbors' performance using the characteristics of the neighbors that are described in table 7.3 (such as location in one of the largest one-hundred cities, number of beds, specialist composition in the area, and patient characteristics). In this specification, we are identifying the effect of neighbors' performance on a hospital's performance by only utilizing variation in neighbors' performance that can be attributed to those neighbors' attributes. This source of identification is valid only insofar as those attributes are not correlated with unobserved characteristics that affect both neighbors' performance and the changes in a hospital's own performance. In other words, we need the fact that a greater share of a hospital's neighbors are teaching hospitals (for example) to be uncorrelated with the *unobservable* determinants of that hospital's quality—having controlled for that hospital's own teaching status and its quality performance last year. This assumption is untestable, but it becomes more difficult to tell a story about the relationship being driven by unobservable common factors. Unfortunately, it is difficult to isolate a source of variation that is not subject to similar concerns.

We present results both for the simple composite quality measure and for the standardized measures. It is easiest to compare coefficient estimates across these specifications in the panels on the right, since the variables of interest are measured in similar units. Examining column (6) of table 7.5, we see that a 1 standard deviation increase in neighbors' quality raises a given hospital's quality score by 0.46 standard deviations. As noted ear-

lier, this effect probably overstates the effect of neighbors on learning. In column (7), where we control for own performance last year, the effect of a hospital's neighbors having quality that is 1 standard deviation higher quality than average (a z -score of 1, with mean equal to 0 by construction) is an increase in quality of 0.16 standard deviations (our preferred specification). The instrumental variable (IV) results in column (8) cannot reject the OLS estimates. This is not the consequence of weak instruments: the first stage of our two-stage least squares regression has an F -statistic of 45.

The analogous results for EOL spending are reported in table 7.6. Here, having neighbors whose EOL spending is 1 standard deviation higher raises a hospital's EOL spending by 0.70 (column [6]) standard deviations when we do not include a lagged dependent variable and 0.17 standard deviations (column [7])—our preferred specification) when we do. These are large spillover effects. In this case, the IV results are virtually identical to our preferred OLS specification—column (8) suggests that having neighbors whose EOL spending is 1 standard deviation higher raises a hospital's own EOL spending by 0.14 standard deviations.

The magnitude of learning is thus quite similar across the two outcomes. This might seem at odds with the fact that there is much more variation within HRRs in quality than in EOL care—a phenomenon that may cause one to think that there is greater within-earning with EOL spending. But it is possible for learning to occur for quality and EOL spending at similar rates even while there is less within-area variation in EOL spending if there is faster innovation in low-value EOL care. This seems consistent with highly intensive care being based on adoption of the latest rapidly developing technology, while most of the quality measures used here have been in existence for much longer.

We performed a number of specification tests to rule out alternative explanations. First, we also estimated models that included the lead (instead of or in addition to lag) of neighbors' performance. Our concern was that if quality is noisily measured from year to year, it may be possible to see a correlation between neighbors' scores last year and performance today. But a pure measurement error story would also imply a correlation between performance today and neighbors' quality a year from today (a model that by construction cannot be causal). In general, we found that the largest effects of neighbors' performance loaded onto lagged performance (results available on request). This specification check suggests that the associations identified in these tables could be causal. As previously discussed, we also tested several other methods of constructing a hospital's neighbors, which produced quite similar results. We also estimated a number of specifications with interaction effects to explore whether certain hospitals (such as teaching hospitals or publicly owned hospitals) had stronger spillover effects on neighbors, but we were not able to estimate these second-order effects with

enough precision. Last, instead of our composite measure of quality, we reestimated the model with the subcomponent measures (such as the treatment of heart attack patients only), again with quite similar results. For example, the increase in standardized quality of treatment for heart attacks (using the IV specification of column [8]) associated with a 1 point increase in neighbors' quality is 0.19 (HRR-clustered standard error 0.09) and for treatment of pneumonia is 0.11 (HRR-clustered standard error 0.07). While it is clearly still possible that our findings are driven by alternative explanations, these results are quite suggestive of agglomeration effects.

7.6 Discussion

The measures that we have used are subject to several limitations. The first is an aspect of the quality measures themselves, since quality measures may in theory penalize hospitals that treat sicker patients. While this concern may be true of quality measures for acute treatments (such as the use of angioplasty after heart attacks), the measures used in this chapter were chosen in part because they are not sensitive to the ability to perform detailed risk adjustment. Moreover, Higashi et al. (2007) have noted that patients with more comorbidities are no less likely to receive higher-quality care of the type used in our quality measures.

Second, we examined process-of-care measures (such as use of beta-blockers) rather than outcome measures (such as mortality). Policymakers have focused on process-of-care measures because they are less sensitive to risk adjustment, but these measures recently have come under fire, with some critics claiming that they do not actually predict health outcomes such as mortality after AMI and CHF (see, for example, the work of Bradley et al. [2006] and Werner and Bradlow [2006]. More recent work by Jha et al. (2007), however, has demonstrated that there is indeed a significant positive relationship between performance on these measures and risk-adjusted survival rates for each of the three conditions measured here and that hospitals with higher performance on these measures achieve better outcomes for their patients.

A third concern is the validity of the end-of-life measure: a very sick patient treated in a high-intensity hospital may have an increased chance of survival and thus not end up in the end-of-life sample. Presumably, this individual experienced above-average expenditures, thus excluding him or her from the sample and thus attenuating the measured differences in spending. Bach, Schrag, and Begg (2004) have also noted that regions with more "low-cost" diseases will appear to experience lower expenditures in the end-of-life cohorts. However, our spending data have already been adjusted for the relative frequency of diseases in each hospital's patient population, mitigating these concerns. Additionally, previous work has shown a correlation

between a hospital's EOL spending and its treatment of acute conditions such as hip fracture and AMI, suggesting that these hospitals treat many different patients with similar intensity.

One might interpret these findings in the context of Baicker and Chandra (2004a, 2004b) and Chandra and Staiger (2007). In this work, we argued that there are two types of health care—one technologically intensive and provided by specialists and the other less intensive. Specialization in one type of care crowds out the ability to deliver the other type of care. This trade-off is best illustrated in the context of competing treatments for the same condition—angioplasty (intensive) versus thrombolytics (less intensive) for heart attacks. It is, however, difficult to view EOL spending as being an alternative therapy for anything; few physicians view it as being productive in almost any setting.

Last, while we have referred to the spillover effects between neighboring hospitals as “learning,” we have not demonstrated that this is the best characterization of the spillover. Competition for local patients could spur hospitals to improve their quality more quickly once their neighbors have done so, for example. This shorthand masks a number of different underlying channels for spillovers that will only be distinguishable with further research.

7.7 Conclusion

Analysis of the health care sector provides a valuable window into the causes of variation in productivity across areas and the role of agglomeration in generating innovation and efficiency. This investigation into variation in the use of high-value, low-cost health care and high-cost, low-value health care has yielded a number of surprising facts. First, there is large variation in the use of both innovations but with different patterns across areas. Variation in the use of high-quality care is not restricted to certain markets—there is variation in care within Boston, Massachusetts, as well as within Birmingham, Alabama—but there is much less variation in the use of low-value care within hospital markets. Second, hospitals seem to learn from their neighbors about both forms of care at similar rates.

This local diffusion has important policy implications. If quality improvements at one hospital diffuse to others in the same hospital market, then there is a case for subsidizing investment in high-quality care, as this positive externality suggests that private investment will be too low. On the other hand, if less productive practice patterns also diffuse through a similar process, then that suggests that hospitals should be discouraged from investing in technologically intensive health care of questionable value. These findings thus have implications both for the optimal design of public subsidization of quality-improving investment and for payments for lower-value care through public insurance programs like Medicare. Both promoting the use

of underutilized high-quality care and discouraging the use of overused low-value care would improve the productivity of health care spending. Understanding the mechanisms through which these practices diffuse and the role that agglomeration economies play in that diffusion is crucial for designing policies that achieve these aims and should further inform the design of a wide array of reforms.

References

- Bach, P. B., D. Schrag, and C. B. Begg. 2004. Resurrecting treatment histories of dead patients: A study design that should be laid to rest. *Journal of the American Medical Association* 292 (22): 2765–70.
- Baicker, K., and A. Chandra. 2004a. Medicare spending, the physician workforce, and beneficiaries' quality of care. *Health Affairs*, April 7 (Web Exclusive, W4-184–W4-197).
- . 2004b. The productivity of physician specialization: Evidence from the Medicare program. *American Economic Review* 94 (2): 357–61.
- Berwick, D. M. 2003. Disseminating innovations in health care. *Journal of the American Medical Association* 289 (15): 1969–75.
- Bradley, E. H., J. Herrin, B. Elbel, R. L. McNamara, D. J. Magid, B. K. Nallamothu, Y. Wang, S.-L. T. Normand, J. A. Spertus, and H. M. Krumholz. 2006. Hospital quality for acute myocardial infarction: Correlation among process measures and relationship with short-term mortality. *Journal of the American Medical Association* 296 (1): 72–78.
- Chandra, A., and D. O. Staiger. 2007. Productivity spillovers in health care: Evidence from the treatment of heart attacks. *Journal of Political Economy* 115 (1): 103–41.
- Coleman, J. S., E. Katz, and H. Menzel. 1957. The diffusion of an innovation among physicians. *Sociometry* 20 (4): 253–70.
- Comin, D., and B. Hobijn. 2004. Cross-country technology adoption: Making the theories face the facts. *Journal of Monetary Economics* 51 (1): 39–83.
- Dartmouth Atlas Project. The Dartmouth Atlas of Health Care. Available at: <http://www.dartmouthatlas.org>.
- Fisher, E. S., D. E. Wennberg, T. A. Stukel, and D. J. Gottlieb. 2004. Variations in the longitudinal efficiency of academic medical centers. *Health Affairs*, October 7 (Web Exclusive, VAR-19–VAR-32).
- Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and E. L. Pinder. 2003a. The implications of regional variations in Medicare spending, part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine* 138 (4): 273–87.
- . 2003b. The implications of regional variations in Medicare spending, part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine* 138 (4): 288–98.
- Hall, B. 2004. Innovation and diffusion. In *The Oxford handbook of innovation*, ed. J. Fagerberg, D.C. Mowery, and R. R. Nelson, 459–84. Oxford: Oxford University Press.
- Hall, R., and C. I. Jones. 2007. The value of life and the rise in health spending. *Quarterly Journal of Economics* 122 (1): 39–72.

- Heidenreich, P. A., and M. McClellan. 2001. Trends in treatment and outcomes for acute myocardial infarction: 1975–1995. *American Journal of Medicine* 110 (3): 165–74.
- Higashi, T., N. S. Wenger, J. L. Adams, C. Fung, M. Roland, E. A. McGlynn, D. Reeves, S. M. Asch, E. A. Kerr, and P. G. Shekelle. 2007. Relationship between number of medical conditions and quality of care. *New England Journal of Medicine* 356 (24): 2496–504.
- Jaffe, A. B., M. Trajtenberg, and R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108 (3): 577–98.
- Jha, A. K., Z. Li, E. J. Orav, and A. M. Epstein. 2005. Care in U.S. hospitals: The Hospital Quality Alliance program. *New England Journal of Medicine* 353 (3): 265–74.
- Jha, A. K., E. J. Orav, Z. Li, and A. M. Epstein. 2007. The inverse relationship between mortality rates and performance in the Hospital Quality Alliance measures. *Health Affairs* 26 (4): 1104–10.
- Jovanovic, B., and Y. Nyarko. 1995. A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings papers on economic activity, microeconomics*: 247–305.
- . 1996. Learning by doing and the choice of technology. *Econometrica* 64 (6): 1299–310.
- Lister, J. 1867. On the antiseptic principle in the practice of surgery. *Lancet* 90 (2299): 353–6.
- Parente, S. L., and E. C. Prescott. 1994. Barriers to technology adoption and development. *Journal of Political Economy* 102 (2): 298–322.
- Rogers, E. M. 1995. *Diffusion of innovations*. 4th ed. New York: Free Press.
- Sacerdote, B. I. 2001. Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics* 116 (2): 681–704.
- Skinner, J., and D. Staiger. 2006. The diffusion of health care technology. Dartmouth College, Department of Economics. Unpublished Manuscript.
- Soumerai, S. B., T. J. McLaughlin, J. H. Gurwitz, E. Guadagnoli, P. J. Hauptman, C. Borbas, N. Morris, et al. 1998. Effect of local medical opinion leaders on quality of care for acute myocardial infarction: A randomized controlled trial. *Journal of the American Medical Association* 279 (17): 1358–63.
- U.S. Department of Health and Human Services, Centers for Medicare and Medicaid Services. Hospital compare. Available at: <http://www.hospitalcompare.hhs.gov>.
- Werner, R. M., and E. T. Bradlow. 2006. Relationship between Medicare's Hospital Compare performance measures and mortality rates. *Journal of the American Medical Association* 296 (22): 2694–702.
- Zucker, L. G., M. R. Darby, and M. B. Brewer. 1998. Intellectual human capital and the birth of U.S. biotechnology enterprises. *American Economic Review* 88 (1): 290–306.