Healthcare Exceptionalism? Performance and Allocation in the U.S. Healthcare Sector

Amitabh Chandra, Amy Finkelstein, Adam Sacarny and Chad Syverson*

January 2016

Abstract

The conventional wisdom in health economics is that idiosyncratic features of the healthcare sector leave little scope for market forces to allocate consumers to higher performance producers. However, we find robust evidence across a variety of conditions and performance measures that higher quality hospitals tend to have higher market shares at a point in time and to expand more over time. Moreover, we find that the relationship between performance and allocation is stronger among patients who have greater scope for hospital choice, suggesting a role for demand in allocation in the hospital sector. Our findings suggest that the healthcare sector may have more in common with "traditional" sectors subject to standard market forces than is often assumed.

JEL classification: D22, D24, I11

Keywords: Hospital Performance; Allocation; Healthcare

*Chandra: Harvard Kennedy School and NBER, Amitabh_Chandra@harvard.edu; Finkelstein: Department of Economics, MIT and NBER; afink@mit.edu; Sacarny: Department of Health Policy and Management, Columbia University Mailman School of Public Health, ajs2102@columbia.edu; Syverson: Chicago Booth and NBER; chad.syverson@chicagobooth.edu

We are grateful to Daron Acemoglu, Nick Bloom, Iain Cockburn, Chris Conlon, Angus Deaton, Mark Duggan, Joe Doyle, Liran Einav, Matthew Gentzkow, Michael Greenstone, Jonathan Gruber, Ben Olken, Jonathan Skinner, Doug Staiger, Scott Stern, Heidi Williams, numerous seminar participants, four referees, and the Editor for helpful comments and advice, and to Maurice Dalton and Nivedhitha Subramanian for expert research assistance. We gratefully acknowledge funding from the National Institute on Aging: P01 AG005842 and P01 AG019783 (Chandra), R01 AG032449 (Finkelstein) and T32-AG000186 (Sacarny). Finkelstein, Sacarny, and Syverson declare that they have no relevant or material financial interests that relate to the research described in this paper. Chandra reports serving on the Congressional Budget Office (CBO) Panel of Health Advisors. Other disclosures are available at https://www.hks.harvard.edu/about/faculty-staff-directory/amitabh-chandra

I Introduction

A classic "signpost of competition" in manufacturing industries is that higher productivity producers are allocated greater market share at a point in time and over time. The conventional wisdom in the healthcare sector, however, is that idiosyncratic, institutional features of this sector dull or eliminate these competitive reallocation forces. Oft-cited culprits include consumers who lack knowledge of or time to respond to the quality and price differences across providers, generous health insurance that insulates consumers from the direct financial consequences of their healthcare consumption decisions, and public sector reimbursement that provides little incentive for providers to achieve productive efficiency. These factors are widely believed to dampen the disciplining force of demand-side competition that exists in most other sectors.

This notion of "healthcare exceptionalism" has a long tradition in health economics. It dates back at least to the seminal article of Arrow (1963), which started the modern field of health economics by emphasizing key features of the health care industry that distinguish it from most other sectors and therefore warrant tailored study. Echoing and advancing the view that demand-side competition does not discipline healthcare providers, Cutler (2010) notes:

> "[T]here are two fundamental barriers to organizational innovation in healthcare. The first is the lack of good information on quality. Within a market, it is difficult to tell which providers are high quality and which are low quality. ... Difficulty measuring quality also *makes expansion of high-quality firms more difficult.* [emphasis added] ... The second barrier is the stagnant compensation system of public insurance plans." (p. 3)

In a similar vein, Skinner (2011) states in his overview article on regional variations in healthcare:

> "[Low productivity producers are] ... unlikely to be shaken out by normal competitive forces, given the patchwork of providers, consumers and third-party payers each of which faces inadequate incentives to improve quality or lower costs..." (p. 47)

In this paper, we question this conventional wisdom by investigating empirically whether and to what extent higher performing hospitals tend to attract greater market share. We look at allocation of Medicare patients for several different health conditions – heart attacks (called acute myocardial infarction, or AMI), congestive heart failure, and pneumonia – and a common pair of surgical procedures – hip and knee replacements – that together account for almost one-fifth of Medicare hospital admissions and hospital spending. Hospital "performance" or "quality" (words that we use interchangeably) is, of course a highly multidimensional object. Broadly speaking, we think of hospital quality as increasing in hospital attributes that increase the utility of patients or their surrogates; hospital quality therefore includes the ability of the hospital to generate good health outcomes, patient beliefs about the hospital's ability to generate good health outcomes, and patient satisfaction with the hospital experience. In practice, we examine several different hospital quality measures: clinical outcomes (survival and readmission), conformance with processes of care (i.e. adherence to well-established practice guidelines), and ex-post measures of patients' satisfaction with their experience (such as whether the room was quiet and whether nurses communicated well).

We find robust evidence that higher performing hospitals – as defined either by the health outcome-based measures or the process of care measures – tend to have greater market share (i.e. more Medicare patients) at a point in time, and experience more growth in market share over time. This positive correlation between quality and market share does not exist, however, when quality is measured by patient self-reported satisfaction with the hospital stay. Importantly, where we do find a positive correlation between quality and market share, these correlations are systematically and substantially stronger among patients who have more scope for choice. Specifically, within a condition the correlation between hospital quality and allocation is stronger for admissions that are transfers from other hospitals than admissions that come via the emergency room. We interpret these results as consistent with a role for consumer demand, either by patients or their surrogates, to affect the allocation of patients to hospitals. Also, consistent with consumer demand in a setting where there is little if any financial consequence of hospital choice for the patient, we find that conditional on hospital performance, the market does not penalize hospitals

with higher inputs – if anything, it rewards them. The normative implications of the reallocation we observe therefore differ for the patient and for a benevolent social planner.

Qualitatively, our results reject the strong form of the healthcare exceptionalism hypothesis: that there are no forces allocating market share to higher quality hospitals. Quantitatively, they suggest an important role for these reallocation forces. For example, we find that reallocation to higher quality hospitals can explain about a quarter of the 3.9 percentage point increase in 30-day survival for AMI over the 1996-2008 period. In other words, AMI survival rates rose almost one percentage point over the period simply because patient flows shifted to higher-quality hospitals. For heart failure and pneumonia – where the secular improvements in survival were, respectively, 0.9 and 3.2 percentage points over this time period – we find a somewhat smaller contribution of reallocation of 18 percent and 6 percent.

The rest of the paper proceeds as follows. Section II describes the analytical framework. Section III discusses our setting and data. Section IV presents our main results on the relationship between hospital quality and market share. Section V presents additional evidence consistent with a demand-based mechanism for these allocation results. The last section concludes.

II Analytical Approach: Static and Dynamic Allocation

Our primary empirical exercise examines the correlation between producer (i.e. hospital) performance and market share at a point in time, and the correlation between producer performance and growth in market share over time. This relationship has been analyzed extensively in a variety of industries and countries as a proxy for the role of competition in these settings (e.g. Olley and Pakes, 1996; Pavcnik, 2002; Escribano and Guasch, 2005; Bartelsman, Haltiwanger and Scarpetta, 2013; Collard-Wexler and De Loecker, 2015). Intuitively, competitive forces exert pressure on lower productivity firms, causing them to either become more efficient, shrink, or exit.

Models of such reallocation mechanisms among heterogeneous-productivity producers have found applications in a number of fields, including industrial organization, trade, and macroeconomics.[1] While these models differ considerably in their specifics, they share a common intuition: greater competition – as reflected in greater consumer willingness or ability to substitute to alternative producers – makes it more difficult for higher-cost, lower-productivity firms to earn positive profits, since demand is more responsive to cost and price differentials across firms. As substitutability increases, purchases are reallocated to higher productivity providers, raising the correlation between productivity and market share at a point in time ("*static allocation*") and causing higher productivity providers to experience more growth over time ("*dynamic allocation*").

The literature to date has focused on the relationship between market share and productivity, or the ratio of output to inputs. However, in the health care setting – and particularly for the Medicare enrollees that are the focus of this study – consumers bear little to none of the costs of production. As a result, it is more sensible to view competition as occurring mainly over output "performance", or quality, rather than productivity per se. In Appendix A we therefore present a model of quality competition among firms that face consumers who are not sensitive to input costs. This model preserves the intuition that consumer ability or willingness to substitute across providers drives the relationship between performance and market allocation. However, in a setting where, due to insurance, consumers have little or no financial stake in their selection, the market need not allocate away from firms that are higher cost for a given level of output.

For the static allocation analysis, we will use the following regression framework:

$$\ln(N_h) = \beta_0^s + \beta_1^s q_h + \gamma_M^s + \varepsilon_h^s \tag{1}$$

where N_h is a measure of the market size of hospital h, γ_M^s are market fixed effects, and q_h is a measure of the quality of hospital h. Thus β_1^s reflects the static

[1] See, for example, Ericson and Pakes (1995); Melitz (2003); and Asplund and Nocke (2006).

relationship between a hospital's quality and its market share within a market. If the coefficient is positive, as has been found with respect to productivity in many U.S. manufacturing industries (e.g., Olley and Pakes, 1996; Hortaçsu and Syverson, 2007; Bartelsman, Haltiwanger and Scarpetta, 2013), it indicates that higher performance producers have a greater share of activity at a point in time. If β_1^s is zero or negative, it indicates that lower quality facilities are the same size or larger than their high quality counterparts and suggests that forces beyond quality competition are driving the allocation of market activity. In the manufacturing productivity literature, $\beta_1^s \leq 0$ has been found in some former Soviet-bloc countries in the early 1990s (Bartelsman, Haltiwanger and Scarpetta, 2013) and in the U.S. steel industry circa 1960-70 (Collard-Wexler and De Loecker, 2015).[2]

The static allocation analysis in equation (1) can reflect the market's ability to reallocate activity from low quality hospitals to higher quality ones, but it shows the outcome of this process rather than the process itself. To measure the actual dynamics of the market's selection and reallocation mechanisms, we examine the relationship between a hospital's quality and its future growth. We will estimate:

$$\Delta_h = \beta_0^d + \beta_1^d q_h + \gamma_M^d + \varepsilon_h^d \tag{2}$$

where Δ_h is a measure of the hospital's growth rate in admissions and all other variables are defined as in equation (1). A positive correlation between quality and growth indicates that higher performance hospitals see larger gains in patient admissions, and points to the operation of a selection and reallocation process. The productivity literature has found widespread evidence in developed country manufacturing and retail that higher productivity producers experience growth in market shares (e.g. Scarpetta et al., 2002; Disney, Haskel and Heden, 2003; and Foster, Haltiwanger and Krizan, 2006).[3]

[2] A positive relationship between producer performance and market share could also reflect increasing returns to scale – increased size could cause performance to rise via e.g. learning by doing. In the health care sector this story is called the "volume-outcome" hypothesis, and we discuss its relationship to our findings when we consider alternative explanations for them in Section V.B.

[3] An even stronger result from this literature is that low productivity producers are more likely to exit the market entirely (see Bartelsman and Doms, 2000 and Syverson, 2011 for surveys). Hospital

Regression equations (1) and (2) form the heart of our empirical analysis. They describe the associations between a hospital's quality and market share and indicate whether forces exist that are favorable to the expansion of higher quality producers. Crucially, they show whether there is a reduced form relationship between our specific quality proxy measures and market allocation. A finding of $\beta_1^s > 0$ or $\beta_1^d > 0$ indicates a correlation between a quality proxy and market share. This may reflect that fact that patients and their surrogates directly value that quality proxy, or that they value attributes of the hospital that are correlated with it. Following the productivity literature, we do not take a stand on which heuristics of consumers generate the observed market allocation.

Although motivated by models in which competitive forces create these reallocation pressures, the static and dynamic correlations are naturally not direct evidence of the impact of competition. After presenting our allocation baseline results, we provide evidence consistent with quality competition as a driver of allocation by examining whether the allocation results are stronger among patients who have more scope for hospital choice. We also discuss possible alternative forces that may mimic the effects of competition, and present evidence suggesting that they are not primarily responsible for the allocation patterns we find in the data.

III Setting: Conditions and Quality Measures

We analyze allocation of Medicare patients for three medical conditions and a pair of common surgical procedures: heart attacks (AMI), congestive heart failure (hereafter heart failure or HF), pneumonia, and hip and knee replacements. Together, they account for 17 percent of Medicare hospital admissions and hospital spending in 2008, our base year for analysis. We selected conditions for which the Centers for Medicare and Medicaid Services (CMS) reports a variety of hospital- and condition-specific quality measures. AMI, HF, and pneumonia are the only three inpatient conditions for which CMS reports all of our quality measures in our base year (2008). Since they are predominantly emergency conditions, we added hip and knee replacement as the only non-emergency (i.e. deferrable) treatment condition

exit is poorly measured in our data, so we eschew this analysis and instead look at hospital growth.

for which one of our condition-specific quality measures was available. We link these hospital quality measures to data on each hospital's market share in treating the conditions at a point in time and over time. In the remainder of this section we describe first our data on allocation across hospitals and then our various proxies for hospital quality.

III.A Patient Data

Our primary data set on hospital size and growth consists of all Medicare Part A (i.e. inpatient hospital) claims for all AMI, heart failure, pneumonia, and hip and knee replacement hospital stays occurring in individuals age 66 and over in the United States in 2008 through 2010. We chose 2008 for our base year because it is the first year that all of our quality and allocation metrics could either be calculated by us or were well-populated by CMS. We avoid using more recent years because doing so would limit our ability to study dynamic allocation. In some of our additional analyses below, we use a similar data set spanning 1996 to 2010 to estimate survival (the quality measure we have going back the furthest in time) and allocation over a longer horizon.[4] As in the primary data set, these data encompass the universe of Medicare fee-for-service admissions for each of our four conditions for individuals age 66 and over in the United States in these years. The data also contain rich information on patient demographic and health characteristics (called risk adjusters on our context). Risk adjustment helps to address concerns that patient selection of hospitals might bias quality metrics.

Panel A of Table 1 shows the prevalence of each condition in the Medicare fee-for-service age 65+ population in 2008. The emergency conditions (AMI, heart failure, and pneumonia) are defined based on the patient's principal diagnosis on the reimbursement claim, which indicates the underlying condition that caused the

[4] In these data, hospitals that convert to critical access facilities (a program for rural hospitals) or that merge may change their Medicare identifiers and spuriously appear to close. We employ data on hospital identifier changes between 1994 and 2010 and group together all identifiers that ever refer to the same facility into one synthetic hospital. For example, if hospital A merges with hospital B and the two facilities begin sharing an identifier, we treat facilities A and B as one synthetic hospital throughout our analysis. We perform this aggregation for the quality measures as well. We thank Jon Skinner for generously sharing these data with us.

admission to the hospital. Hip and knee replacement patients are defined as patients who received a total hip or knee replacement procedure. Heart failure is the most common (accounting for over half a million patients per year, or 6 percent of Medicare discharges) and AMI is the least common (totaling about a quarter of a million patients per year or about 3 percent of discharges). Over 70 percent of AMI, heart failure, and pneumonia patients are admitted through the emergency room. By contrast, only 2 percent of hip and knee replacement admissions come via the emergency room, which is why we consider this condition non-emergent.

III.B Quality metrics

Hospital quality is multi-dimensional and includes any aspects of the hospital that might affect the patient or his surrogate's utility. We employ a variety of proxies intended to capture various dimensions of hospital quality: The measures, which are described in more detail in Appendix B, are each drawn from or based on publicly reported hospital-specific quality measures. They all are currently being used by CMS as the basis for financial incentives for hospitals.

Our first two hospital quality metrics capture condition-specific health outcomes: risk-adjusted 30-day survival rates and risk-adjusted 30-day readmission rates. Both are measured for Medicare patients with a given condition at a given hospital. Through the Hospital Value Based Purchasing Program and Readmissions Reduction Program, respectively, CMS is now adjusting its payments to hospitals to reward those that provide high quality care on these two dimensions (Rau, 2013).

Risk-adjusted readmission is the only condition-specific measure which CMS reports for hip and knee replacement. Hip and knee surgeries are the second most common surgical conditions to occur before re-hospitalizations in Medicare (Jencks, Williams and Coleman, 2009) and readmission (adjusted for patient risk factors) is a well accepted quality metric for hip and knee replacement (e.g. Jencks, Williams and Coleman, 2009; Grosso et al., 2012). In 2015, the hip and knee replacement measure was added to the Readmissions Reduction Program to incentivize facilities to keep patients out of the hospital during recovery (Kahn et al., 2015).

Our third quality measure is a condition-specific "process of care" measure

which captures the hospital's conformance with established clinical guidelines. Specifically, it measures the shares of eligible patients who received certain evidence-based interventions. The data pertain to all patients irrespective of their insurer and so are not limited to patients covered by Medicare. The processes "were identified with respect to published scientific evidence and consistency with established clinical-practice guidelines" (Williams et al., 2005). For example, the AMI processes cover the administering of aspirin, ACE inhibitors, smoking cessation advice, β blockers, and angioplasty. The measures have been widely analyzed in the medical, health policy, and health economics literature (e.g. Jencks et al., 2000; Jencks, Huff and Cuerdon, 2003; Jha et al., 2007; Werner and Bradlow, 2006; Skinner and Staiger, 2007, 2015). They are also now used to adjust payments, with the Medicare Value Based Purchasing Program rewarding hospitals for high levels and growth in the process measures (Blumenthal and Jena, 2013).

Our final quality measure captures overall (hospital-level) patient satisfaction with the hospital experience on a variety of dimensions, such as whether nurses communicated well or the rooms were quiet. The measures come from the 2008 HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems), a survey that hospitals administer to their patients following discharge. All patients are included, not just those covered by Medicare, and unlike the other metrics the results are not disaggregated by health condition. The survey results are processed and reported by CMS; the survey instrument is condensed into 10 measures of the patient's experience and perceived quality of care. CMS performs an adjustment for interview mode (e.g. mail, telephone, etc.) and patient characteristics. Like the process of care measures, high and growing survey scores are now being rewarded by the Value Based Purchasing Program (Blumenthal and Jena, 2013).

The four quality measures capture distinct aspects of hospital performance. Risk-adjusted survival is arguably the key endpoint for emergent conditions and has been the health outcome of choice for a large economics and medical literature (see e.g. Andersen et al., 2003 for a typical medical trial example and Cutler et al., 1998 for a classic example of survival as an endpoint in economics and health). Risk-adjusted readmission is widely used as a proxy for medical errors and inappropriate discharge (e.g. Anderson and Steinberg, 1984; Axon and Williams, 2011;

Jencks, Williams and Coleman, 2009). The process of care measures are designed to measure interventions that the facility should deliver to all appropriate patients; the study of processes of care has long been motivated by the concept that hospitals may have more control over them than over health outcomes like survival or readmission, since hospitals have limited influence over which patients they treat and how patients comply with care after discharge (Donabedian, 1966). Patient satisfaction is designed to capture patients' self-reports of ex-post satisfaction with aspects of their hospital experience (Giordano et al., 2010).

III.C Summary statistics

Sample restrictions and potential measurement error in quality measures

In all of our analyses, we limit the sample for each condition to hospitalizations among patients who have not had an inpatient stay for that condition in the prior year. We call these hospitalizations index events.[5] We exclude patients who are poorly observed in our data because their Medicare coverage is incomplete (i.e. they failed to enroll in both parts A and B of Medicare) or they were enrolled in a private Medicare Advantage plan. These patients cannot be tracked well over time, so even when we observe their hospitalizations, we cannot assign them to index events. In all of our allocation analyses, we exclude hospitals with no index admission for that condition in 2008.[6] In addition to the above restrictions which apply to all of our analyses, we make some additional condition- and quality metric-specific restrictions and adjustments as described below.

[5]Focusing on index events is useful for our allocation exercises because it allows us to think of each observation as an episode of care, treating readmissions and other health expenditures endogenous to the course of treatment in the initial stay as part of that episode rather than as new events. For example, a second admission to the hospital (within a year) for the treatment of an AMI will not count as an index event, so a hospital that frequently readmits its patients will not appear to capture market share as a result.

[6]This restriction introduces a potential concern about selection on the dependent variable (the logged number of patients in 2008) in the static analysis in equation (1); this is not a concern for the subsequent dynamic analysis. We therefore explored the sensitivity of our static allocation results to an alternative, Tobit-style truncated regression which adjusts for the truncation under a normality assumption. We found that the static allocation results were slightly strengthened by this adjustment (see Appendix Table A1).

The combination of a relatively small number of patients in some hospitals together with the stochastic nature of some of the quality outcomes means that our quality metrics may be estimated with error. Such estimation error may cause attenuation bias in our analysis of the relationship between market share and hospital quality in equations (1) and (2). We take a number of steps to help address this concern. First, in constructing our quality metrics, we aggregate data for our condition-specific measures (risk-adjusted survival, risk-adjusted readmission, and process of care) over the three-year period 2006-2008. Second, we restrict our sample to hospitals with a minimum number of patients per condition over the three-year measurement period; the cutoff threshold varies across our quality measures as described in Appendix B. For example, for risk-adjusted survival, we follow CMS and restrict to hospitals with at least 25 patients for that condition over 2006-2008.

Third, for our clinical outcomes (survival and re-admission), we apply the standard shrinkage or "smoothing" techniques of the empirical Bayes literature (e.g. Morris, 1983) to adjust for estimation error in our hospital-specific estimates. McClellan and Staiger (2000) introduced this approach into the healthcare literature when estimating quality differences across hospitals, and it has since been widely applied in the education literature for estimating and analyzing teacher or school value added measures (e.g. Kane and Staiger, 2001; Jacob and Lefgren, 2007). The intuition behind it is that when a hospital's quality is estimated to be far above (below) average, it is likely to be suffering from positive (negative) estimation error. Therefore, the expected level of quality, given the estimated quality, is a convex combination of the estimate and the mean of the underlying quality process. The relative weight that the estimate gets in this convex combination varies inversely with the noise of the estimate (which is based on the standard error of the hospital fixed effect). Appendix C provides a detailed description of the procedure.[7]

[7] In practice, as we show in Appendix Table A2 and Appendix Section C.5, our core findings using these quality metrics remain statistically significant without the empirical Bayes adjustment, although naturally the magnitude is attenuated.

Static and Dynamic Allocation

Table 1 - Summary Statistics on Allocation Metrics across Conditions

Condition	(1) AMI	(2) Heart Failure	(3) Pneumonia	(4) Hip/Knee
Panel A - Composition of all Medicare Discharges in 2008				
Number of patients in 2008	263,485	545,363	475,756	350,536
Share through Emergency Dept	0.71	0.76	0.76	0.02
Share of all Medicare discharges	0.03	0.06	0.05	0.04
Share of Medicare hospital spending	0.04	0.05	0.04	0.05
Number of hospitals in 2008	4,257	4,547	4,607	3,297
Panel B - Static Allocation: Patients in 2008				
Patients (Index Events)	190,189	308,122	354,319	267,557
Average No. of Patients per Hospital	65.8	76.6	81.9	101.7
SD of Patients per Hospital	67.6	78.2	70.8	118.0
Hospitals	2,890	4,023	4,325	2,632
Average No. of Hospitals per Market	9.4	13.1	14.1	8.6
Panel C - Dynamic Allocation: Growth in Patients from 2008 to 2010				
Average Growth Rate across Hospitals	-0.17	-0.10	-0.13	-0.03
SD across Hospitals	0.42	0.38	0.36	0.46
Hospitals	2,890	4,023	4,325	2,632

Panel A is calculated on a 100% sample of age 65+ fee-for-service Medicare patients in 2008 and counts all patients with the condition, not just the index events that are the subject of the remainder of this study and Panels B and C. The sample in Panels B and C is all hospitals that had at least 1 index admission in 2008 for the condition shown in the column heading and had a valid risk-adjusted survival rate for that condition (risk-adjusted readmission for hip/knee replacement). There are 306 hospital markets, called Hospital Referral Regions (HRRs). Growth is calculated based on the formula in equation (3) that restricts values to between -2 and 2.

Panels B and C of Table 1 present some summary statistics on our static and dynamic allocation measures, respectively. As discussed, the hospital sample varies by the condition and quality metric; for illustrative purposes we report allocation statistics for the hospitals for which we construct the risk-adjusted survival metric (for the emergency conditions) or risk-adjusted readmission metric (for hip and

knee replacement). There are fewer patients and hospitals in these panels than in Panel A because here we limit to index event hospitalizations.

For our static allocation analysis in equation (1), our measure of hospital market size N_h is the number of Medicare patients with the given condition in 2008 who were treated in hospital h – in other words, this is a count of the index events that can be attributed to the hospital. Across the conditions, Panel B shows that the average hospital treated between 66 and 102 Medicare patients in 2008. The standard deviation of hospital size ranges from 68 to 120.[8]

Panel C reports summary statistics on growth in patients from 2008-2010 (i.e. Δ_h). We define this variable as:

$$\Delta_h = \frac{N_{h,2010} - N_{h,2008}}{\frac{1}{2}\left(N_{h,2010} + N_{h,2008}\right)} \tag{3}$$

where $N_{h,t}$ is the number of Medicare patients with the given condition treated by hospital h in year t. Our measure of the hospital's two-year growth rate thus divides the change in the number of patients between the two years by the average number of patients in these two years.[9] Panel C shows substantial dispersion in this growth rate across the facilities, with the standard deviation of the measure ranging from 36 to 46 percentage points.

For all the conditions, Panel C also shows that the average hospital experiences a negative growth in the number of patients between 2008 and 2010. The largest decline occurs for AMI, where the average hospital treats 17 percent fewer patients in 2010 compared to 2008, and the smallest occurs for hip and knee replacement, at 3 percent. The overall decline in patients reflects three factors. First, there is a secular decline in inpatient admissions for these conditions overall (not just in

[8]Hospital size distributions have a long tail. For example, the 10th size percentile hospital treats 10 AMI patients, the 90th percentile treats 151 patients, and the 99th percentile treats 322 patients (not shown).

[9]This transformation of the standard percentage growth rate metric bounds growth between -2 (exit) and +2 (growth from an initial level of 0). An attraction of this transformation is that it reduces the chance that the results are skewed by a few fast-growing but initially small hospitals that would have very large percentage growth rates. This growth rate transformation has been used in other contexts to avoid unnecessary skewness in the growth rate measure; see, for example, Davis, Haltiwanger and Schuh (1998).

Medicare) over this time period.[10] Second, Medicare Advantage, the program that allows Medicare enrollees to receive private insurance, expanded between 2008 and 2010, and these enrollees are excluded from our sample.[11] Finally, our quality measures require the hospital to have at least 1 patient in 2008 and enough patients in 2006-2008 to calculate the measure accurately (see Appendix B), so regression to the mean will also reduce average growth.

We follow the literature in defining a hospital market as a Hospital Referral Region (HRR).[12] Our sample includes 306 HRRs. On average, the emergent conditions have 9 to 14 hospitals per HRR while hip and knee replacement has 9 hospitals per HRR. In Appendix Table A3, we show that the great majority (87 percent to 90 percent) of patients stay within their market of residence when receiving treatment for the emergent conditions and a slightly lower share (84 percent) stay in their market for hip and knee replacement.

Quality metrics

Table 2 presents basic summary statistics on the quality metrics. It shows clinically and economically meaningful dispersion across hospitals in all of the measures.

[10] See http://goo.gl/FJ0Nvy and http://goo.gl/zeHg2J for this data.

[11] See http://kff.org/medicare/state-indicator/enrollees-as-a-of-total-medicare-population/

[12] The *Dartmouth Atlas of Healthcare* divides the United States into HRRs, which are determined at the ZIP code level through an algorithm that reflects commuting patterns to major referral hospitals. HRRs, which are akin to empirically defined markets for healthcare, may cross state and county borders. A complete list of HRRs can be found at http://www.dartmouthatlas.org/.

Table 2 - Summary Statistics on Quality Metrics across Conditions

Condition	(1) AMI	(2) Heart Failure	(3) Pneumonia	(4) Hip/Knee
Panel A - Risk-Adjusted Survival Rates (30 Days): Patients in 2006-2008				
Average 30-Day Survival Rate	0.82	0.89	0.88	
(SD of Risk-Adjusted Measure)	(0.03)	(0.02)	(0.02)	
Hospitals in Risk-Adjusted Measure	2,890	4,023	4,325	
Panel B - Risk-Adjusted Readmission Rates (30 Days): Patients in 2006-2008				
Average 30-Day Readmission Rate	0.21	0.21	0.16	0.06
(SD of Risk-Adjusted Measure)	(0.03)	(0.02)	(0.02)	(0.02)
Hospitals in Risk-Adjusted Measure	2,322	3,904	4,264	2,632
Panel C - Processes of Care: Shares of Patients Receiving Appropriate Treatments in 2006-2008				
Average Score	0.93	0.83	0.88	
(SD)	(0.05)	(0.14)	(0.07)	
Hospitals	2,398	3,666	3,920	
Average No. of Processes Reported	4.40	3.30	6.22	
Panel D - Patient Survey: Survey Covers All Patients in 2008 (Not Limited to Particular Condition)				
Avg Overall Rating (1-3, higher is better)	2.53	2.53	2.53	2.53
(SD)	(0.14)	(0.14)	(0.14)	(0.14)
Hospitals	3,498	3,598	3,610	3,061

Sample restrictions are specific to the condition and quality metric; see text for more details of the metric definitions and sample restrictions. Summary statistics are reported across hospitals. In Panels A and B, the standard deviations are of the risk-adjusted measures and are empirical-Bayes-adjusted to account for measurement error (see Appendix Section C.3.1). In Panel D, the number of hospitals differs across conditions even though the patient survey metric is not condition-specific because we calculate the ratings on the subset of hospitals that reported at least 1 patient with the condition in 2008.

Panel A shows 30 day survival rates, which range from a hospital-level average of 82 percent for AMI to 89 percent for heart failure; as a non-emergent condition, survival is not considered a relevant metric for hip and knee replacement. The standard deviation of these rates across hospitals ranges from 2 to 3 percentage points after adjusting for patient risk factors, suggesting that some facilities are capable of generating higher survival than others.

Panel B shows that the average hospital-level readmission rate ranges from 6 percent for hip and knee replacement to about 21 percent for AMI and heart failure. Like survival, the cross-facility standard deviations are 2 to 3 percentage points after adjustment for patient risk factors.

Panel C reports on the process of care measure. In our allocation results, we combine the condition-specific individual process of care scores into a single composite, standardized (i.e. mean 0 and standard deviation 1), condition-specific score. To give a sense of the metric, we present here for each condition a score that is generated by taking each hospital's average utilization of the condition's processes, then averaging the result across hospitals. The reported score of 0.93 for AMI means that for the average hospital, the average utilization rate across the 6 AMI treatments is 93 percent. Compliance with the processes is lower for heart failure (83 percent) and pneumonia (88 percent). The dispersion in compliance across hospitals is larger than in risk-adjusted survival and readmission – it ranges from 5 percentage points for AMI to 14 percentage points for heart failure.

The patient survey is reported in Panel D. In order to capture the full breadth of questions included in the survey in our allocation analyses, we mimic our approach for the process of care metric and use a standardized (i.e. mean 0, standard deviation 1) composite of all the survey questions. To give a flavor for the measures, Panel D reports the results of one of the questions: a patient-reported overall rating of the hospital. The table reports the average score across hospitals when low, medium, and high are valued at 1, 2, and 3 respectively. The average patient at the average hospital gives between a medium and a high rating, and this is true even for the hospital two standard deviations below the average.

We examined the correlation of quality measures across hospitals and conditions. For a given hospital quality measure, hospital quality is strongly positively correlated across conditions (see Appendix Table A4); for example, the within-hospital correlation of risk-adjusted readmission between the four conditions ranges from 0.44 to 0.94.

Table 3 examines the correlation of hospital quality measures within each condition (though as stated, the patient survey covers all patients). Higher values of all these quality measures are desirable, except for risk-adjusted readmission. Most

of the correlations are of the expected sign: risk-adjusted survival and process of care are positively correlated, and risk adjusted readmission and process of care are negatively correlated.[13] However the correlations are substantially below 1, suggesting that these measures may be capturing different dimensions of the hospital experience.[14]

Table 3 - Correlation of Quality Metrics within Condition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	AMI				HF			
Metric	Risk-Adj Survival	Risk-Adj Readm	Process of Care Z	Patient Survey Z	Risk-Adj Survival	Risk-Adj Readm	Process of Care Z	Patient Survey Z
Risk-Adjusted Survival	1.00				1.00			
	[2,890]				[4,023]			
Risk-Adjusted Readmission	0.03	1.00			0.35	1.00		
	[2,322]	[2,322]			[3,904]	[3,904]		
Process of Care Z-Score	0.24	-0.25	1.00		0.17	-0.15	1.00	
	[2,346]	[2,214]	[2,398]		[3,607]	[3,578]	[3,666]	
Patient Survey Z-Score	-0.06	-0.26	0.18	1.00	-0.18	-0.36	0.01	1.00
	[2,799]	[2,293]	[2,370]	[3,498]	[3,447]	[3,398]	[3,392]	[3,598]
	Pneumonia				Hip/Knee Replacement			
Metric	Risk-Adj Survival	Risk-Adj Readm	Process of Care Z	Patient Survey Z	Risk-Adj Survival	Risk-Adj Readm	Process of Care Z	Patient Survey Z
Risk-Adjusted Survival	1.00							
	[4,325]							
Risk-Adjusted Readmission	0.08	1.00				1.00		
	[4,264]	[4,264]				[2,632]		
Process of Care Z-Score	0.08	-0.18	1.00					
	[3,871]	[3,847]	[3,920]					
Patient Survey Z-Score	-0.03	-0.36	0.18	1.00		-0.23		1.00
	[3,527]	[3,503]	[3,512]	[3,610]		[2,542]		[3,061]

Hospitals used to calculate correlation in brackets. All quality metrics are condition-specific except the patient survey, which is only available as an all-patient average. Correlations involving risk-adjusted survival and readmission are adjusted to account for measurement error (see Appendix Section C.3.2).

[13] Risk-adjusted survival and risk-adjusted readmission are positively correlated; this ostensibly surprising pattern has been previously documented (see e.g. Gorodeski, Starling and Blackstone, 2010) and at least partly reflects the fact that mortality and readmission are competing risks, since patients who die cannot be readmitted.

[14] Hospitals are multi-product firms that treat many different conditions. Understanding why performance correlates across outputs is beyond the scope of this study, but is a potentially fruitful avenue for future research.

Patient satisfaction does not have a systematic correlation with our other quality measures. As has been found previously in the literature (see e.g. Jha et al., 2008 and Boulding et al., 2011), it is positively correlated with hospital performance as measured by readmission and process of care. However, we find that it is negatively correlated with risk-adjusted survival rates.[15] These ambiguous findings for patient satisfaction are not new to the quality measurement literature, and align with concerns of physicians who question the value of patient satisfaction as an informative measure of hospital quality (Manary et al., 2013).

IV Allocation Results

IV.A Static and Dynamic Allocation

Table 4 presents our central results. The left-hand panel shows static allocation results based on the estimation of equation (1). These results relate the hospital's log number of patients for a given condition in 2008, $\ln(N_h)$, to a given quality measure for that hospital in 2008, q_h. Because we include market (HRR) fixed effects, this estimate is within market, relating a hospital's market share of patients with a given condition to its quality relative to other hospitals in its market. Each panel shows results from a separate regression using the reported quality measure.

[15] In Appendix Table A5 we also find that patient satisfaction tends to be negatively correlated with the Bloom et al. (2012) measure of hospital management quality for the several hundred hospitals for which this measure is available, while both risk-adjusted survival and process of care scores are positively correlated with hospital management scores; there is no clear pattern with respect to risk-adjusted readmission. We are extremely grateful to Nick Bloom for providing us with these measures.

Table 4 - Allocation across Conditions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation				Dynamic Allocation			
Measure \ Condition	AMI	HF	Pneu	Hip/Knee	AMI	HF	Pneu	Hip/Knee
Risk-Adjusted Survival								
Coef on Survival Rate	17.496	15.360	5.140		1.533	0.774	1.220	
	(0.995)	(1.320)	(0.777)		(0.379)	(0.501)	(0.354)	
Hospitals	2,890	4,023	4,325		2,890	4,023	4,325	
Risk-Adjusted Readmission								
Coef on Readmission Rate	-9.162	-10.346	0.499	-21.037	-1.428	-2.300	-1.138	-1.112
	(1.621)	(1.782)	(1.575)	(2.027)	(0.611)	(0.651)	(0.679)	(0.836)
Hospitals	2,322	3,904	4,264	2,632	2,322	3,904	4,264	2,632
Process of Care Z-Score								
Coef on Process Z-Score	0.319	0.332	0.211		0.048	0.043	0.026	
	(0.026)	(0.016)	(0.015)		(0.010)	(0.009)	(0.009)	
Hospitals	2,398	3,666	3,920		2,398	3,666	3,920	
Patient Survey Z-Score								
Coef on Survey Z-Score	-0.321	-0.252	-0.210	0.057	-0.065	-0.003	0.007	0.037
	(0.052)	(0.038)	(0.030)	(0.051)	(0.015)	(0.011)	(0.011)	(0.022)
Hospitals	3,498	3,598	3,610	3,061	3,498	3,598	3,610	3,061

The static allocation results are estimated using equation (1), a hospital-level regression of log-patients in 2008 on market fixed effects and the quality measure named in the row. The dynamic allocation results are estimated using equation (2), which is an identical regression except for the dependent variable, which is now growth in patients from 2008 to 2010. Growth is defined as in equation (3). Standard errors are bootstrapped with 300 replications and are clustered at the market level. Risk-adjusted survival and readmission are reported in percentage points (e.g. a value of 0.1 is 10 percentage points); process of care and patient survey metrics are reported in standard deviation units (e.g. a value of 1 is 1 standard deviation).

We find a statistically significant and positive relationship between hospital quality and market share for three of the four quality metrics (risk-adjusted survival, risk-adjusted readmission, and process of care). This suggests that, within a market, more market share (patients) tends to be allocated to higher quality hospitals at a point in time. For AMI patients (column 1), our estimates indicate that a 1 percentage point increase in a hospital's risk-adjusted AMI survival rate is associated with a 17 percent higher market share (or equivalently, due to the presence of market fixed effects, 17 percent more patients), a 1 percentage point reduction in the hospital's readmission rate is associated with 9 percent more patients, and a 1 standard deviation increase in the use of consensus AMI treatments (processes of care)

is associated with 32 percent more patients; all of these results are statistically significant. For heart failure patients (column 2) results are similar in magnitude and statistical significance. For pneumonia patients (column 3) the results are smaller in magnitude but still statistically significant for risk-adjusted survival and process of care; they are wrong-signed but insignificant for risk-adjusted readmission. For hip and knee replacement (column 4) we only observe the risk-adjusted readmission measure, which is statistically significant with the expected sign.

The right-hand panel shows dynamic allocation results based on the estimation of equation (2). These estimates examine the within-market relationship between a hospital's quality q_h and its subsequent two-year growth Δ_h, as defined in equation (3). Again, each panel shows results from a separate regression using the reported quality measure. The results once more tend to show a statistically significant positive relationship between measures of hospital quality and market share, with the exception of the patient satisfaction survey. For example, for AMI the results indicate that a 1 percentage point increase in the hospitals's risk-adjusted survival rate is associated with 1.5 percentage points higher growth in AMI patients relative to other hospitals in the same market. A hospital with a 1 percentage point lower risk-adjusted AMI readmission rate would tend to grow its AMI patient load 1.5 percentage points faster than other hospitals in the market, and a 1 standard deviation increase in utilization of AMI processes of care is associated with 4.8 percentage points higher growth. All of these results are statistically significant. The results are similar for the other three conditions – with higher risk-adjusted survival, lower readmission, and better process of care scores associated with greater two-year growth – and they are mostly (but not always) statistically significant.

The patient survey score is an exception to our general finding that higher quality hospitals tend to be larger (in the static allocation results) and grow faster (in the dynamic allocation results) than their peers. As previously discussed, hospitals' scores on the patient satisfaction survey are negatively correlated with some of our other quality metrics, and there is debate over the survey's value as a measure of hospital quality. These facts may explain our findings. Alternatively, the fact that market share appears correlated with the health and process of care measures rather than patient satisfaction could reflect which factors drive the demand of patients or

their surrogates. For example, patients may not know or may not value features on the survey such as how quiet rooms are at night. It is also possible that – as the one quality metric that is not condition-specific – the patient satisfaction measure has less relevance for the condition-specific allocation decisions.

To probe a little further on the different quality metrics, we also analyzed allocation putting the whole vector of quality metrics on the right-hand side of equations (1) and (2). Appendix Table A6 reports the results. Not surprisingly, given that these variables are highly correlated (see Table 3), the magnitude of the coefficients on the individual quality metrics often attenuate and, for many of the dynamic analyses, are no longer statistically significant.[16] Overall, however, they suggest an association between market share and each individual quality measure, conditional on the others, that is qualitatively similar to the unconditional correlations shown in Table 4.

We also considered the relationship between allocation for a given condition and quality measures for multiple conditions included in the regression simultaneously. Appendix Table A8 shows the results. Again, we find that multiple quality measures tend to matter. While own-condition quality usually remains a significant predictor of allocation, allocation generally loads onto several conditions' quality measures, and the AMI quality measure is often the most quantitatively important. These results likely reflect that within-hospital quality is highly correlated across conditions (see Table A4), and that our condition-specific quality measures are each noisy measures of underlying condition quality. The AMI measures may offer relatively more precise signals in comparison to the measures of the other conditions, or they may be more salient to consumers.

Robustness

We examined the robustness of our main allocation findings in Table 4 along a number of dimensions. As previously discussed, we show our static allocation results

[16] Part of the weakening of the results is also due to the fact that we limit the multivariate analysis in Appendix Table A6 to the subset of hospitals that report all four quality measures. Appendix Table A7 shows that when we run the allocation results separately for each quality measure as in Table 4 but limited to this subset of hospitals for which all four quality measures are available, the coefficients are somewhat attenuated relative to our baseline results.

are robust to alternative ways of handling the truncation on sample size (see footnote 6 and Appendix Table A1). We also show that our core allocation findings for survival and readmission remain without the empirical Bayes adjustment, although naturally the magnitude is attenuated (see footnote 7 and Appendix Table A2). Furthermore, we find that our static allocation analysis is not sensitive to an alternative specification, Poisson regression (see Appendix Table A9).

Finally, we explored the sensitivity of our findings to how we handle risk adjustment. A potential concern with both the survival and readmission quality measures is that they may be capturing heterogeneity in patient health across hospitals; this concern is muted for the process of care measures, which exclude patients who were inappropriate for each standard of care. Fortunately, there exist rich data on the relevant health characteristics of the patients (called risk-adjusters) which we use in creating our survival and readmission metrics. Of course, such risk adjustment is only as good as the observable characteristics on which it is based. Risk adjustment for AMI based on our observables has recently been cross-validated by research exploiting ambulance catchment areas as a source of exogenous variation in the allocation of patients to hospitals (Doyle, Graves and Gruber, 2014). In Appendix Table A10 we show that the results are only slightly affected if we instead use coarser risk adjustment (age/race/sex only) or no risk adjustment at all. These results help mitigate concerns that additional risk adjustment would attenuate or eliminate our findings. Moreover, we showed in previous work that additional risk adjustment with extremely rich data (available for a subset of AMI patients) has little effect on allocation results (Chandra et al., 2013).

IV.B Benchmarking the magnitude of reallocation

IV.B.1 Quality vs. distance

As one way of benchmarking the magnitude of our allocation results, we compared the reallocation of market share associated with higher quality to the reallocation associated with shorter distance between patient and hospital. Distance-to-hospital is a classic hospital attribute that has been extensively analyzed as a measure of hospital "price", with the general finding that individuals consider greater distance

to the hospital as a disamenity (see e.g. Luft et al., 1990; Town and Vistnes, 2001; Gaynor and Vogt, 2003; Tay, 2003; Romley and Goldman, 2011).

To compare allocation on quality to that on distance, we adapt our static allocation analysis in the spirit of the existing distance-to-hospital choice literature. Specifically, we specify the utility function of consumer p for hospital h as:

$$U_{ph} = \rho_1 \cdot distance_{ph} + \rho_2 \cdot distance_{ph}^2 + \theta \cdot q_h + \varphi_{ph}. \tag{4}$$

U_{ph} is the utility of a potential choice, $distance_{ph}$ is the distance from the patient to the hospital (entering as a quadratic; operationally we measure the distance from the patient's ZIP code of residence to the hospital's ZIP code), and q_h is the hospital's quality metric. There is also a component of utility that is idiosyncratic to the patient-hospital pair φ_{ph}. We assume that φ_{ph} is distributed type 1 extreme value, which means the problem can be readily estimated as a conditional logit, as is standard in the hospital choice literature (e.g. Dranove and Satterthwaite, 2000).

When consumers maximize this utility function, the realized choice probabilities are:

$$\Pr(C_p = h) = \frac{\exp\left(\rho_1 \cdot distance_{ph} + \rho_2 \cdot distance_{ph}^2 + \theta \cdot q_h\right)}{\sum_{h' \in H_{M(p)}} \exp\left(\rho_1 \cdot distance_{ph'} + \rho_2 \cdot distance_{ph'}^2 + \theta \cdot q_{h'}\right)}, \tag{5}$$

where C_p indicates the hospital that the patient chose for treatment and $H_{M(p)}$ is the patient's choice set of hospitals; we define the choice set as all hospitals in the patient's hospital market $M(p)$ that treat at least 1 patient with that condition in 2008 and for which we observe the relevant quality metric. Because of our definition of the patient's choice set, our analysis – unlike the allocation analysis in Table 4 – excludes any patient who left her market of residence for treatment. As shown in Appendix Table A11, this restriction excludes 10 to 16 percent of patients depending on the condition, but does not affect the basic static allocation results.

Table 5 presents the results from the conditional logit choice model. Across the columns are different health conditions. The panels report the marginal rates of

substitution (MRS) between a given quality metric and distance; each panel shows the results from a separate regression using the reported quality measure (Appendix Table A12 presents the raw logit coefficients). The reported marginal rate of substitution of quality for distance, evaluated at the average distance traveled by patients for that condition, is derived from the conditional logit estimates as:

$$MRS = \frac{\partial U_{ph} / \partial q_h}{\partial U_{ph} / \partial distance_{ph}} = \frac{\theta}{\rho_1 + 2 \cdot \rho_2 \cdot \overline{distance}}. \tag{6}$$

When the MRS is negative, it implies that the quality measure is a good, i.e. that patients are willing to travel farther to gain access to more quality. When it is positive, it implies that the quality metric is a bad.

Table 5 - Choice Model of Patient Allocation across Conditions

Condition	(1) AMI	(2) HF	(3) Pneumonia	(4) Hip/Knee
Mean Miles to Chosen Hospital	12.48	8.27	7.49	13.16
SD Miles to Chosen Hospital	20.06	13.25	11.92	18.85
Risk-Adjusted Survival				
MRS(1 pp risk-adjusted survival, miles)	-1.793	-1.029	-0.378	
	(0.158)	(0.129)	(0.057)	
Patients	165,005	275,671	317,904	
Risk-Adjusted Readmission				
MRS(1 pp risk-adjusted readmission, miles)	1.138	1.040	0.451	2.385
	(0.173)	(0.122)	(0.109)	(0.268)
Patients	158,086	274,667	317,374	222,673
Process of Care Z-Score				
MRS(1 SD process of care, miles)	-4.418	-2.238	-1.325	
	(0.383)	(0.221)	(0.110)	
Patients	158,032	270,773	309,623	
Patient Survey Z-Score				
MRS(1 SD patient survey, miles)	0.324	-0.093	0.036	-1.604
	(0.388)	(0.205)	(0.151)	(0.382)
Patients	167,429	266,915	298,185	224,451

This table reports the marginal rates of substitution (MRSs) of quality for distance derived from the conditional logit model (see equation 6). For the survival and readmission rates, the MRS given by equation (6) is divided by 100 to put it into percentage point terms. Only one quality measure is used at a time in each logit model. Standard errors are analytic and clustered at the market level.
The sample is all patients with the condition in 2008 who stayed in their market of residence for treatment. The choice set for a patient is all hospitals in his market with the quality measure available that treated at least one patient in 2008. The mean and SD miles statistics are taken from the patients in the column's risk-adjusted survival sample (risk-adjusted readmission for hip/knee replacement). All MRSs in a column are evaluated at this mean.

Qualitatively, our results are what would be expected given the existing literature and the results from our static allocation analysis: hospital choices are consistent with a willingness to travel longer distances in order to receive treatment at facilities with better health outcomes and processes, but not higher patient survey scores. Quantitatively, the results indicate that the average AMI patient is willing to travel 1.8 more miles (about one-tenth of the standard deviation of distance traveled for AMI) to gain access to a hospital with 1 percentage point greater risk-adjusted survival, 1.1 miles for 1 percentage point lower risk-adjusted readmission, and 4.4

miles for 1 standard deviation unit greater use of processes of care.[17]

These results are broadly similar to other estimates of "willingness to travel" to a different hospital. Perhaps most directly comparable to our estimates is Tay (2003)'s analysis of hospital choice for Medicare patients with AMI in 3 states in 1994. She finds that distance, hospital mortality rates, and hospital complication rates are all disamenities in patients' hospital choices. Her results imply, for example, that younger white male patients are willing to travel 8.0 miles to access a hospital with a 1 percentage point lower mortality rate and 1.7 miles to access a 1 percentage point lower complication rate. In addition, Romley and Goldman (2011) look at hospital demand for Medicare patients with pneumonia within Los Angeles over 2000-2004 and estimate a a willingness to travel ranging from 2.4 to 3.9 miles to move from the hospital at the 25th to 75th percentile of the distribution of hospitals' revealed utility.

IV.B.2 Contribution to survival gains

Another way to benchmark the allocation results is to explore their contribution to the secular improvements in survival gains for individuals hospitalized with these conditions. To do so, we expand our analysis period to track risk-adjusted survival from 1996 through 2008. The conventional wisdom is that the driving forces behind survival gains over this time period are a combination of 'high-tech' and 'low-tech' adoption decisions by hospitals (Cutler, 2005; Chandra and Skinner, 2012). But average survival gains can also come from reallocation of patients toward hospitals that achieve better outcomes. We investigate the extent to which the observed growth in average survival can be attributed to reallocation to higher quality hospitals as opposed to quality improvements within hospitals.

We use the approach of Foster, Haltiwanger and Krizan (2001) and Foster, Haltiwanger and Syverson (2008), which is itself a modification of the decomposition first derived in Baily et al. (1992). Specifically, we decompose the change in the average risk-adjusted 30-day survival in a market as follows:

[17]The results are qualitatively similar for the other conditions, though magnitudes are often smaller – indicating that for these conditions, patients are somewhat less willing to travel additional distance for a given increment in quality.

$$
\begin{aligned}
(7) \qquad \Delta\bar{q}_t \;=\; & \underbrace{\sum_{h\in C_t} \theta_{h,t-1}\Delta q_{h,t}}_{\text{within}} + \underbrace{\sum_{h\in C_t} \left(q_{h,t-1} - \bar{q}_{t-1}\right)\Delta\theta_{h,t}}_{\text{between}} + \underbrace{\sum_{h\in C_t} \Delta q_{h,t}\Delta\theta_{h,t}}_{\text{cross}} \\
& + \underbrace{\sum_{h\in M_t} \theta_{h,t}\left(q_{h,t} - \bar{q}_{t-1}\right)}_{\text{entry}} - \underbrace{\sum_{h\in X_t} \theta_{h,t-1}\left(q_{h,t-1} - \bar{q}_{t-1}\right)}_{\text{exit}}
\end{aligned}
$$

where $\bar{q}_t$ is the market-share-weighted average 30-day survival across hospitals in the market in year t, and Δ is the difference operator; thus the left-hand side is the change in weighted average patient survival in the market between two periods. On the right hand side, $q_{h,t}$ is the risk-adjusted survival rate for hospital h in year t and $\theta_{h,t}$ is its market share, i.e. the share of patients in the market with the condition who were treated at that hospital. C_t is the set of hospitals that were open in both $t-1$ and t; M_t is the set of hospitals that entered the market in t, and X_t is the set of hospitals that exited the market between $t-1$ and t.[18]

The decomposition in equation (7) divides average survival growth into five terms. The first term, "within", reflects changes in average survival in the market due to survival improvements among continuing hospitals holding their market shares constant. These are the survival gains that would have been attained in absence of any reallocation. The remaining terms reflect reallocation effects on average survival in the market. The second term, "between", shows how much of the rise is due to patients reallocating to hospitals that were already of high quality. The middle term, "cross", captures the covariance between gains in survival and gains in market share; it indicates whether hospitals that raised quality also grew their patient loads. The final two terms are, respectively, gains in survival due to entering hospitals having better performance than the previous average and the gains due to lower than average performance hospitals exiting. Any of these five terms could of course be negative if changes in survival, shares, or the composition of hospitals were such as to detract from average survival rates in the market.

[18] Because measurement error in risk-adjusted survival does not cause bias in any of the terms of this decomposition, we do not empirical-Bayes-adjust the survival measure when computing these metrics.

We look at the long difference between $t = 2008$ (our baseline period) and $t - 1 = 1996$. To do this, we replicated our sample selection and risk-adjusted survival measure in the 1996 data.[19] Thus $\Delta\bar{q}_t$ represents the change in 30-day survival between 2008 and 1996 for the market.[20] After conducting the decomposition for each market, we average each component over all markets weighting by the initial number of patients in the market in 1996. The resulting averages reflect the extent to which each of the five components accounted for survival gains over the 12-year period.

Table 6 displays the results. The first row shows the substantial secular improvement in survival gains for individuals hospitalized with these conditions. Average 30-day survival increased across all HRRs by 3.9 percentage points for AMI, 0.9 percentage points for heart failure, and 3.2 percentage points for pneumonia. The academic literature has focused on progress in AMI-survival – presumably because it is most dramatic – and attributed the improvements to technological progress. The literature has credited medically intensive interventions such as stents and reperfusion therapy and low-cost medical interventions such as aspirin and β blockers (Fibrinolytic Therapy Trialists' Collaborative Group, 1994; Keeley and Hillis, 2007; Chandra and Skinner, 2012).

[19] Specifically, just as in our baseline analysis for 2008 we measure risk-adjusted survival using 2006-2008 data and market share with 2008 patient counts, so for our 1996 analysis we measure risk-adjusted survival using 1994-1996 data and market share with 1996 patient counts. We use the same approach to define the sample and implement the risk-adjustment for the 1996 analysis as for the 2008 analysis.

[20] We centered the risk-adjusted survival measure $q_{h,t}$ in each market so that its market-share weighted average $\bar{q}_t$ equalled that of the raw, market-year average survival measure.

Table 6 - Decomposition of Gains in Survival Over Time

	(1)	(2)	(3)	(4)	(5)	(6)
	Contributions in Pctage Points			Contributions as Share of Total		
Condition	AMI	HF	Pneu	AMI	HF	Pneu
Total Change in Wtd Survival	0.0389	0.0092	0.0316	1.00	1.00	1.00
Within	0.0298	0.0076	0.0297	0.77	0.82	0.94
Between	0.0000	-0.0004	-0.0004	0.00	-0.05	-0.01
Cross	0.0062	0.0015	0.0015	0.16	0.16	0.05
Entry	0.0021	0.0006	0.0010	0.05	0.07	0.03
Exit	-0.0007	0.0000	0.0002	-0.02	0.01	0.01

This table decomposes the gains in risk-adjusted survival for the emergent conditions over our full sample window (between 1996 and 2008) using the decomposition shown in equation (7). Columns 1-3 show the contribution of each component to gains in survival in percentage points (e.g. a value of 0.1 is 10 percentage points). Columns 4-6 show the share of total gains that can be attributed to each component. The exit component enters negatively, so a negative value indicates that exit accounts for a gain in survival.
The decomposition is performed for each market, then averaged together weighted by the market's size in 1996. Risk-adjusted survival is calculated from a regression of survival on hospital fixed effects and patient risk-adjusters. A separate regression is run for each of the year groups 1994-1996 (yielding the 1996 survival rates) and 2006-2008 (yielding the 2008 survival rates).

Consistent with this conventional wisdom, we find that within-hospital upgrading in quality accounts for the bulk of the AMI improvements, explaining 77 percent of gains in risk-adjusted AMI survival over our 12 year period. However, we also find a quantitatively important role for reallocation; 23 percent of the secular improvement in AMI survival can be explained by reallocation of patient flows toward higher quality hospitals. The cross term explains the bulk of the reallocation gains, contributing 0.6 percentage points (16 percent of the total gains). Entry of high-performance hospitals can explain about 0.2 percentage points (5 percent) of survival gains, and exit of low-performance hospitals can explain about 0.1 percentage points (about 2 percent) of survival gains.[21]

[21]Due to our sample constructions, "exit" and "entry" need not be literal hospital entry and exit. They also reflect a reduction in condition-specific sample size below the inclusion cutoff (specifically, at least 25 patients in the three year period used to estimate risk-adjusted survival and at least 1 patient in the final year; see Section III.B).

To put the role of reallocation in AMI survival gains in perspective, it is instructive to note that the nearly 1 percentage point improvement in AMI survival over the 1996-2008 period that we attribute to reallocation is about half of the magnitude of the survival gains attributed to each of two major breakthroughs in AMI treatment: reperfusion and primary angioplasty. Reperfusion (including e.g. fibrinolytics) started being widely used in the early 1990s and has been estimated to raise 30-day survival by 2 percentage points (Fibrinolytic Therapy Trialists' Collaborative Group, 1994). Primary angioplasty diffused over the 1990s and has been estimated to increase 30-day AMI survival by 2 percentage points over reperfusion therapy (Stone, 2008; Keeley, Boura and Grines, 2003).

For heart failure and pneumonia – where the secular improvements in survival are noticeably smaller – we find a somewhat smaller contribution of reallocation. Table 6 indicates that reallocation can explain about 18 percent of the 0.9 percentage point secular improvement in heart failure survival and about 6 percent of the 3.2 percentage point improvement for pneumonia.

Using our long survival panel to explore whether the relationship between market allocation and hospital performance has evolved over time, we find that the alignment of market allocation with hospital performance appears to have increased since 1996.[22] This could be because information about hospital quality and hospital outcomes has become more available to patients and their surrogates. The study sample corresponds to a period in which tools like CMS Hospital Compare made it much easier for patients to ascertain the quality of a hospital, and research on the impact of report cards has found some evidence of consumers responding to information therein (Dranove et al., 2003; Dranove and Sfekas, 2008). Alternatively, patients' willingness or ability to switch among hospitals may have risen for other reasons. For example, the consolidation of either insurers or provider groups may have induced consumers to reallocate to higher quality providers. Clearly, these

[22] Appendix Table A13 shows these results. We report the static and dynamic allocation analysis of equations (1) and (2) for five separate periods: 1996, 1999, 2002, 2005, and 2008. We find that allocation tends to become more directed toward high-survival hospitals over the sample. AMI has the cleanest such pattern; the magnitudes of both static and dynamic allocation increase monotonically over the sample. Heart failure and pneumonia are noisier, especially for the dynamic allocation, but their overall trends are in the same direction.

channels are speculative; we do not have the necessary data to pin down the mechanism in this study, but we see this as a natural and interesting area for future work.

IV.C Private vs. Social Preferences

Thus far we have examined the correlation of market share with a variety of "quality" and "output" metrics, with no attention to hospital inputs or costs. Uniquely in the health care sector relative to the rest of the economy, consumers absorb little to none of the costs of their hospital choice. The Medicare patients we analyze all have insurance with limited cost-sharing, much of which in turn is covered by supplemental (public or private) coverage (MEDPAC, 2012). As a result, an output-based quality metric is likely the relevant one from the perspective of consumer demand; we would not expect patients or their surrogates (family members, physicians, etc.) to substitute away from high cost hospitals.

A benevolent social planner, on the other hand, would want to allocate toward hospitals with high quality (output) *conditional* on inputs. How the social planner would trade off higher output at the cost of higher inputs would, of course, depend on the the social welfare function. We therefore analyze how allocation correlates with productivity, which we define as the hospital's ability to generate survival conditional on the inputs it uses in the treatment process. Because the social planner is concerned about conditional survival while the patient values unconditional survival, there may be a wedge between the socially and privately optimal hospital choice.

To analyze whether the market re-allocates toward some measure of input-adjusted outcomes, we define a patient-level health production function of the following form:

$$Y_p^s = A_h \left(\prod_k R_{pk}^{\lambda_k} \right) X_p^{\mu} \Xi_p. \tag{8}$$

where the leading term, A_h, measures the exponential of total factor productivity (TFP) of hospital h; Y_p^s is the output generated by the hospital in treating patient

p; and X_p is a measure of hospital inputs used to treat the patient. All production functions relate outputs to inputs; our particular function uses 30-day survival as a measure of output (in fact, Y_p^s is technically the exponential of this indicator) and a single index of resources spent on the patient as inputs.[23] The parameter μ is the elasticity of 30-day survival with respect to risk-adjusted inputs. Because patients are inherently heterogeneous, survival may also depend on characteristics of the patient, which could potentially also be correlated with input choices. In addition, the marginal effect of inputs on survival may vary with patient characteristics. To capture both of these effects, we follow the literature and adjust inputs for a vector of observable patient-level risk factors, R_{pk}, where k indexes the factors. The risk factors are the same as those used in the calculation of risk-adjusted survival described in Appendix B. Finally, the expression Ξ_p is a patient-level error term that accounts for random variations in health outcomes.

The hospital production function model in equation (8) allows variation across hospitals in the marginal health product of inputs but constrains hospitals to have the same elasticity of output with respect to input (i.e., μ is common across hospitals). Our empirical specification therefore allows the "marginal return to inputs" curve to vary across hospitals, as suggested by Chandra and Staiger (2007) and Garber and Skinner (2008).

Taking logs and using lowercase letters to represent the logarithm of uppercase letters yields our estimating equation for the hospital production function:

$$y_p^s = a_h + \sum_k \lambda_k r_{pk} + \mu x_p + \xi_p \tag{9}$$

This equation is identical to what we use to estimate risk-adjusted survival (see Appendix B for more details) with one key difference: it controls for logged hospital inputs x_p. Hospital productivity a_h is risk- and resources-adjusted survival – or equivalently, hospital output conditional on inputs (which include patient health inputs, i.e. the risk-adjusters we used previously to construct risk-adjusted survival,

[23] This sort of single-input production function is unusual but convenient; one could reasonably interpret the single input as an index of the use of multiple inputs that go into producing health.

as well as resource inputs).

A key challenge in calculating productivity is constructing an appropriate measure of hospital inputs. These inputs are the resources utilized in the treatment of the patient – labor inputs like physicians and nurses and capital inputs like the operating theater and diagnostic scanners. Measuring inputs is challenging in most industries, and healthcare is no exception. In practice, we consider two imperfect measures of these inputs.

Our first input measure is federal expenditures, i.e. total Medicare dollar payments to hospitals for inpatient services used in the treatment of the patient during the first 30 days following the admission. Medicare pays hospitals for each patient stay; its reimbursement for a hospital stay is based on the diagnosis causing the admission, whether the patient has other complicating medical conditions, the geographic location of the hospital, the type of hospital (e.g. whether it is an academic medical center) and, to some extent, what is done to the patient in the hospital (MedPAC, 2011). For example, within a given hospital, Medicare reimbursement for an AMI admission will vary depending on the presence of a complicating condition like heart failure or stroke and whether the patient receives various intensive treatments such as a bypass operation; Medicare spending over the 30 days following the index admission will also depend on whether the patient has multiple hospital stays, as each stay triggers an additional Medicare payment.

Using federal expenditures as our input measure allows us to construct a measure of output (survival) per dollar of federal expenditures. Given the social cost of raising public funds, this is a natural and useful productivity metric. However, it has the disadvantage that it captures variation in inputs coming both from hospital-specific prices and CMS's estimates of the real resources used by the hospital for treatment. Our second input measure addresses this concern by purging the federal spending measure of pricing variation to create a "resource" measure of inputs. Specifically, we define hospital inputs for a patient as the sum of diagnostic-related group (or DRG) weights during the first 30 days following a heart attack. These DRG weights reflect CMS's assessment of the resources necessary to treat a patient as a function of the patient's medical conditions and procedures received. This approach is standard in the literature as a way of purging measures of care utilization

of administrative price variation (see e.g. Skinner and Staiger, 2015; Gottlieb et al., 2010). Nonetheless, it is a highly imperfect measure of inputs, as it does not reflect actual inputs used but rather CMS-defined expected inputs based on the treatment approach chosen.

We limit our analysis to AMIs. We do this because for the other conditions, the vast majority of patients fall into just one or two DRGs. As a result, for these conditions, there is little useful variation in the input measures for us to exploit.

We estimate the hospital production function in equation (9) for each input measure using the same sample and risk-adjusters as we used to estimate risk-adjusted survival (described in Appendix B).[24] Having estimated the production function models for each input measure, we then extract the hospital fixed effects to create our estimates of productivity.

Table 7 examines static and dynamic allocation of AMI patients as a function of these hospital productivity measures (i.e. the a_h estimated in equation 9).[25] Columns 1 and 2 consider static allocation with respect to productivity by the "federal dollars" and "real resources" metrics, respectively. Both results imply that a hospital that can generate 1 percentage point greater survival holding inputs constant (i.e. has 1 percentage point greater TFP) is expected to be about 18 percent larger than other hospitals in its market. Columns 5 and 6 repeat this analysis looking at dynamic allocation; for both measures, we find that having 1 percentage point greater TFP is associated with 1.5 percentage points higher growth in patients between 2008 and 2010, compared to other hospitals in the market.

[24]Estimating this model with the federal dollars inputs measure yields $\hat{\mu}$=0.035 (standard error = 0.001). In other words, a 10 percent higher payment to the hospital for treating an AMI patient is associated with a 0.35 percentage point increase (0.4 percent of the hospital average survival rate of 0.82) in the probability of post-AMI 30-day survival. Using the resource-based input measure to estimate the hospital production function yields $\hat{\mu}$=0.042 (standard error = 0.001); increasing resources on a patent by 10 percent is associated with a 0.42 percentage point (0.5 percent) increase in the probability of post-AMI 30 day survival.

[25]We empirical-Bayes-adjust the hospital productivity objects when using them in the regressions of Table 7 so that our coefficients are not biased due to measurement error. The empirical Bayes approach is described in Appendix C and is identical to our approach for risk-adjusted survival and readmission.

Table 7 - Allocation of AMI with Respect to AMI Productivity and its Components

Measure	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation for AMI				Dynamic Allocation for AMI			
Productivity (Fed $)	17.637				1.491			
	(1.118)				(0.420)			
Productivity (Resources)		17.540				1.471		
		(1.013)				(0.386)		
Risk-Adjusted ln(Fed $)			1.447				0.246	
			(0.169)				(0.064)	
Risk-Adjusted ln(Resources)				0.620				0.468
				(0.406)				(0.162)
Risk-Adjusted Survival			17.940	19.789			1.479	1.559
			(1.192)	(1.297)			(0.446)	(0.441)
Hospitals	2,890	2,890	2,890	2,890	2,890	2,890	2,890	2,890

This table extends the analysis of Table 4 but is limited to static and dynamic allocation for AMI. It shows how allocation is related to AMI productivity or its two components (risk-adjusted survival and risk-adjusted log-inputs). Productivity is defined as risk- and inputs-adjusted survival; see Section 4.3 and equation (9). We consider two input measures, "federal expenditures" and "resources", also defined in the text. Standard errors are bootstrapped with 300 replications and are clustered at the market level.
The standard deviation of productivity is 0.03 (Fed $ or Resources), of risk-adjusted log-inputs is 0.22 (Fed $) and 0.07 (Resources), and of risk-adjusted survival is 0.04 — this number differs from that of Table 2 because it comes from estimating the joint distribution of survival and inputs, not survival alone.

These results show that market reallocation occurs in the direction of higher productivity hospitals. However, as noted, if patients or their surrogates are making the allocation decision, they should not have reason to penalize hospitals which, conditional on output, are high input utilizers. In other words, they value health outcomes, not input-adjusted outcomes or "productivity". The social planner, however, would penalize hospitals that, conditional on health outcomes, had higher inputs (costs). We therefore examined the (conditional) correlation between allocation and risk-adjusted survival and risk-adjusted inputs.[26]

The results – shown in the remaining columns of Table 7 – are consistent with the social planner's goal of rewarding hospitals with high performance after adjust-

[26]Risk-adjusted input use is defined as the hospital fixed effect from the regression of log-inputs on the patient risk-adjusters and hospital fixed effects. Risk-adjusted survival is defined as it was previously (see Appendix B). As with our productivity estimates, we use an empirical Bayes correction to adjust our estimates of risk-adjusted survival and of risk-adjusted inputs for measurement error; our procedure, described in Appendix Section C.4, accounts for the correlation in measurement error between these two objects.

ing for inputs. However, the market does not reallocate away from hospitals with higher input use conditional on survival as the social planner would want.[27] Put another way, higher productivity hospitals do treat more patients and grow faster, but this is a result of the coincidence of high productivity with high output or quality (which our results indicate patients do respond to) rather than patients systematically avoiding hospitals that use a lot of inputs to achieve high output.

This result illustrates the divergence between the goals of the social planner and the consumer in a market where consumers face few costs. Indeed, it appears that, conditional on output, higher inputs are also associated with greater patient flow, although the magnitude of reallocation based on inputs is substantially smaller than that based on output.[28] It may be the case that consumers (correctly or not) view inputs conditional on outputs as a signal of unobserved quality, or that they have preferences for high intensity care regardless of its medical value, as has been conjectured to have occurred in the past when hospitals competed on technology in a "medical arms race" (e.g. Kessler and McClellan, 2000). With nearly full insurance for hospital treatment, there is no demand-side force in the market for AMI treatment to counteract these preferences and align the consumer's preferences with those of the social planner. To push back against this consumer insensitivity, in 2015 Medicare began penalizing hospitals with high costs by adding per-beneficiary spending to the measures it uses to adjust payments in its Value-Based Purchasing

[27] Columns 3, 4, 7 and 8 show that the market reallocates both to hospitals with higher survival (conditional on inputs) and higher inputs (conditional on survival). The coefficient on productivity can remain positive when used as a single index (columns 1, 2, 5, and 6) because the coefficients on risk-adjusted inputs are smaller than those on risk-adjusted survival and because risk-adjusted inputs are a smaller component of variation in productivity than risk-adjusted survival, since they enter after being multiplied by $\mu \ll 1$.

[28] For example, in the static allocation analysis, column 4 indicates that a 1 standard deviation increase in outputs (i.e. risk-adjusted survival) is associated with a 69 percent rise in hospital size, while a 1 standard deviation increase in real resources-based inputs is associated with just a 5 percent rise in size.

In the static analysis, we find substantially larger reallocation toward inputs measured by federal dollars (column 3) than real resources (column 4), but this difference is an artifact of the CMS payment approach; the Medicare payment formula explicitly pays academic medical centers and safety net hospitals more than other hospitals for the same patient, and these facilities tend to be large, which magnifies the correlation between federal dollars and size. The "real resources" input measure accounts for this concern by using the same DRG weight (our measure of inputs) across all hospitals for each type of case.

Program (Kahn et al., 2015; QualityNet, 2015).

V Mechanism

The previous section showed, through multiple quantitative lenses, evidence of reallocation to higher quality hospitals. A natural concern with these findings is that they purport to show the results of a reallocation process without giving evidence of a mechanism by which such reallocation can occur. In this section we provide additional evidence consistent with a demand-based mechanism and investigate leading alternative explanations.

V.A Evidence of demand-based mechanism

We view the need to find additional signposts of demand-driven allocation as particularly important given that intuition suggests that there may be little scope for patient choice. In particular, for our three emergency conditions, AMI, heart failure and pneumonia, about three-quarters of admissions come via the emergency department. Relatedly, there is a long tradition in health economics – dating back at least to McClellan, McNeil and Newhouse (1994)'s analysis of hospital choice by AMI patients – of using the distance from the patient's residence to the nearest hospital as an instrument for which hospital the patient goes to. How, then, is it plausible that patients' (or their surrogates') demand is playing an allocative role in such emergency situations?

As discussed in Section II and Appendix A, a key comparative static of a demand-based allocation mechanism is that as consumer's ability or willingness to substitute across producers increases, the static and dynamic allocation results should grow stronger. To investigate this prediction, we segmented our patients within each condition into two groups with arguably different scope for exercising choice over hospitals. We based the division on the manner in which the patients were admitted as inpatients: through the emergency department (ED) or as a non-emergency transfer from another hospital.[29]

[29]The source of admission is determined by whether the hospital reported positive charges for use

AMI is a condition that demands immediate treatment, and most AMI patients are admitted through the hospital's ED. However, a small but still substantial subset of AMI patients arrive at the hospital after being initially treated and potentially stabilized at another facility. The practice of transferring certain patients to facilities with more advanced treatment capabilities is recommended by the American Heart Association (AHA; see Neumar et al., 2015 and O'Connor et al., 2015). For example, the AHA recommends that hospitals unable to perform angioplasty transfer patients whose heart attacks are appropriate for such treatment to facilities able to perform it. In a typical case, the patient presents at the initial hospital's ED, a clinician confirms that treatment of the AMI would be improved by transfer to another facility, and the patient is then transported and directly admitted to the receiving facility.

We identify transfer patients as those whose index inpatient stay did not require an admission through the ED and which occurred immediately after an encounter at another hospital; in Table 8 we show that transfers comprise 16 percent of the AMI sample – about two-thirds of the AMI patients who are not admitted to the hospital through the ED.[30] To confirm that our findings are not specific to AMI, we undertake the same breakdown for heart failure and pneumonia patients, although the clinical rationale behind these transfers is less standardized and, relatedly, we have a far smaller share of transfer patients for these conditions. There is effectively no variation in the source of admission for hip and knee replacements – 98 percent are non-transfer, non-ED admissions – so we exclude them from the analysis.

---

of the ED.

[30]Immediately prior to their index stay, transfer AMI patients had an ED encounter at another facility for any reason (including but not limited to AMI) or an inpatient encounter at another facility that was not for AMI (for example, chest pain). Patients are indexed to the first inpatient hospitalization that indicated AMI as the underlying cause of their admission, and hospitalizations that occur within 1 year of a prior AMI stay are not counted as index events. For these reasons, a given patient's episode of care cannot enter both the ED and the transfer samples.

The excluded category of patients from this analysis are the one-third of non-ED AMI patients who are not transfers – Appendix Table A14 shows allocation relationships for these patients, which do tend to be positive. These patients may have been directly admitted to the hospital upon physicians' orders, were brought to the hospital for some other reason and subsequently experienced an AMI during their stay, or were miscoded in the data. We limit our analysis to transfer patients instead of the full set of non-ED patients because transfer patients more clearly select hospitals for the purpose of AMI treatment.

Table 8 - Travel Distance for ED and Non-ED Transfer Patients across Conditions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Condition	AMI		Heart Failure		Pneumonia		Hip/Knee	
Source of admission	ED	Transfer	ED	Transfer	ED	Transfer	ED	Transfer
Share of patients in 2008	0.76	0.16	0.75	0.03	0.77	0.01	0.02	0.00
Median miles traveled	5.43	33.77	5.06	30.15	5.12	25.20	5.88	29.18
Mean miles traveled	40.89	66.45	34.27	61.38	37.65	52.21	44.24	61.01
Share treated at nearest hospital	0.52	0.03	0.53	0.07	0.54	0.12	0.50	0.14

This analysis considers 2008 patients who were treated at hospitals in the baseline allocation sample (i.e. hospitals with at least 1 patient in 2008). ED patients were admitted through the hospital's emergency department (see footnote 29). Non-ED transfer patients were admitted directly after a stay at another hospital (see footnote 30). Distances are from the ZIP code centroid of the patient to the ZIP code centroid of the hospital.

Table 8 contains two important findings. First, even among patients admitted through the ED – where both intuitively and according to the statistics on distance travelled there is less scope for choice – only half are treated at the nearest hospital. This fact helps illustrate that a demand-based mechanism can exist for these patients, even if their ability to choose a hospital is more constrained. Some decision-maker, whether that is the patient, his family, his doctor, or the ambulance driver, is still exercising active choice for a large share of ED admissions. In other words, the hospital distance instrument of McClellan, McNeil and Newhouse (1994) may be predictive, but does not have an R^2 of 1; there may be scope for choice even among patients admitted via the ED.

Second, Table 8 provides empirical evidence consistent with our motivating assumption that, within a condition, transfer patients have greater scope for choice than ED patients. The table shows distance traveled by the patient to the hospital, with distance defined as the number of miles between the patient's ZIP code of residence and the hospital's ZIP code. Within a condition, transfer patients travel significantly farther for treatment. For example, the median AMI patient admitted through the ED is treated about 5 miles from the patient's home, while the median AMI transfer patient is treated 34 miles from his home. Relatedly, about 52 percent of AMI patients admitted through the ED are treated at the hospital nearest to them, compared to only 3 percent of AMI transfer patients. The travel distance differentials are similar across ED and transfer patients for the other two emergent

conditions.

Motivated by this empirical corroboration that transfer patients have more scope for choice than ED patients, Table 9 repeats our static and dynamic allocation analysis separately for ED patients and non-ED transfer patients. While the quality metrics are the same as in our baseline analysis of Table 4, to construct the left hand side allocation measures, we use only ED patients or only non-ED transfer patients.[31] The results in Table 9 are consistent with a demand-based mechanism: within a condition, where there is more scope for patient choice – both intuitively and according to the data – the static and dynamic allocation results are substantially and statistically significantly larger.[32]

[31] One potential concern with the static analysis is selection on the dependent variable; the transfer sample is much smaller than the ED sample and transfers appear in roughly one-third to one-half of analysis sample hospitals. We therefore use a fixed effects Poisson model in this section to avoid dropping zeroes, keeping a common sample within-condition between the two patient groups. The Poisson model coefficients have the same interpretation as our baseline linear regression coefficients (Wooldridge, 2002). Our baseline allocation results are not sensitive to the choice of model (see Appendix Table A9) and our ED vs. transfer results remain even under the linear model, though as expected they are somewhat attenuated due to differential truncation (see Appendix Table A15). Furthermore, Appendix Table A16 shows that we find the same pattern of results if instead we replicate the conditional logit analysis of hospital choice in equation (4) separately for patients admitted through the ED and patients admitted as transfers.

[32] For completeness, we include the patient survey results although we found no allocation toward higher quality by this metric. The patient survey continues to be negatively correlated with static allocation for the three conditions, though this relationship is often weak. Its association with growth in patient loads is mostly positive but imprecisely measured; we find no evidence that this gradient differs between ED and transfer patients.

Table 9 - Allocation for ED and Non-ED Transfer Patients across Conditions

	(1)	(2)	(3)	(4)	(5)	(6)
Condition	AMI		Heart Failure		Pneumonia	
Source of admission	ED	Transfer	ED	Transfer	ED	Transfer
Share of patients in 2008	0.76	0.16	0.75	0.03	0.77	0.01
Risk-Adjusted Survival						
Static Allocation	14.489	42.532	15.727	50.673	7.168	14.049
	(1.022)	(2.609)	(1.586)	(4.664)	(0.983)	(2.941)
P-value of test for equality	0.000		0.000		0.009	
Hospitals	2,881	2,881	4,023	4,011	4,325	4,275
Dynamic Allocation	0.572	7.258	2.300	13.935	3.423	4.454
	(0.496)	(1.260)	(0.799)	(2.635)	(1.006)	(1.793)
P-value of test for equality	0.000		0.000		0.562	
Hospitals	1,384	1,384	1,438	1,438	1,451	1,451
Risk-Adjusted Readmission						
Static Allocation	-8.128	-25.550	-11.265	-37.988	-1.647	1.089
	(1.730)	(5.921)	(2.329)	(6.744)	(2.021)	(6.252)
P-value of test for equality	0.001		0.000		0.653	
Hospitals	2,304	2,304	3,903	3,892	4,264	4,214
Dynamic Allocation	0.148	-1.837	-1.812	-11.381	-0.439	-2.517
	(0.704)	(2.034)	(0.994)	(2.727)	(1.419)	(2.965)
P-value of test for equality	0.285		0.001		0.500	
Hospitals	1,342	1,342	1,434	1,434	1,445	1,445
Process of Care Z-Score						
Static Allocation	0.326	1.179	0.377	0.754	0.262	0.214
	(0.021)	(0.090)	(0.025)	(0.058)	(0.018)	(0.043)
P-value of test for equality	0.000		0.000		0.261	
Hospitals	2,379	2,379	3,665	3,653	3,920	3,869
Dynamic Allocation	0.008	0.216	0.063	0.295	0.079	0.098
	(0.021)	(0.045)	(0.027)	(0.053)	(0.024)	(0.042)
P-value of test for equality	0.000		0.000		0.638	
Hospitals	1,360	1,360	1,433	1,433	1,428	1,428
Patient Survey Z-Score						
Static Allocation	-0.157	-0.034	-0.141	-0.090	-0.137	-0.203
	(0.035)	(0.072)	(0.032)	(0.060)	(0.028)	(0.057)
P-value of test for equality	0.051		0.349		0.170	
Hospitals	3,498	3,498	3,598	3,586	3,610	3,559
Dynamic Allocation	0.019	0.052	0.038	0.012	-0.002	0.008
	(0.020)	(0.037)	(0.018)	(0.052)	(0.023)	(0.051)
P-value of test for equality	0.430		0.600		0.851	
Hospitals	1,397	1,397	1,423	1,423	1,396	1,396

This table repeats the analysis of Table 4, but the left hand side of these regressions considers hospital size and growth counting only ED patients in the odd-numbered columns and only non-ED transferred patients in the even-numbered columns. We use a fixed effects Poisson model for static allocation to avoid differential truncation between the two patient groups. To make the Poisson model analogous to our baseline static allocation model, its regressand is the count of patients, not the logarithm.
The static allocation sample is the baseline analysis sample from Table 4. Poisson sample sizes may differ from baseline sample sizes because markets with one hospital or markets with all zero patient counts are excluded. The dynamic allocation sample is the subset of baseline hospitals with at least one ED patient and non-ED transfer patient in 2008. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Focusing on AMI, there is a consistent pattern that the gradient between quality and market share is larger for the transfer patients than ED patients both at a point in time (static allocation) and over time (dynamic allocation). For the condition-specific measures, the differences are statistically significant in five of the six comparisons. For example, a hospital with 1 percentage point greater risk-adjusted survival will tend to treat 14 percent more ED patients than other hospitals in its market, but the gradient is even greater for transfers, where 1 percentage point higher risk-adjusted survival is associated with 43 percent more of these patients. Moreover, a hospital with 1 percentage point greater risk adjusted survival will tend to grow its ED patient load by 0.5 percentage points over the next 3 years, but its transfer patient load will tend to grow by 7 percentage points.

As noted, the share of transfer patients is much smaller for heart failure and even more so for pneumonia, since transfers are less central to the standard treatment protocol for these conditions. Nonetheless, the higher gradient between quality and market share for transfer patients than ED patients continues to hold for heart failure, both in the static and dynamic analysis. For pneumonia, the results are largely insignificantly different between ED and transfer patients, though the lack of transfer patients makes accurately measuring allocation difficult for this population.

These findings are consistent with an allocation mechanism based on demand by patients or their surrogates. However, they fall short of showing the specific pathway by which patients know which hospitals offer high quality. This ambiguity is not unique to our study. Indeed, a long-standing question in the field – dating back at least to Arrow (1963) – is how patients can acquire information on provider quality. One possibility is some form of market-learning; hospitals acquire a reputation for good outcomes and this reputation spreads through physicians' professional networks and patients' social networks, where it influences patients and their surrogates to request treatment at hospitals that are better at producing survival. Indeed, in a related setting, Johnson (2011) finds that cardiac specialists who have higher risk-adjusted survival rates for their patients are less likely to stop practicing. She interprets this and related evidence as consistent with a model of market learning by the referring physician. Patients or their family members may also obtain such information themselves; there is some evidence, for example, that patients respond

to provider report cards (e.g., Dranove et al., 2003 and Dranove and Sfekas, 2008).

Another possibility is that choice of hospital is not based on our hospital quality measures per se, but rather on other characteristics that are correlated with hospital quality. One potential scenario, backed by clinical guidelines for AMI, is that patients or their surrogates seek structural features of the hospital like cardiac catheterization labs which they believe will directly improve the hospital's quality. It is also possible that higher quality hospitals have better non-health amenities like nicer lobbies, which in turn influences hospital demand (though our null results with respect to the patient survey would seem to lean against this explanation). We take no stand on whether consumers actively seek quality itself, seek features of the hospital that generate quality, or seek attributes that correlate but do not cause quality. All such explanations imply a role for patient demand in causing the healthcare sector to re-allocate toward higher quality producers, although they naturally have very different counterfactual implications.

V.B Alternative Explanations

We consider two key alternative, non-demand based explanations for the allocation results. The first is reverse causality: with increasing returns to scale, causality runs from hospital scale to hospital quality, rather than vice versa. This is a general issue for interpreting the static allocation measure in any industry. In the particular context of health care, the "volume-outcome" hypothesis conjectures that treating more patients improves provider performance. Not surprisingly, it has proven challenging to establish empirically whether an observed positive correlation between provider volume and outcomes is causal (see e.g. Epstein, 2002 for a discussion of the interpretation difficulties in this literature).

The volume-outcomes hypothesis is unable to explain the totality of our results. First, the hypothesis alone provides no prediction for dynamic allocation: in an environment with increasing returns to scale but without quality-sensitive consumers, better hospitals will tend to be bigger at a point in time, but they do not necessarily grow more over time. Additional assumptions would be required for the presence of a volume-outcome relationship to explain the dynamic allocation results. For ex-

ample, if hospital-level changes in patient flows were autocorrelated and there were strong returns to scale, an innovation in patient traffic before the base period would yield both high performance in the base period (from the scale relationship) and further growth in patients (from the autocorrelation). Second, even if a more nuanced scenario like this one were operative, it would not explain why patients who have more scope for choice are systematically more likely to choose high quality hospitals. Since we find in Table 9 that transfer patients are much more likely to present at high quality facilities than ED patients, even if there are increasing returns to scale, patient preferences appear to drive at least some of the observed allocation.

A second non-demand-based explanation for our findings is a "mechanical" one in which patients simply go to the nearest hospital without considering performance at all. With mechanical assignment of patients to the nearest hospital, our static and dynamic allocation results could be produced spuriously if, for example, within a market, more densely populated (e.g. urban) areas have both higher quality hospitals and faster population growth. Mechanical assignment of many patients to hospitals based on proximity seems a particularly natural alternative given the famous McClellan, McNeil and Newhouse (1994) use of distance as an instrumental variable for which hospital treats a given AMI patient.

In practice, however, this type of strict mechanical allocation rule does not seem able to explain our findings. We produce a counterfactual allocation of patients based on this mechanical allocation rule. Specifically, we assign each patient to his nearest hospital instead of the one at which we observe treatment. This approach substantially alters the allocation of patients across hospitals, since across the four conditions, 44 percent to 62 percent go to a hospital that is *not* the closest one to them (see Appendix Table A3). Importantly, under this counterfactual allocation, our static and dynamic results for the condition-specific quality measures either substantially attenuate or reverse.[33]

[33] We present these results in Appendix Tables A17 and A18. Consider, for example, the static and dynamic allocation results for AMI (column 2 of each table). Compared to the baseline results (column 1 of each table), the static allocation coefficients attenuate by a factor of 8 for risk-adjusted survival and attenuate to statistical insignificance for risk-adjusted readmission and processes of care. For dynamic allocation, the outcomes and process coefficients all attenuate to insignificance. The patient survey continues to have a negative correlation with static and dynamic allocation, though the coefficients attenuate by about one-third.

VI Conclusion

This paper has examined the relationship between firm performance and market allocation in healthcare, focusing specifically on hospital treatment of Medicare patients. We examine three conditions and a pair of common surgical procedures – AMI, heart failure, pneumonia, and hip and knee replacement – that together account for almost one-fifth of Medicare hospitalizations and hospital spending. For all of these conditions, we find robust evidence that higher quality hospitals – as measured by risk-adjusted survival, risk-adjusted readmissions, and adherence to well-established clinical practice guidelines – tend to attract greater market share at a point in time and to grow more over time. The one exception to this favorable pattern of reallocation is that hospitals that score better on ex-post measures of patients' satisfaction with their experience (such as whether the room was quiet) do not attract greater market share; as we discussed, there is some debate in the literature over whether patient satisfaction scores are an informative measure of hospital quality (see e.g. Manary et al., 2013**).**

We provide several ways of quantifying these allocation results. For example, focusing on the AMI condition and the risk-adjusted survival quality metric, our estimates suggest that, within a market, hospitals with a 1 percentage point higher risk-adjusted AMI survival rate have a 17 percent higher market share at a point in time and a 1.5 percentage point higher growth in market share over the next 2 years. We estimate that AMI patients are willing to travel an additional 1.8 miles (about 0.1 standard deviation of distance traveled) to go to a hospital with a 1 percentage point higher risk-adjusted survival rate. Looking over the 1996-2008 period, we estimate that reallocation is responsible for about one quarter of the 3.9 percentage point gain in AMI survival, with the other three-quarters due to within-hospital quality improvements.

We present additional evidence that is consistent with a demand-based mechanism at work. In particular, we use a patient's source of admission to the hospital – as a transfer patient admitted after being stabilized elsewhere compared to a patient admitted through the emergency department – to classify patients who have more or less scope for hospital choice. We find a robust pattern that the relationship

between performance and market share, both at a point in time and over time, is stronger within a condition for patients who have greater scope for hospital choice.

While quality is clearly a predictor of allocation, patients continue to present at hospitals far from the quality frontier. This finding raises the question of what drives the apparent imperfect substitutability between facilities. The culprit could be any number of factors – for example, patients or their surrogates may have unobserved tastes for certain hospitals, or there may be information frictions that make it difficult for some patients to observe hospital performance. Patients may also have a pure disutility of travel, and in the case of emergent conditions, delays due to travel have further mortality consequences. Additional research will be helpful to tease apart the role of these various factors in the allocation process.

In addition, our estimated reallocation relationships stop short of indicating what economic or policy forces could be unleashed to create still greater reallocation to high quality producers. We see a great opportunity for further work that tries to estimate the causal impact of competition – or other factors – on allocation in healthcare. Another important but challenging area for future work is what policies could encourage reallocation in a manner consistent with the objectives of a benevolent social planner. Consumers face very little, if any, financial consequences of their hospital choice. Consistent with this, we find that while the market reallocates to higher quality hospitals, conditional on quality it does not reallocate away from higher cost hospitals. Naturally this fact has important normative implications for the social (rather than private) efficiency consequences of the allocation forces we observe. Perhaps in recognition of this wedge between the social planner and the patient, in 2015 Medicare began penalizing hospitals that have high costs (Kahn et al., 2015; QualityNet, 2015).

Taken together, our results suggest that healthcare may have more in common with "traditional" sectors than is commonly recognized in popular discussion and academic research. In this sense, our results are in the same spirit as Skinner and Staiger (2007)'s finding of a common "innovativeness factor" across healthcare and other sectors within a geographic area; they showed that areas of the country that were early adopters of hybrid corn in the 1930s and 1940s were also early adopters of β blockers for heart attacks at the beginning of the current century. Such findings

suggest that, going forward, research on the determinants of performance in the health care sector may benefit from more attention to the insights, both theoretical and empirical, from research about productivity and allocation in other industries. By the same token, the health care sector may likewise be a useful laboratory for insights about other industries.

References

Andersen, Henning R., Torsten T. Nielsen, Klaus Rasmussen, Leif Thuesen, Henning Kelbaek, Per Thayssen, Ulrik Abildgaard, Flemming Pedersen, Jan K. Madsen, Peer Grande, Anton B. Villadsen, Lars R. Krusell, Torben Haghfelt, Preben Lomholt, Steen E. Husted, Else Vigholt, Henrik K. Kjaergard, Leif Spange Mortensen, and DANAMI-2 Investigators. 2003. "A Comparison of Coronary Angioplasty with Fibrinolytic Therapy in Acute Myocardial Infarction." *The New England Journal of Medicine*, 349(8): 733–42.

Anderson, Gerard F., and Earl P. Steinberg. 1984. "Hospital Readmissions in the Medicare Population." *The New England Journal of Medicine*, 311(21): 1349–1353.

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *The American Economic Review*, 53(5): pp. 941–973.

Asplund, Marcus, and Volker Nocke. 2006. "Firm Turnover in Imperfectly Competitive Markets." *The Review of Economic Studies*, 73(2): 295–327.

Axon, R. Neal, and Mark V. Williams. 2011. "Hospital readmission as an accountability measure." *JAMA*, 305(5): 504–505.

Baily, Martin N., Charles Hulten, David Campbell, Timothy Bresnahan, and Richard E. Caves. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity, Microeconomics*, 1992: 187–267.

Bartelsman, Eric J., and Mark Doms. 2000. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature*, 38(3): 569–594.

Bartelsman, Eric J., John C. Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *The American Economic Review*, 103(1): 305–334.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen. 2012. "Management Practices Across Firms and Countries." *The Academy of Management Perspectives*, 26(1): 12–33.

Blumenthal, Daniel, and Anupam B Jena. 2013. "Hospital value-based purchasing." *Journal of Hospital Medicine*, 8(5): 271–7.

Boulding, William, Seth W Glickman, Matthew P Manary, Kevin A Schulman, and Richard Staelin. 2011. "Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days." *The American Journal of Managed Care*, 17(1): 41–8.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2013. "Healthcare Exceptionalism? Productivity and Allocation in the U.S. Healthcare Sector." National Bureau of Economic Research Working Paper 19200.

Chandra, Amitabh, and Douglas O. Staiger. 2007. "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks." *Journal of Political Economy*, 115(1): 103–140.

Chandra, Amitabh, and Jonathan Skinner. 2012. "Technology Growth and Expenditure Growth in Health Care." *Journal of Economic Literature*, 50(3): 645–80.

Collard-Wexler, Allan, and Jan De Loecker. 2015. "Reallocation and Technology: Evidence from the US Steel Industry." *American Economic Review*, 105(1): 131–71.

Cutler, David M. 2005. *Your Money Or Your Life: Strong Medicine for America's Health Care System.* Oxford University Press.

Cutler, David M. 2010. "Where Are The Health Care Entrepreneurs? The Failure of Organizational Innovation in Health Care." National Bureau of Economic Research Working Paper 16030.

Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler. 1998. "Are Medical Prices Declining? Evidence from Heart Attack Treatments." *The Quarterly Journal of Economics*, 113(4): 991–1024.

Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1998. *Job Creation and Destruction.* The MIT Press.

Disney, Richard, Jonathan Haskel, and Ylva Heden. 2003. "Restructuring and Productivity Growth in UK Manufacturing." *Economic Journal*, 113(489): 666–694.

Donabedian, Avedis. 1966. "Evaluating the Quality of Medical Care." *The Milbank Quarterly*, 44(3): 166–206.

Doyle, Joseph, John A. Graves, and Jonathan Gruber. 2014. "Paying for Quality in Healthcare." Mimeo.

Dranove, David, and Andrew Sfekas. 2008. "Start Spreading the News: A Structural Estimate of the Effects of New York Hospital Report Cards." *Journal of Health Economics*, 27(5): 1201–7.

Dranove, David, and Mark A. Satterthwaite. 2000. "The industrial organization of health care markets." In *Handbook of Health Economics*. Vol. 1B, ed. Anthony J. Culyer and Joseph P. Newhouse, Chapter 20, 1093–1139. Elsevier.

Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. 2003. "Is More Information Better? The Effects of "Report Cards" on Health Care Providers." *Journal of Political Economy*, 111(3): 555–588.

Elliott, Marc N., Alan M. Zaslavsky, Elizabeth Goldstein, William Lehrman, Katrin Hambarsoomians, Megan K. Beckett, and Laura Giordano. 2009. "Effects of survey mode, patient mix, and nonresponse on CAHPS hospital survey scores." *Health Services Research*, 44(2, Part I): 501–18.

Epstein, Arnold M. 2002. "Volume and Outcome–It Is Time to Move Ahead." *The New England Journal of Medicine*, 346(15): 1161–4.

Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *The Review of Economic Studies*, 62(1): 53–82.

Escribano, Alvaro, and J. Luis Guasch. 2005. "Assessing The Impact Of The Investment Climate On Productivity Using Firm-Level Data: Methodology And The Cases Of Guatemala, Honduras, And Nicaragua." World Bank Policy Research Working Paper 3621.

Fibrinolytic Therapy Trialists' Collaborative Group. 1994. "Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients." *Lancet*, 343(8893): 311–22.

Foster, Lucia, John C. Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *The American Economic Review*, 98(1): 394–425.

Foster, Lucia, John C. Haltiwanger, and C. J. Krizan. 2001. "Aggregate Productivity Growth. Lessons from Microeconomic Evidence." In *New Developments in Productivity Analysis*. 303–372. National Bureau of Economic Research.

Foster, Lucia, John C. Haltiwanger, and C. J. Krizan. 2006. "Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s." *The Review of Economics and Statistics*, 88(4): 748–758.

Garber, Alan M., and Jonathan Skinner. 2008. "Is American Health Care Uniquely Inefficient?" *The Journal of Economic Perspectives*, 22(4): 27–50.

Gaynor, Martin, and William B. Vogt. 2003. "Competition among Hospitals." *The RAND Journal of Economics*, 34(4): 764–785.

Giordano, Laura A., Marc N. Elliott, Elizabeth Goldstein, William G. Lehrman, and Patrice A. Spencer. 2010. "Development, implementation, and public reporting of the HCAHPS survey." *Medical Care Research and Review*, 67(1): 27–37.

Gorodeski, Eiran Z., Randall C. Starling, and Eugene H. Blackstone. 2010. "Are All Readmissions Bad Readmissions?" *The New England Journal of Medicine*, 363(3): 297–298.

Gottlieb, Daniel J., Weiping Zhou, Yunjie Song, Kathryn Gilman Andrews, Jonathan S. Skinner, and Jason M. Sutherland. 2010. "Prices Don't Drive Regional Medicare Spending Variations." *Health Affairs*, 29(3): 537–43.

Grosso, Laura M., Jeptha P. Curtis, Zhenqiu Lin, Lori L. Geary, Smitha Vellanky, Carol Oladele, Lesli S. Ott, Craig Parzynski, Susannah Bernheim, Lisa G. Suter, Elizabeth E. Drye, and Harlan M. Krumholz. 2012. "Hospital-level 30-Day All-Cause Risk-Standardized Readmission Rate Following Elective Primary Total Hip Arthroplasty (THA) And/Or Total Knee Arthroplasty (TKA)." Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation.

Hortaçsu, Ali, and Chad Syverson. 2007. "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." *Journal of Political Economy*, 115(2): 250–301.

Jacob, Brian A., and Lars Lefgren. 2007. "What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers." *The Quarterly Journal of Economics*, 122(4): 1603–1637.

Jencks, Stephen F., Edwin D. Huff, and Timothy Cuerdon. 2003. "Change in the quality of care delivered to Medicare beneficiaries, 1998-1999 to 2000-2001." *JAMA*, 289(3): 305–12.

Jencks, Stephen F., Mark V. Williams, and Eric A. Coleman. 2009. "Rehospitalizations among patients in the Medicare fee-for-service program." *The New England Journal of Medicine*, 360(14): 1418–28.

Jencks, Stephen F., Timothy Cuerdon, Dale R. Burwen, Barbara Fleming, Peter M. Houck, Annette E. Kussmaul, David S. Nilasena, Diana L. Ordin, and David R. Arday. 2000. "Quality of medical care delivered to Medicare beneficiaries: A profile at state and national levels." *JAMA*, 284(13): 1670–6.

Jha, Ashish K., E. John Orav, Jie Zheng, and Arnold M. Epstein. 2008. "Patients' perception of hospital care in the United States." *The New England Journal of Medicine*, 359(18): 1921–31.

Jha, Ashish K., E. John Orav, Zhonghe Li, and Arnold M. Epstein. 2007. "The inverse relationship between mortality rates and performance in the Hospital Quality Alliance measures." *Health Affairs*, 26(4): 1104–10.

Johnson, Erin. 2011. "Ability, Learning, and the Career Path of Cardiac Specialists." Mimeo.

Kahn, Charles N., Thomas Ault, Lisa Potetz, Thomas Walke, Jayne Hart Chambers, and Samantha Burch. 2015. "Assessing Medicare's Hospital Pay-For-Performance Programs And Whether They Are Achieving Their Goals." *Health Affairs*, 34(8): 1281–1288.

Kane, Thomas J., and Douglas O. Staiger. 2001. "Improving School Accountability Measures." National Bureau of Economic Research Working Paper 8156.

Keeley, Ellen C., and L. David Hillis. 2007. "Primary PCI for myocardial infarction with ST-segment elevation." *The New England Journal of Medicine*, 356(1): 47–54.

Keeley, Ellen C., Judith A. Boura, and Cindy L. Grines. 2003. "Primary Angioplasty versus Intravenous Thrombolytic Therapy for Acute Myocardial Infarction: A Quantitative Review of 23 Randomised Trials." *Lancet*, 361(9351): 13–20.

Kessler, Daniel P., and Mark B. McClellan. 2000. "Is Hospital Competition Socially Wasteful?" *The Quarterly Journal of Economics*, 115(2): 577–615.

Luft, Harold S., Deborah W. Garnick, David H. Mark, Deborah J. Peltzman, Ciaran S. Phibbs, Erik Lichtenberg, and Stephen J. McPhee. 1990. "Does quality influence choice of hospital?" *JAMA*, 263(21): 2899–906.

Manary, Matthew P., William Boulding, Richard Staelin, and Seth W. Glickman. 2013. "The patient experience and health outcomes." *The New England Journal of Medicine*, 368(3): 201–3.

McClellan, Mark, and Douglas Staiger. 2000. "Comparing the Quality of Health Care Providers." In *Frontiers in Health Policy Research, Volume 3*. , ed. Alan M. Garber, 113–136. MIT Press.

McClellan, Mark, Barbara J. McNeil, and Joseph P. Newhouse. 1994. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality? Analysis Using Instrumental Variables." *JAMA*, 272(11): 859–66.

MEDPAC. 2012. "Health Care Spending and the Medicare Program." http://www.medpac.gov/documents/reports/jun12_entirereport.pdf.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): 1695–1725.

Morris, Carl N. 1983. "Parametric Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association*, 78(381): 47–55.

Neumar, Robert W, Michael Shuster, Clifton W Callaway, Lana M Gent, Dianne L Atkins, Farhan Bhanji, Steven C Brooks, Allan R de Caen, Michael W Donnino, Jose Maria E Ferrer, Monica E Kleinman, Steven L Kronick, Eric J Lavonas, Mark S Link, Mary E Mancini, Laurie J Morrison, Robert E O'Connor, Ricardo A Samson, Steven M Schexnayder, Eunice M Singletary, Elizabeth H Sinz, Andrew H Travers, Myra H Wyckoff, and Mary Fran Hazinski. 2015. "Part 1: Executive Summary: 2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care." *Circulation*, 132(18 Suppl 2): S315–67.

O'Connor, Robert E, Abdulaziz S Al Ali, William J Brady, Chris A Ghaemmaghami, Venu Menon, Michelle Welsford, and Michael Shuster. 2015. "Part 9: Acute Coronary Syndromes: 2015 American Heart Association Guidelines Update for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care." *Circulation*, 132(18 Suppl 2): S483–500.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–97.

Pavcnik, Nina. 2002. "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants." *The Review of Economic Studies*, 69(1): 245–276.

QualityNet. 2015. "Hospital Value-Based Purchasing Overview." https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228772039937.

Rau, Jordan. 2013. "Methodology: How Value Based Purchasing Payments Are Calculated." http://khn.org/news/value-based-purchasing-medicare-methodology/.

Romley, John A., and Dana P. Goldman. 2011. "How Costly is Hospital Quality? A Revealed-Preference Approach." *The Journal of Industrial Economics*, 59(4): 578–608.

Scarpetta, Stefano, Philip Hemmings, Thierry Tressel, and Jaejoon Woo. 2002. "The Role of Policy and Institutions for Productivity and Firm Dynamics: Evidence from Micro and Industry Data." OECD Economics Working Paper 329.

Skinner, Jonathan, and Douglas Staiger. 2015. "Technology Diffusion and Productivity Growth in Health Care." *Review of Economics and Statistics*, 97(5): 951–964.

Skinner, Jonathan S. 2011. "Causes and Consequences of Regional Variations in Health Care." In *Handbook of Health Economics*. Vol. 2, ed. Mark V. Pauly, Thomas G. Mcguire and Pedro P. Barros, 45–93. Elsevier.

Skinner, Jonathan S., and Douglas O. Staiger. 2007. "Technology Adoption from Hybrid Corn to Beta-Blockers." In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*. , ed. Ernst R. Berndt and Charles R. Hulten, 545–570. University of Chicago Press.

Stone, Gregg W. 2008. "Angioplasty strategies in ST-segment-elevation myocardial infarction: part I: primary percutaneous coronary intervention." *Circulation*, 118(5): 538–51.

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature*, 49(2): 326–365.

Tay, Abigail. 2003. "Assessing Competition in Hospital Care Markets: The Importance of Accounting for Quality Differentiation." *The RAND Journal of Economics*, 34(4): pp. 786–814.

Town, Robert, and Gregory Vistnes. 2001. "Hospital competition in HMO networks." *Journal of Health Economics*, 20(5): 733–53.

Werner, Rachel M., and Eric T. Bradlow. 2006. "Relationship between Medicare's hospital compare performance measures and mortality rates." *JAMA*, 296(22): 2694–702.

Williams, Scott C., Stephen P. Schmaltz, David J. Morton, Richard G. Koss, and Jerod M. Loeb. 2005. "Quality of care in U.S. hospitals as reflected by standardized measures, 2002-2004." *The New England Journal of Medicine*, 353(3): 255–64.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data. Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Appendix To:

Healthcare Exceptionalism? Performance and Allocation in the U.S. Healthcare Sector

Amitabh Chandra, Amy Finkelstein, Adam Sacarny and Chad Syverson

January 2016

For Online Publication

Appendix A Analytical Framework

As mentioned in the text, models of reallocation mechanisms among heterogeneous-productivity producers have found applications in a number of fields, including industrial organization, trade, and macroeconomics. While these models differ considerably in their specifics, they share an archetypal mechanism that connects the extent of competition in the market (as reflected in consumers' willingness or ability to substitute among producers) to the shape of the productivity distribution among market producers. Here we sketch out a model with such a mechanism that fits the specific feature of the healthcare sector that consumers bear little if any of the financial costs of their firm choice.

We assume hospitals are heterogeneous in two dimensions: quality and costs. These two dimensions of heterogeneity may be correlated (e.g., higher quality hospitals might tend have higher costs on average), but this is not necessary for our results. While both quality and costs are likely to be at least in part affected by hospitals' choices, we follow the majority of the productivity literature and assume that they are exogenous and fixed.[34]

When choosing their hospital, we assume that patients care about quality but, conditional on quality, do not care about costs. The assumption that patients are not sensitive to hospital costs is a natural one, given that Medicare and supplemental insurance shields patients from paying for most of the care they receive, and likely all of the incremental cost associated with their hospital choice. Of course, while patients do not care about costs, a social planner does. A benevolent social planner would desire both high quality and low costs. The social planner would trade off between them based upon the parameters of the social welfare function. Thus there may be a wedge between the privately and socially optimal hospital choice.

Producers (indexed by h) earn profits which depend positively on their idiosyncratic quality levels q_h (higher quality firms earn higher profits because they draw more patients), negatively on their costs c_h, and negatively on the number (or mass, in models with a continuum of firms) of producers in the industry N.[35] Hence $\pi_h = \pi(q_h, c_h, N)$, with $\frac{\partial \pi}{\partial q_h} > 0$, $\frac{\partial \pi}{\partial c_h} < 0$, and $\frac{\partial \pi}{\partial N} < 0$. The monotonic relationship between quality and profits implies that, for any given N, there is locus of

[34] This model is a more generic and looser version of the type of multidimensional-heterogeneity-producer model in Foster, Haltiwanger and Syverson (2008).

[35] Standard presentations of these models consider profit-maximizing firms. Although we keep this terminology to be more familiar relative to the existing literature, we note that in the context of hospitals, it might be more appropriate to consider firms as earning (and maximizing) "surplus" rather than "profits". This more general terminology recognizes that many hospitals are legally structured as nonprofits. All that is required for the conceptual framework to carry over is for surplus to be increasing in quality (again because all else equal it increases the patient traffic at a hospital). Nonprofit hospitals are often modeled in the literature as having an objective function that is a convex combination of profits and other objectives; thus on the margin they should respond qualitatively the same way as for-profit hospitals to factors like competition. And indeed a large empirical literature finds essentially no evidence of differential behavior across for-profit and non-profit hospitals, calling into question whether the non-profit label has any substantive meaning for behavioral responses (see Sloan, 2000 for a review of this literature).

critical cutoff quality and cost levels at which hospital profits are zero. Along this locus, quality is a monotonically increasing function of costs, because higher costs require a higher quality level for a hospital to earn zero profit.

Call this locus $q^*(c,N)$, where we have expressed it as the critical value of quality necessary to earn zero profit as a function of costs and the number of hospitals in the market. Only producers with quality levels at or above $q^*(c,N)$ will operate in equilibrium.

The zero-profit cutoff locus is endogenously determined by a free entry condition, where ex-ante identical potential entrants consider whether to pay a sunk cost σ to take idiosyncratic quality and cost draws from a known joint distribution of q and c, $G(\cdot)$, with an upper bound quality of $\overline{q}$ and a lower and upper bound costs $\underline{c}$ and $\overline{c}$, respectively. The expected value of entry, which equals zero by the free entry condition, is:

$$V^e = \int_{\underline{c}}^{\overline{c}} \int_{q^*(c,N)}^{\overline{q}} \pi(q,c,N)\, g(q,c)\, dqdc - \sigma = 0$$

The expected profits from entry depend upon the equilibrium number of entrants N in two ways. First, an increase in N shifts upward the zero-profit cutoff quality level $q^*(c,N)$, reducing the probability that the entrant's quality and cost draws are high and low enough (respectively) to earn nonnegative profits, and thus making successful entry less likely. Second, a higher number of firms N also reduces the producer's profits if it does enter. Thus expected profits fall monotonically in N. In equilibrium, the number of firms choosing to pay the entry cost yields a number of entrants N that, through these two effects, exactly equates the expected profit from taking a quality and cost draw to the sunk entry cost.

The endogeneity of $q^*(c,N)$ means the industry quality and cost distribution observed in the data is determined in equilibrium. Specifically, it is a truncation of $G(\cdot)$, the underlying distribution from which potential entrants take quality and cost draws, where the truncation locus is $q^*(c,N)$. Changes in market primitives that shift the equilibrium location of $q^*(c,N)$ therefore shift the observed joint distribution of quality and costs.

The primitive that underlies these results is the extent to which patients are able or willing to substitute to alternate hospitals in order to obtain higher quality. The specific mechanism through which primitives map into substitutability may vary from, for example, the extent of information available to patients or their surrogates, to differences in travel costs. The particulars of the mechanism aren't important here; what matters are the effects on the equilibrium.

This framework has several predictions that we examine empirically. In equilibrium, if patients have some ability to substitute across alternate producers (hospitals), there is a robust prediction that the market will allocate patients to higher quality hospitals on average, so that there is a correlation between quality and market share at a point in time ("*static allocation*"). In addition, over

time higher quality hospitals will be more likely to grow in market share ("*dynamic allocation*"). Our empirical work in Section IV focuses on examining these static and dynamic equilibrium allocations. The model also generates the comparative static prediction that these static and dynamic equilibrium allocation results will be stronger where patients have greater ability to substitute to alternate producers. In Section V we test this comparative static prediction by comparing allocation results for patients admitted through the emergency department and patients admitted as non-emergency transfers from another hospital. Stratifying on the method of admission to the hospital offers one way to distinguish among patients with different abilities to substitute to alternate producers.[36]

Finally, we note that endogenous selection based on patients' preferences for quality also has implications for equilibrium cost levels. Even if quality and cost draws are uncorrelated in $G(\cdot)$, factors that tend to truncate the equilibrium quality distribution at a higher level will also raise average observed costs, because hospitals with higher quality can have higher costs before becoming unprofitable. Thus when patients are not sensitive to costs and choose based solely on quality, the equilibrium will tend to allocate toward both higher quality and higher cost firms. As noted, there may therefore be a wedge between the privately and socially optimal allocations.

[36]This model is static, so the effects of changes in competition on equilibrium should be thought of as comparing two different markets or the same market across different long-run steady states. However, several of the models in the literature are explicitly dynamic and have similar predictions about the effect of competition on the productivity of entrants and growth of incumbents (e.g. Hopenhayn, 1992; Asplund and Nocke, 2006).

Appendix B Quality Measures

This section provides details on the definition and construction of each of our four quality metrics.

B.1 Risk-adjusted survival

Risk-adjusted survival is arguably the key endpoint for emergent conditions and has been the health outcome of choice for a large economics and medical literature. CMS publicly reports risk-adjusted survival measures for heart attacks, heart failure and pneumonia for hospitals that treat at least 25 patients with the condition in the 3 year window it uses for the analysis. We calculate our own risk-adjusted survival scores for these conditions in order to have control over the regression model (we use a fixed effects linear regression while CMS uses a random effects logit), shrinkage approach (we use empirical Bayes adjustment while CMS uses the best linear unbiased predictors, or BLUPs, of the random effects), and risk adjustment (we test sensitivity to alternative sets of risk-adjusters while CMS only publishes one approach). The CMS data are also reported as ratios of observed mortality rates relative to expected mortality rates, which is a nonlinear transformation of the hospital random effects – and one that is not designed to produce unbiased coefficients when placed on the right-hand side of our allocation regressions.[37]

Mimicking the CMS measure, for each hospital with at least 25 patients with the condition between 2006 and 2008, we estimate a risk-adjusted survival rate – the probability that a Medicare patient would survive 30 days after being treated for the condition at the hospital. Specifically, we start with the patient-level sample of initial hospitalizations for the condition, or index events, from 2006-2008. Then, we regress 30-day survival (counting from the patient's hospital admission date) on a rich set of observable information about the patient, including age/race/sex interactions and indicators for being hospitalized for 25 conditions in the past year, as well as hospital fixed effects.[38] The inclusion of risk-adjusters is standard practice in the literature and is designed to minimize the impact of differences in patient health across hospitals on survival rates.

We extract the hospital fixed effects, which become the risk-adjusted survival rate estimates for the hospitals for a given condition. Since these estimates include measurement error, they may

[37] Since we use different regression models, shrinkage approaches, risk-adjustment approaches, and transformations of hospital effects, we do not expect the correlations between our measures and the CMS measures to be 100%. Still, in Appendix Table A19, we find that our risk-adjusted survival and readmission measures are highly correlated with the CMS measures, with correlation coefficients ranging from 0.66 to 0.82.

[38] The age groups are 66-69, 70-74, 75-79, 80-84, 85-89, 90-94, and 95+. The race and age groups are white/non-white and male/not male. The risk-adjusters are: heart failure, myocardial infarction, unstable angina, chronic atherosclerosis, respiratory failure, hypertensive heart disease, valvular heart disease, arrhythmia, hypertension, stroke, cerebrovascular disease, renal failure disease, dialysis, COPD, pneumonia, diabetes, protein calorie malnutrition, dementia, paralysis and disability, peripheral vascular disease, metastatic cancer, trauma, substance abuse, major psychiatric disorder, and chronic liver disease.

produce biased coefficients when included on the right hand side of our allocation regressions; for example, classical measurement error will cause attenuation bias toward zero. In Appendix C, we describe the empirical Bayes shrinkage approach we use to correct for measurement error bias. We discuss and show how the shrinkage method affects our results in Appendix Section C.5 and Appendix Table A2 – consistent with the presence of measurement error bias resulting in attenuation, the correction expands nearly all of our allocation gradients.

In our analyses studying risk-adjusted survival over the long horizon, we calculate the measure for 1996, 1999, 2002, and 2005 in addition to the baseline 2008 measure. As in the 2008 calculations, we aggregate over three years, e.g. the 1996 measure includes patients from 1994-1996. Results using this longer sample of risk-adjusted survival rates are presented in Section IV.B.2.

B.2 Risk-adjusted readmission

This measure, defined and estimated similarly to risk-adjusted survival, indicates the probability that an average Medicare patient would be readmitted within 30 days after discharge from her initial hospital stay. It is widely used as a proxy for medical errors and inappropriate discharge. Mimicking the CMS measure, the sample of patients is the same as that for risk-adjusted survival, with the addition of the following exclusion criteria: if the patient dies during the initial hospital stay, is transferred from her initial hospital to another inpatient facility, or leaves the hospital against medical advice, the patient is removed from the sample. Per the CMS approach, these exclusions help to remove patients who either could not be readmitted or whose readmissions might not be due to the index hospital's quality of care. We then use the same regression, risk adjustment, and empirical Bayes method as in risk-adjusted survival. For hip and knee replacement, an indicator for whether the patient received a hip replacement is also included as a risk-adjuster to allow for differential readmission rates depending on which joint is being replaced.

B.3 Process of care

Publicly reported "process of care" measures give the shares of eligible patients who received certain evidence-based interventions. Hospitals report their utilization of these processes to CMS, which publishes the information online and uses it to adjust hospital payments.[39] The data pertain to all eligible patients irrespective of their insurer, and are not limited to patients covered by Medicare. Patients for whom the interventions are contraindicated are not counted in the numerator or denominator of the shares. We consider the process measures for specific inpatient conditions that were reliably reported from 2006 through 2008: 6 AMI measures, 4 heart failure measures, and 7 pneumonia measures.

[39]These data can be downloaded at https://data.medicare.gov/data/hospital-compare

The processes "were identified with respect to published scientific evidence and consistency with established clinical-practice guidelines" (Williams et al., 2005). The AMI measures have their origins in the Cooperative Cardiovascular Project, a large study of AMI among Medicare beneficiaries that was conducted in the 1990s; the metrics for heart failure and pneumonia can be traced back to the practice guidelines of professional organizations that focus on these conditions (Jencks et al., 2000). The AMI processes cover the administering of aspirin (one measure for arrival and another for discharge), ACE inhibitors, smoking cessation advice, β blockers, and angioplasty (percutaneous coronary intervention, or PCI).[40] The heart failure processes cover providing discharge instructions for care, evaluation of left ventricular systolic dysfunction, ACE inhibitors, and smoking cessation advice. The pneumonia processes cover providing oxygenation assessments, pneumococcal vaccines, blood cultures before antibiotics, smoking cessation advice, timely antibiotics, the most appropriate antibiotics for the particular infection, and influenza vaccines.

To reduce the dimensionality of the process measures before including them in our regressions, for each condition we generate a hospital-level composite measure of adherence to the condition's processes. We start with the publicly reported hospital-level data from 2006 to 2008; we combine 3 years to reduce measurement error. To further reduce measurement error in the composite score, we remove any individual process score for which the hospital had fewer than 50 eligible patients over the 3 year window (since each process has different contraindications, a hospital may have more than 50 patients in one score for a condition and fewer than 50 patients in another). We use a higher patient count threshold than for risk-adjusted survival and readmission because process scores are not empirical-Bayes-adjusted by CMS to account for measurement error; since the scores have all-payer coverage, the patient counts tend to be larger than for the risk-adjusted outcomes and this threshold is therefore less restrictive.

We standardize each process score (among the set of hospitals that reported it for at least 50 patients over the 3 years) to have mean 0 and standard deviation 1. For each condition, we average together the condition's individual process standardized scores to create a composite score, then standardize that composite score. The result is a condition-specific composite score with zero mean and unit variance defined on the set of hospitals that reported 50 or more patients for at least one process of care for that condition over 2006-2008.

The process of care data are only available at the hospital level so it is not possible to perform the kind of detailed risk adjustment when generating these quality metrics that we could for survival or readmission. However, one advantage of these metrics is that they are designed to measure interventions or experiences that the facility should deliver to essentially all of its patients; patients

[40] An additional AMI measure, Thrombolytics at Arrival, was missing for 60-80% of hospitals each year, and was removed from the analysis.

who are inappropriate for the intervention are excluded. As a result, risk adjustment is not obviously relevant or required for the process of care measures – while it may not be possible for a hospital to prevent all AMI deaths or readmissions, it is possible in theory for a facility to administer β blockers to all appropriate patients at discharge. Indeed, the study of processes of care has been justified by hospitals' ability to directly control these measures of quality, since quality scores based on clinical outcomes include factors like the hospital's patient population and patient compliance with post-discharge care that hospitals are less able to manage (Donabedian, 1966).

B.4 Patient satisfaction

Patient satisfaction is measured by the 2008 HCAHPS (Hospital Consumer Assessment of Healthcare Providers and Systems), which hospitals administer to their patients after discharge.[41] The survey is given to a sample of all of the hospital's patients, not just patients who were covered by Medicare; unlike our other quality measures, the public data covers all patients, and is not offered in a condition-specific manner. The survey results are processed and reported by CMS; the survey instrument is condensed into ten measures of the patient's experience and perceived quality of care. The ten measures are: communication with nurses, communication with doctors, responsiveness of hospital staff, pain management, communication about medicines, cleanliness of hospital environment, quietness of hospital environment, discharge information, overall hospital rating, and recommend the hospital. The average hospital reports scores for all 10 survey questions. CMS adjusts the results for interview mode (e.g. mail, telephone, etc.) and a set of patient characteristics. Its adjustment for mode uses data from a randomized trial comparing survey responses by mode, while the adjustment for patient characteristics comes from a model that is estimated quarterly from hospitals' submissions (Giordano et al., 2010; Elliott et al., 2009).

We generate a composite score of hospital performance on the patient survey by aggregating together the 10 questions into a measure with mean zero and unit variance. For all of the questions but one, the publicly reported data indicates the share of patients responding that the hospital provided high, medium or low quality. One question (discharge information) is reported as the shares of patients responding yes or no. We assign these responses to numeric values (3/2/1 for the three-level questions or 1/0 for the yes/no question, with higher values always better) and compute an average response for each hospital. Then, following the same method we use for process of care, we generate standardized scores for each question, average together the standardized scores, and standardize the result.

[41] For an overview of the design and implementation of HCAHPS, see Giordano et al. (2010).

Appendix C Empirical Bayes Adjustment

In this appendix we describe the empirical Bayes (EB) procedure we use to adjust our estimates of risk-adjusted survival, risk-adjusted readmission, risk-adjusted inputs, and input-adjusted risk-adjusted survival (which we call productivity) for measurement error. This procedure is based on Morris (1983). For another example see Jacob and Lefgren (2007).

For hospital h, its quality measure (risk-adjusted survival, risk-adjusted readmission, or input-adjusted risk adjusted survival; we also will refer to risk-adjusted inputs in this manner) is called q_h. These objects are the "true" quality values and their distribution is the "underlying" distribution of quality. We denote by $\hat{q}_h$ the estimate of quality; it equals quality plus an error term η_h:

$$\hat{q}_h = q_h + \eta_h$$

The goal of the EB procedure is to adjust the estimates of quality so that the presence of the error term does not introduce bias when the quality estimates are included as regressors in our allocation regressions (see equations 1 and 2). The procedure adjusts the estimates by shrinking them toward the mean of the true, underlying distribution. True quality is not observable, but we show in this appendix that its distribution is estimable. We also show how this shrinkage estimator fixes the attenuation bias that measurement error could otherwise introduce into our regressions.

In this appendix we use bold lowercase greek and roman letters to refer to vectors and uppercase greek and roman letters for matrices. Non-bold lowercase letters refer to scalars.

C.1 Background on Empirical Bayes Procedure

C.1.1 Statistical Background

We start with an overview of the EB procedure assuming that all parameters of the distributions are known, and refer to the EB-adjusted estimated quality as q_h^{EB}. We then describe the feasible EB-adjusted estimate, which we denote $q_h^{EB(f)}$.

Suppose that the estimated quality is independently normally distributed around the true quality with known variance π_h^2:

$$\hat{q}_h | q_h, \pi_h^2 \sim N\left(q_h, \pi_h^2\right) \text{ independently}$$

One can think of π_h^2 as the variance of the measurement error of the estimate.

We also assume that the true quality q_h is independently normal with underlying mean $\mathbf{x}_h'\boldsymbol{\lambda}$ (a known, linear function of the hospital's covariates) and underlying variance σ^2 (known and common across hospitals).

The *prior distribution* of quality q_h – the distribution before conditioning on the estimated

quality – is therefore:

$$q_h|\mathbf{x}_h, \boldsymbol{\lambda}, \sigma^2 \sim N\left(\mathbf{x}_h'\boldsymbol{\lambda}, \sigma^2\right) \text{ independently}$$

Conditioning on the estimated quality $\hat{q}_h$ produces the *posterior distribution* of q_h:

$$q_h|\hat{q}_h, \mathbf{x}_h, \boldsymbol{\lambda}, \sigma^2, \pi_h^2 \sim N\left(q_h^{EB}, \pi_h^2\left(1-b_h\right)\right) \tag{A1}$$

q_h^{EB} denotes the EB-adjusted quality. This object is the expected value of q_h conditional on the estimated value $\hat{q}_h$ and the parameters $\boldsymbol{\lambda}, \sigma^2$, and π_h^2 and is given by the formula:

$$q_h^{EB} = \left(1-b_h\right)\hat{q}_h + b_h\mathbf{x}_h'\boldsymbol{\lambda} \tag{A2}$$

$$b_h = \pi_h^2/\left(\pi_h^2+\sigma^2\right) \tag{A3}$$

The adjustment amounts to attenuating the estimate $\hat{q}_h$ toward the prior mean $\mathbf{x}_h'\boldsymbol{\lambda}$. As the variance of the measurement error π_h^2 rises, the EB correction increasingly disregards the value of the estimate and closes in on the prior mean.

C.1.2 Feasible Version of Procedure

This section describes how we implement the EB procedure when parameters must be estimated.

The value $\hat{q}_h$ is the estimated hospital fixed effect from the regression used to estimate quality (see Appendix B for a description of the regressions used to estimate risk-adjusted survival and readmission; see equation 9 for the regression used to estimate input-adjusted risk-adjusted survival, a.k.a. "productivity"). We estimate a standard error for the fixed effect assuming homoscedastic disturbances in the first-step patient-level regression; under the homoscedasticity assumption, the standard error is our estimate of the standard deviation of the asymptotic distribution of $\hat{q}_h$. We estimate π_h^2 by squaring the standard error and call this value $\hat{\pi}_h^2$.

We estimate the underlying parameters of the quality distribution, $\boldsymbol{\lambda}$ and σ^2, using the method outlined in section 5 of Morris (1983). We fix yearly estimates:

$$\begin{aligned}
\hat{\boldsymbol{\lambda}} &:= \left(X'WX\right)^{-1}X'WQ \\
\hat{\sigma}^2 &= \max\left\{0, \frac{\sum_h W_h\left\{\left(\frac{n_H}{n_H-n_X}\right)\left(\hat{q}_h-\mathbf{x}_h'\boldsymbol{\lambda}\right)^2-\hat{\pi}_h^2\right\}}{\sum_h w_h}\right\} \\
w_h &= \frac{1}{\hat{\pi}_h^2+\hat{\sigma}^2}
\end{aligned}$$

where X is the stacked $\mathbf{x}_h'$, W is a diagonal matrix of the w_h, and Q is the stacked $\hat{q}_h$ for year t. n_H is the number of hospitals, or equivalently the number of $\hat{q}_h$. n_X is the number of regressors,

i.e. the dimensionality of $\mathbf{x}_h$.

$\hat{\boldsymbol{\lambda}}$ is a WLS regression of the $\hat{q}_h$ on $\mathbf{x}_h$. $\hat{\sigma}^2$ is the weighted average of the squared deviations of $\hat{q}_h$ from $\mathbf{x}_h'\hat{\boldsymbol{\lambda}}$ less the weighted average of $\hat{\pi}_h^2$. The weights are w_h, giving more weight to observations with less measurement error. The max operator ensures that $\hat{\sigma}^2$ is always nonnegative in finite samples.

$\hat{\boldsymbol{\lambda}}$ and $\hat{\sigma}^2$ are simultaneously determined in these equations, so they are estimated by the following iterative procedure. We start by fixing $w_h = 1 \,\forall\, h$, then iterate the following to convergence:

1. Compute $\hat{\boldsymbol{\lambda}}$ and then a new estimate $\hat{\sigma}^2$
2. If this is the second or greater iteration and $\hat{\sigma}^2$ has converged, exit. Otherwise, fix new weights w_h and return to step 1

The (feasible) best estimate of the posterior mean $q_h^{EB(f)}$ is given in Morris (1983) by the formula of equations (A2) and (A3) with a degrees of freedom adjustment :

$$\begin{aligned} q_h^{EB(f)} &= \left(1-\hat{b}_h\right)\hat{q}_h + \hat{b}_h \mathbf{x}_h'\hat{\boldsymbol{\lambda}} \\ \hat{b}_h &= \left(\frac{n_H - n_X - 2}{n_H - n_X}\right)\left(\frac{\hat{\pi}_h^2}{\hat{\pi}_h^2 + \hat{\sigma}^2}\right) \end{aligned}$$

The variance of the quality distribution unconditional on covariates, called $\hat{\varsigma}^2$, is given by the following formula:

$$\text{(A4)} \qquad \hat{\varsigma}^2 = \max\left\{0, \frac{\sum_h w_h \left\{\left(\frac{n_H}{n_H-1}\right)(\hat{q}_h - \bar{q}) - \hat{\pi}_h^2\right\}}{\sum_h w_h}\right\}$$

$$\text{(A5)} \qquad \bar{q} = \frac{\sum_h w_h \hat{q}_h}{\sum_h w_h}$$

Where $\bar{q}$ is the weighted mean quality.

C.2 Implementation of Empirical Bayes Adjustment

We assume that the underlying mean of quality is equal to a market fixed effect, i.e. $\mathbf{x}_h'\boldsymbol{\lambda} = \tau_M$, where M indexes markets. Thus $\mathbf{x}_h$ becomes a vector of 306 indicators for whether hospital h was in each of the 306 markets and $\boldsymbol{\lambda}$ is a vector of the 306 market fixed effects. We then perform the EB procedure, producing estimates of the underlying market means $\hat{\boldsymbol{\lambda}}$ and conditional – i.e. within-market – variance $\hat{\sigma}^2$. Running the procedure also yields EB-adjusted estimated quality measures $q_h^{EB(f)}$ and also can be used to produce the unconditional – i.e. national – estimated

variance $\hat{\varsigma}^2$, as described above. When we compute quality metrics for multiple years, for example in the case of Appendix Table A13, we perform the EB adjustment separately for each year. That is, we allow each year to have its own market means $\hat{\boldsymbol{\lambda}}$ and conditional variance $\hat{\sigma}^2$.

Our procedure ensures that when the EB-adjusted quality is used in our main regressions (equations 1 and 2 in the main text), which have market fixed effects, all regressors are orthogonal to the measurement error term.

C.3 Reported Statistics Involving Quality Metrics

C.3.1 Standard Deviation

To estimate the standard deviation of quality in Table 2, we rely on the estimates of the underlying national variance of quality $\hat{\varsigma}^2$ that the procedure computes.[42] The root of these estimates is taken, forming $\hat{\varsigma}$.

The EB adjustment produces $\hat{\varsigma}^2$ by taking the weighted empirical variance of the $\hat{q}_h$ and subtracting the weighted average squared standard error $\hat{\pi}_h^2$ (see equations A4 and A5). Hospitals with larger standard errors receive lower weights. In effect, this process takes the variance of the noisy quality estimates and subtracts off the variance due to measurement error.

C.3.2 Correlations

In Table 3 and Appendix Table A4 we report correlations adjusted for measurement error. The raw correlation between two quality measures potentially suffers from two sources of bias. First, the variance terms in the denominator are upward-biased if either quality measure is estimated with measurement error, as in the fixed effects approach that we use for risk-adjusted survival and readmission. Second, the covariance term in the numerator may also be biased if the two quality metrics were estimated using the same samples of patients (e.g. risk-adjusted survival and readmission for the same condition), since the sampling error in one fixed effect may be correlated with the sampling error in the other.

Our empirical correlation estimate corrects for these two sources of bias, and is calculated as the following:

[42] While it might seem natural to instead estimate the standard deviation of the EB-adjusted values, this would cause us to erroneously under-estimate dispersion. True quality is composed of a best prediction (the EB-adjusted quality) and the prediction error. These two components are orthogonal. The variance of true quality is thus strictly greater than the variance of EB-adjusted quality (see Jacob and Lefgren, 2007).

$$
\begin{aligned}
\tilde{\text{Cor}}_h\left(\hat{q}_{A,h},\hat{q}_{B,h}\right) &= \frac{\tilde{\text{Cov}}_h\left(\hat{q}_{A,h},\hat{q}_{B,h}\right)}{\sqrt{\tilde{\text{Var}}_h\left(\hat{q}_{A,h}\right)\tilde{\text{Var}}_h\left(\hat{q}_{B,h}\right)}} \\
\tilde{\text{Cov}}_h\left(\hat{q}_{A,h},\hat{q}_{B,h}\right) &= \hat{\text{Cov}}_h\left(\hat{q}_{A,h},\hat{q}_{B,h}\right) - \hat{\mathbb{E}}_h\left[\pi_{AB,h}\right] \\
\tilde{\text{Var}}_h\left(\hat{q}_{X,h}\right) &= \hat{\text{Var}}_h\left(\hat{q}_{X,h}\right) - \hat{\mathbb{E}}_h\left[\pi^2_{X,h}\right]
\end{aligned}
$$

where $\hat{q}_{A,h}$ is one estimated quality score for hospital h and $\hat{q}_{B,h}$ is another, $\pi_{AB,h}$ is the estimate of the measurement error covariance between the two quality scores, and $\pi^2_{X,h}$ is the estimate of the measurement error variance for measure $X \in \{A,B\}$. Tildes indicate estimates that have been adjusted for measurement error and hats indicate raw sample averages and variances. In other words, we take the covariance of the two raw quality scores and subtract the average covariance of the measurement error, and we take the variance of each quality score and subtract its average measurement error variance.

When the two quality scores come from the same patient sample, $\pi_{AB,h}$ is derived by making a homoscedasticity assumption on the covariance of the error terms in the two first-step regressions that produce the quality measures. Each pair of patient error terms, one for each regression, is assumed to be drawn from a distribution with a common variance-covariance matrix; error terms across patients are uncorrelated. We then estimate the variance-covariance matrix of the two hospital fixed effects and $\pi_{AB,h}$ is set to the covariance. If the two quality scores are derived from different patient samples (i.e. risk-adjusted survival for AMI and heart failure) or if only one is estimated by us from patient data (i.e. risk-adjusted survival for AMI and process of care for AMI) $\pi_{AB,h}$ is set to 0.

$\pi^2_{X,h}$ is the squared standard error of the fixed effect for measure X, described in Appendix Section C.1.2. It is set to 0 if the measure is not estimated by us from patient data, like for process of care scores or the patient survey.

C.3.3 Static and Dynamic Allocation Regressions

The allocation metrics use noisy estimates of quality on the right-hand side of regressions, and they rely on EB adjustment to correct for measurement error. Jacob and Lefgren (2007) show that with the adjustment, these regressions are estimated consistently.

Suppose that there is a relationship between growth Δ_h, market fixed effects γ_M, and quality q_h:

$$\Delta_h = \gamma_M + \delta q_h + \varepsilon_h$$

where $\mathbb{E}\left[\varepsilon_h | x_h, q_h\right] = 0$ (x_h is a vector of indicators for the markets – the design matrix for the

market fixed effects). The left-hand side variable could alternatively be the number of patients as in the static allocation regression.

Since we do not observe true quality q_h, we use the estimate $\hat{q}_h = q_h + \eta_h$, where η_h is measurement error. Then substituting into the equation:

$$\Delta_h = \gamma_M + \delta \hat{q}_h + (\varepsilon_h - \delta \eta_h)$$

This regression generally produces a biased estimate of δ due to the correlation between $\hat{q}_h$ and η_h in the error term. We use the EB-adjusted quality q_h^{EB} to eliminate this correlation. Equation (A1) implies:

$$\mathbb{E}\left[q_h | \hat{q}_h, \mathbf{x}_h, \boldsymbol{\lambda}, \sigma^2, \pi_h^2\right] = q_h^{EB}$$

We represent the prediction error of the EB procedure as v_h:

$$q_h = q_h^{EB} + v_h$$

By construction the prediction error is orthogonal to q_h^{EB} and any regressor included in x_h – i.e. the market fixed effects:

$$\mathbb{E}\left[v_h | q_h^{EB}, \mathbf{x}_h, \boldsymbol{\lambda}, \sigma^2, \pi_h^2\right] = 0$$

($\hat{q}_{ht}$ is replaced by q_h^{EB} because given the parameters, knowing one determines the other)

The regression of Δ_h on market effects and q_h^{EB} adds only δv_h to the original error term ε_h:

$$\Delta_h = \gamma_M + \delta q_h^{EB} + (\varepsilon_h - \delta v_h)$$

Therefore there is no correlation between any of the regressors and the new error term. The unbiasedness of δ follows.

C.4 Multivariate Empirical Bayes Procedure

In some cases we run regressions with multiple imprecisely measured quality metrics on the right-hand side, each estimated from the same sample of patients. In these cases, the measurement error across the quality metrics is likely to be correlated, making the EB procedure we used for a single quality metric insufficient to restore unbiasedness to the regression estimates – one EB-adjusted quality metric is uncorrelated with its own prediction error, but it may be correlated with the prediction error of the other EB-adjusted quality metric. For these regressions, the quality metrics must be EB-adjusted jointly.

Let $\mathbf{q}_h$ be the "true" vector of quality for hospital h and let $\hat{\mathbf{q}}_h$ be the estimate of the vector. The multivariate method assumes that the two quality estimates are distributed jointly normal around

true quality with covariance matrix Π_h:

$$\hat{\mathbf{q}}_h|\mathbf{q}_h,\Pi_h \sim \mathrm{N}(\mathbf{q}_h,\Pi_h) \text{ independently} \tag{A6}$$

The underlying, or prior, distribution of quality is also jointly normal:

$$\mathbf{q}_h|\mathbf{x}_h,\Lambda,\Sigma \sim \mathrm{N}(\Lambda\mathbf{x}_h,\Sigma) \text{ independently} \tag{A7}$$

Conditioning on the estimated quality vector yields the posterior distribution (we present the formula given in Murphy, 2007 after some algebraic manipulation):

$$\mathbf{q}_h|\hat{\mathbf{q}}_h,\mathbf{x}_h,\Lambda,\Sigma \sim \mathrm{N}\left(\mathbf{q}_h^{EB},(I-B_h)\Pi_h\right) \tag{A8}$$

Where:

$$\mathbf{q}_h^{EB} = (I-B_h)\hat{\mathbf{q}}_h + B_h\Lambda\mathbf{x}_h \tag{A9}$$

$$B_h = \Pi_h(\Pi_h+\Sigma)^{-1} \tag{A10}$$

One can think of $\mathbf{q}_h$ as the hyperparameter for the mean of $\hat{\mathbf{q}}_h$. The above formulas give the posterior distribution of the hyperparameter after conditioning on realization $\hat{\mathbf{q}}_h$.

C.4.1 Feasible Version of Procedure

To implement the EB adjustment, we begin by fixing values of $\hat{\mathbf{q}}_h$. Each member of the vector equals the estimated fixed effect from a patient-level quality regression, e.g. for $\hat{\mathbf{q}}_h = \left(\hat{q}_{A,h},\hat{q}_{B,h}\right)'$, $\hat{q}_{A,h}$ could be hospital h's fixed effect from the risk-adjusted survival patient-level regression while $\hat{q}_{B,h}$ could be the fixed effect from the readmission regression.

To construct $\hat{\Pi}_h$ (the estimate of Π_h), we assume homoscedastic disturbances in each first-step quality regression, but we extend the assumptions to account for multiple measures. We treat the set of first-step regressions as a SUR and assume that each patient's disturbances are drawn from a distribution with a common covariance matrix. That is, we allow a patient's disturbance term in one quality regression to be correlated with her disturbance term in another. Disturbances across patients are assumed to be uncorrelated. (We make the same assumption in Appendix Section C.3.2 to estimate correlations between quality measures.)

Under these assumptions, we extract hospital-level estimates of the covariance of the measure-

ment error between the quality measures, for example between the hospital's risk-adjusted survival score and its risk-adjusted readmission score. These estimates become the off diagonal values of $\hat{\Pi}_h$. We also estimate standard errors on the hospital fixed effects in each regression – under our homoscedasticity assumption these standard errors are the same as in the single quality-metric approach. The squared standard errors are estimates of the variance of the measurement error of each quality metric; these values become the diagonals of $\hat{\Pi}_h$.

Next we must estimate Σ and Λ. Combining equations (A6) and (A7) we have the distribution of $\hat{\mathbf{q}}_h$ unconditional on $\mathbf{q}_h$ (in Bayesian terms called the prior predictive distribution):

$$\hat{\mathbf{q}}_h|\mathbf{x}_h,\Lambda,\Sigma,\Pi_h \sim \mathrm{N}\left(\Lambda\mathbf{x}_h,\Sigma+\Pi_h\right) \text{ independently}$$

The full vector of measured quality $\hat{\mathbf{q}}$ – the stacked $\hat{\mathbf{q}}_h$ – therefore follows a multivariate normal distribution as well. We now show how to represent this joint distribution so that we can build its likelihood function.

For the simultaneous EB adjustment of k quality measures at once, we define X as the stacked $I_k \otimes \mathbf{x}_h'$, $\boldsymbol{\lambda}$ as the rows of Λ transposed to column vectors and stacked to create one vector of coefficients, and Π as the block diagonal matrix formed with Π_h on the diagonals. Then $\hat{\mathbf{q}}$ is distributed:

$$\hat{\mathbf{q}}|X,\Lambda,\Sigma,\Pi \sim \mathrm{N}\left(X\boldsymbol{\lambda},I_{n_h}\otimes\Sigma+\Pi\right)$$

In our model, the number of parameters in $\boldsymbol{\lambda}$ is large relative to the sample size – 306 market fixed effects per quality measure and about 3,000 hospitals per measure. An ML estimate of Σ would therefore have significant bias in finite samples due to the loss of degrees of freedom from estimating $\boldsymbol{\lambda}$. We estimate Σ by REML instead of MLE to avoid this bias.

The REML likelihood function is:[43]

$$\text{(A11)} \qquad \mathscr{L}\left(\tilde{\Sigma};\Pi\right) = -\frac{1}{2}\ln\left|\tilde{\Theta}\right| - \frac{1}{2}\ln\left|X'\tilde{\Theta}^{-1}X\right| - \frac{1}{2}\left(\hat{\mathbf{q}}-X\tilde{\boldsymbol{\lambda}}\right)'\tilde{\Theta}^{-1}\left(\hat{\mathbf{q}}-X\tilde{\boldsymbol{\lambda}}\right)$$

$$\text{(A12)} \qquad \tilde{\boldsymbol{\lambda}}\left(\tilde{\Theta}\right) = \left(X'\tilde{\Theta}^{-1}X\right)^{-1}X'\tilde{\Theta}^{-1}\hat{\mathbf{q}}$$

$$\text{(A13)} \qquad \tilde{\Theta}\left(\tilde{\Sigma};\Pi\right) = I_{n_h}\otimes\tilde{\Sigma}+\Pi$$

$\hat{\Sigma}$ is the maximizer of the likelihood function (with unknown Π replaced by the known $\hat{\Pi}$) and $\hat{\boldsymbol{\lambda}}$ is given by equation (A12):

[43] This likelihood function is derived and given in Diggle et al. (2002). The MLE likelihood function is defined identically but omits the $\frac{1}{2}\ln\left|X'\tilde{\Theta}^{-1}X\right|$ term. Maximizing this likelihood would yield unbiased estimates of λ but not Σ.

$$\begin{aligned}\hat{\tilde{\Sigma}} &= \arg_{\tilde{\Sigma}} \max \mathscr{L}\left(\tilde{\Sigma};\hat{\Pi}\right) \\ \hat{\boldsymbol{\lambda}} &= \tilde{\boldsymbol{\lambda}}\left(\tilde{\Theta}\left(\hat{\tilde{\Sigma}};\hat{\Pi}\right)\right)\end{aligned}$$

The feasible estimate of the posterior quality vector is given by equations (A9) and (A10) with unknown parameters replaced by their estimates ($\hat{\Lambda}$ is constructed from $\hat{\boldsymbol{\lambda}}$ by splicing the vector back into matrix form):

$$\begin{aligned}\mathbf{q}_h^{EB(f)} &= \left(I-\hat{B}_h\right)\hat{\mathbf{q}}_h + \hat{B}_h\hat{\Lambda}\mathbf{x}_h \\ \hat{B}_h &= \hat{\Pi}_h\left(\hat{\Pi}_h+\hat{\Sigma}\right)^{-1}\end{aligned}$$

C.4.2 Implementation

We perform bivariate EB adjustment in two cases. The first is our multivariate allocation regressions, where we regress hospital size and growth on risk-adjusted survival and readmission simultaneously (see the discussion of Section IV.A and Appendix Table A6). The second is our regressions that study allocation with respect to productivity decomposed into risk-adjusted survival and risk-adjusted spending (see Section IV.C and Table 7). In these cases, we allow the underlying mean of each quality measure to equal a market fixed effect, so e.g. $\Lambda\mathbf{x}_h = \begin{pmatrix}\tau_M^1 \\ \tau_M^2\end{pmatrix}$. As in the approach with a single quality metric, $\mathbf{x}_h$ is a vector of 306 indicators for whether a hospital is in each of the 306 markets.

We perform the EB procedure, extracting the matrix of underlying market means $\hat{\Lambda}$ and underlying variance $\hat{\Sigma}$, then producing EB-adjusted quality vectors $\mathbf{q}_h^{EB(f)}$. The $\mathbf{q}_h^{EB(f)}$ become regressors in the multivariate allocation and decomposed productivity regressions replacing the noisy estimates $\hat{\mathbf{q}}_h$. By the result given in section C.3.3 (replacing scalars q_h and η_h with vectors $\mathbf{q}_h$ and $\boldsymbol{\eta}_h$), the EB adjustment restores consistency to the coefficients of interest in these regressions.

C.5 Comparison of estimates

We run all of our baseline regression analyses with the EB-adjusted productivities $q_h^{EB(f)}$. Appendix Table A2 explores the impact of the EB correction on our main results, reproducing the EB-adjusted main results from Table 4 without the EB correction.

To produce the uncorrected allocation metrics, we use the estimates $\hat{q}_h$ rather than $q_h^{EB(f)}$ in our regressions. Due to measurement error in the estimates, we generally expect the allocation

metrics computed without the EB correction to be attenuated. The attenuation result is well-known under classical measurement error, though when measurement error is non-classical it is possible for coefficients to be expanded rather than attenuated. Our EB approach allows each hospital's measurement error to be different (based on the squared standard error of the hospital fixed effect from the first-step regression), so it is robust to this violation of the classical measurement error assumption.

The results show that the EB correction has a substantial effect on our baseline estimates, and our findings are consistent with measurement error causing attenuation. Comparing our baseline (EB-adjusted) estimates to the un-adjusted versions, we see that the allocation gradients are substantially greater with the correction. For example, in column 1 of Appendix Table A2, a 1 percentage point rise in risk-adjusted AMI survival is associated with 17% more patients when the EB correction is used, but only 7% more patients when we use the raw quality metric. The expansion of the coefficient is more substantial for some of the other metrics – e.g. a 1 percentage point fall in risk-adjusted HF readmission is associated with 10% more patients in our baseline analysis, but only 1% more patients when we drop the EB correction.

A quantitatively large impact of the EB correction (i.e. a large amount of measurement error, to the extent that it is classical) is not surprising in light of results from other applications. For example, looking at estimates of teacher fixed effects in value added regressions, Jacob and Lefgren (2007) estimate a ratio of the unadjusted standard deviation to the EB-adjusted estimate of the standard deviation of about 1.3 to 1.6. We find ratios ranging from 1.5 to 2.1.

References (Additional for Appendices)

Diggle, Peter J., Patrick Heagerty, Kung-Yee Liang, and Scott L. Zeger. 2002. *Analysis of Longitudinal Data. Oxford Statistical Science Series*, Oxford University Press.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5): pp. 1127–1150.

MedPAC. 2011. "Hospital Acute Inpatient Services Payment System." http://www.medpac.gov/documents/MedPAC_Payment_Basics_11_hospital.pdf.

Murphy, Kevin P. 2007. "Conjugate Bayesian analysis of the Gaussian distribution." http://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf.

Sloan, Frank Allen. 2000. "Not-for-profit ownership and hospital behavior." In *Handbook of Health Economics*. Vol. 1, ed. Anthony J. Culyer and Joseph P. Newhouse, Chapter 21, 1141–1174. Elsevier.

Tables (Additional for Appendices)

Table A1 - Sensitivity of Static Allocation Results to Truncation

Condition	(1)	(2) AMI	(3)	(4)	(5) Heart Failure	(6)	(7)	(8) Pneumonia	(9)	(10)	(11) Hip/Knee Replacement	(12)
Measure \ Model	Baseline	Trunc OLS	Trunc MLE	Baseline	Trunc OLS	Trunc MLE	Baseline	Trunc OLS	Trunc MLE	Baseline	Trunc OLS	Trunc MLE
Risk-Adjusted Survival	17.496	10.075	14.050	15.360	11.714	16.950	5.140	4.299	5.748			
	(0.995)	(0.750)	(1.085)	(1.320)	(1.120)	(1.713)	(0.777)	(0.654)	(0.884)			
Hospitals	2,890	1,910	1,910	4,023	2,918	2,918	4,325	3,578	3,578			
Risk-Adjusted Readmission	-9.162	-7.639	-10.556	-10.346	-9.993	-13.895	0.499	-1.093	-1.410	-21.037	-16.078	-22.119
	(1.621)	(1.449)	(2.037)	(1.782)	(1.617)	(2.335)	(1.575)	(1.302)	(1.680)	(2.027)	(1.781)	(2.536)
Hospitals	2,322	1,909	1,909	3,904	2,918	2,918	4,264	3,575	3,575	2,632	2,108	2,108
Process of Care Z-Score	0.319	0.285	0.493	0.332	0.230	0.378	0.211	0.151	0.217			
	(0.026)	(0.022)	(0.035)	(0.016)	(0.015)	(0.030)	(0.015)	(0.014)	(0.019)			
Hospitals	2,398	1,893	1,893	3,666	2,892	2,892	3,920	3,450	3,450			
Patient Survey Z-Score	-0.321	0.054	0.077	-0.252	-0.064	-0.088	-0.210	-0.092	-0.112	0.057	0.071	0.102
	(0.052)	(0.027)	(0.040)	(0.038)	(0.026)	(0.035)	(0.030)	(0.023)	(0.029)	(0.051)	(0.029)	(0.042)
Hospitals	3,498	1,901	1,901	3,598	2,808	2,808	3,610	3,197	3,197	3,061	2,104	2,104

This table repeats the static allocation analysis of Table 4 and shows its sensitivity to truncation in patient counts. For each condition, the left column ("Baseline") repeats our baseline results. The middle column ("Trunc OLS") shows static allocation limiting to hospitals with at least 25 patients in 2008, i.e. truncating at 25 patients; this specification attenuates results, as is generally the case when outcomes are truncated. The right column ("Trunc MLE") shows results from running a Tobit-style maximum-likelihood truncated regression on the "Trunc OLS" sample. This model adjusts for the truncation under a normality assumption; as expected, this correction always expands coefficients beyond the truncated OLS model and often beyond our baseline results. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A2 - Sensitivity of Allocation Results to Empirical Bayes Adjustment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation				Dynamic Allocation			
Condition	AMI	HF	Pneu	Hip/Knee	AMI	HF	Pneu	Hip/Knee
Panel A - Risk-Adjusted Survival								
Baseline (EB-Adjusted)	17.496	15.360	5.140		1.533	0.774	1.220	
	(0.995)	(1.320)	(0.777)		(0.379)	(0.501)	(0.354)	
Raw (No EB Adjustment)	6.833	3.761	1.957		0.645	0.084	0.340	
	(0.342)	(0.425)	(0.403)		(0.175)	(0.199)	(0.175)	
Hospitals	2,890	4,023	4,325		2,890	4,023	4,325	
Raw SD / Corrected SD	1.597	1.788	1.547		1.597	1.788	1.547	
Panel B - Risk-Adjusted Readmission								
Baseline (EB-Adjusted)	-9.162	-10.346	0.499	-21.037	-1.428	-2.300	-1.138	-1.112
	(1.621)	(1.782)	(1.575)	(2.027)	(0.611)	(0.651)	(0.679)	(0.836)
Raw (No EB Adjustment)	-1.699	-1.043	0.755	-6.492	-0.307	-0.217	-0.189	-0.431
	(0.395)	(0.346)	(0.427)	(0.727)	(0.197)	(0.162)	(0.195)	(0.382)
Hospitals	2,322	3,904	4,264	2,632	2,322	3,904	4,264	2,632
Raw SD / Corrected SD	1.870	2.132	1.864	1.794	1.870	2.132	1.864	1.794

This table shows the sensitivity of the allocation results of Table 4 to the empirical Bayes adjustment procedure. In each panel, we first repeat the baseline allocation results in which the quality metric is empirical-Bayes-adjusted. We then show the same allocation models using the raw quality metric without empirical Bayes adjustment. Lastly, we show the ratio of the raw standard deviation of the quality measure to its standard deviation after correcting for measurement error (see Appendix Section C.3.1). Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A3 - Distance Traveled across Conditions

	(1)	(2)	(3)	(5)
Condition	AMI	Heart Failure	Pneumonia	Hip Fracture
Median miles traveled	7.0	5.4	5.2	9.1
Mean miles traveled	45.0	33.7	35.8	41.9
Share treated at nearest hospital	0.43	0.52	0.56	0.38
Share staying in market	0.87	0.89	0.90	0.84
Index Events	190,189	308,122	354,319	267,557
Hospitals	2,890	4,023	4,325	2,632

Distances are miles from the centroid of the patient's ZIP code to the centroid of the hospital's ZIP code. The sample is patients in 2008 at hospitals that had a valid risk-adjusted survival rate (risk-adjusted readmission for hip/knee replacement).

Table A4 - Correlation of Measures Across Conditions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Risk-Adjusted Survival			Risk-Adjusted Readmission			
Condition	AMI	HF	Pneu	AMI	HF	Pneu	Hip/Knee
AMI	1.00			1.00			
	[2,890]			[2,322]			
HF	0.58	1.00		0.76	1.00		
	[2,888]	[4,023]		[2,320]	[3,904]		
Pneumonia	0.45	0.70	1.00	0.66	0.94	1.00	
	[2,883]	[4,006]	[4,325]	[2,313]	[3,880]	[4,264]	
Hip/Knee Replacement				0.44	0.51	0.44	1.00
				[2,020]	[2,524]	[2,544]	[2,632]

	Process of Care		
Condition	AMI	HF	Pneu
AMI	1.00		
	[2,398]		
HF	0.59	1.00	
	[2,397]	[3,666]	
Pneumonia	0.44	0.63	1.00
	[2,386]	[3,637]	[3,920]

Hospitals used to calculate correlation in brackets. Correlations involving survival or readmission are adjusted for measurement error (see Appendix Section C.3.2).

Table A5 - Correlation of Management Scores and Quality

Measure \ Condition	(1) AMI	(2) HF	(3) Pneu	(4) Hip/Knee
Risk-Adjusted Survival	0.19	0.21	0.17	
	[198]	[274]	[289]	
Risk-Adjusted Readmission	0.09	-0.09	0.45	0.21
	[158]	[267]	[288]	[190]
Process of Care Z-Score	0.12	0.35	0.28	
	[162]	[244]	[263]	
Patient Survey Z-Score	-0.01	-0.02	-0.02	0.05
	[227]	[233]	[233]	[202]

Each cell shows the correlation between our measure and the Bloom-Van Reenen average management score at the hospital. Hospitals used to calculate correlation in brackets. Correlations involving survival or readmission are adjusted for measurement error (see Appendix Section C.3.2).

Table A6 - Allocation across Conditions - Multivariate Approach to Measuring Quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation				Dynamic Allocation			
Measure \ Condition	AMI	HF	Pneu	Hip/Knee	AMI	HF	Pneu	Hip/Knee
Risk-Adjusted Survival	12.866	17.839	4.237		0.217	1.827	1.269	
	(1.421)	(1.902)	(0.764)		(0.600)	(0.586)	(0.357)	
Risk-Adjusted Readmission	-6.108	-14.155	-1.120	-20.716	-0.814	-2.264	-0.580	-0.757
	(1.983)	(2.214)	(1.552)	(2.108)	(0.745)	(0.668)	(0.644)	(0.845)
Process of Care Z-Score	0.196	0.188	0.196		0.036	0.014	0.015	
	(0.027)	(0.022)	(0.014)		(0.011)	(0.010)	(0.009)	
Patient Survey Z-Score	-0.113	-0.292	-0.166	-0.022	0.002	-0.014	-0.004	0.025
	(0.040)	(0.028)	(0.022)	(0.034)	(0.013)	(0.009)	(0.009)	(0.016)
Hospitals	2,193	3,316	3,454	2,542	2,193	3,316	3,454	2,542

This table repeats the analysis of Table 4 but uses all available quality measures for the condition at once (plus the patient survey, which is not condition-specific). The allocation sample for each regression is all hospitals with the displayed quality measures and at least one patient in 2008. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A7 - Allocation across Conditions - Constant Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation				Dynamic Allocation			
Measure \ Condition	AMI	HF	Pneu	Hip/Knee	AMI	HF	Pneu	Hip/Knee
Risk-Adjusted Survival	12.082	13.870	4.192		0.526	1.364	1.170	
	(0.929)	(1.296)	(0.745)		(0.401)	(0.452)	(0.323)	
Hospitals	2,193	3,316	3,454		2,193	3,316	3,454	
Risk-Adjusted Readmission	-9.909	-10.519	-1.372	-20.742	-1.127	-1.969	-0.733	-0.988
	(1.600)	(1.686)	(1.562)	(2.097)	(0.567)	(0.531)	(0.561)	(0.835)
Hospitals	2,193	3,316	3,454	2,542	2,193	3,316	3,454	2,542
Process of Care Z-Score	0.327	0.300	0.185		0.045	0.030	0.018	
	(0.023)	(0.018)	(0.015)		(0.009)	(0.008)	(0.008)	
Hospitals	2,193	3,316	3,454		2,193	3,316	3,454	
Patient Survey Z-Score	0.074	-0.172	-0.121	0.068	0.017	0.004	0.004	0.029
	(0.030)	(0.031)	(0.028)	(0.035)	(0.012)	(0.007)	(0.008)	(0.015)
Hospitals	2,193	3,316	3,454	2,542	2,193	3,316	3,454	2,542

This table repeats the analysis of Table 4 but uses a constant sample of hospitals across the quality metrics. The sample for each regression is all hospitals with a risk-adjusted survival rate (if calculated for that condition), risk-adjusted readmission rate, process of care score (if calculated for that condition), patient survey score, and at least one patient in 2008. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A8 - Allocation with Respect to All Condition-Specific Quality Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation				Dynamic Allocation			
Measure \ Condition	AMI	HF	Pneu	Hip/Knee	AMI	HF	Pneu	Hip/Knee
Risk-Adjusted Survival Z-Score								
AMI	0.553	0.335	0.183		0.048	0.014	0.031	
	(0.030)	(0.022)	(0.021)		(0.014)	(0.009)	(0.010)	
HF	0.198	0.124	0.031		-0.001	0.010	0.004	
	(0.036)	(0.030)	(0.026)		(0.017)	(0.011)	(0.012)	
Pneumonia	0.034	0.080	0.052		0.014	0.010	0.032	
	(0.034)	(0.024)	(0.019)		(0.014)	(0.013)	(0.011)	
Hospitals	2,882	2,882	2,882		2,882	2,882	2,882	
Risk-Adjusted Readmission Z-Score								
AMI	-0.171	-0.112	-0.122	-0.265	-0.031	-0.035	-0.018	-0.001
	(0.047)	(0.038)	(0.035)	(0.056)	(0.019)	(0.014)	(0.015)	(0.023)
HF	-0.215	-0.146	-0.086	-0.282	-0.027	-0.043	-0.028	-0.007
	(0.044)	(0.031)	(0.030)	(0.054)	(0.019)	(0.017)	(0.015)	(0.021)
Pneumonia	0.104	0.084	0.041	-0.055	-0.006	-0.014	-0.018	-0.008
	(0.042)	(0.031)	(0.027)	(0.050)	(0.019)	(0.017)	(0.015)	(0.024)
Hip/Knee Replacement	-0.079	-0.067	-0.086	-0.284	-0.010	-0.016	0.002	-0.024
	(0.034)	(0.024)	(0.021)	(0.041)	(0.013)	(0.012)	(0.012)	(0.016)
Hospitals	2,018	2,018	2,018	2,018	2,018	2,018	2,018	2,018
Process of Care Z-Score								
AMI	0.336	0.209	0.165		0.058	0.046	0.014	
	(0.031)	(0.021)	(0.020)		(0.013)	(0.011)	(0.011)	
HF	-0.054	-0.103	-0.219		-0.061	-0.027	0.001	
	(0.049)	(0.040)	(0.042)		(0.025)	(0.020)	(0.018)	
Pneumonia	0.001	0.094	0.205		0.036	0.026	0.025	
	(0.051)	(0.041)	(0.037)		(0.020)	(0.020)	(0.016)	
Hospitals	2,386	2,389	2,390		2,386	2,389	2,390	

This table repeats the analyses of Table 4 but rather than using the one (non-standardized, except for process of care) quality measure for the left-hand side condition, it uses Z-scores of all condition specific quality measures. Each column in a panel represents one regression, e.g. the top panel of column (1) regresses hospital size for AMI on risk-adjusted survival Z-scores for AMI, HF, and pneumonia. Standard errors are analytic and clustered at the market level.

Table A9 - Sensitivity of Static Allocation to Poisson Regression Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Condition	AMI		HF		Pneumonia		Hip/Knee	
Measure \ Model	Baseline	Poisson	Baseline	Poisson	Baseline	Poisson	Baseline	Poisson
Risk-Adjusted Survival	17.496	18.631	15.360	17.532	5.140	6.478		
	(0.995)	(1.109)	(1.320)	(1.504)	(0.777)	(0.829)		
Hospitals	2,890	2,881	4,023	4,023	4,325	4,325		
Risk-Adjusted Readmission	-9.162	-10.685	-10.346	-13.145	0.499	-1.922	-21.037	-25.811
	(1.621)	(2.086)	(1.782)	(2.175)	(1.575)	(1.756)	(2.027)	(3.466)
Hospitals	2,322	2,304	3,904	3,903	4,264	4,264	2,632	2,626
Process of Care Z-Score	0.319	0.422	0.332	0.332	0.211	0.199		
	(0.026)	(0.024)	(0.016)	(0.020)	(0.015)	(0.014)		
Hospitals	2,398	2,379	3,666	3,665	3,920	3,920		
Patient Survey Z-Score	-0.321	-0.131	-0.252	-0.123	-0.210	-0.114	0.057	0.130
	(0.052)	(0.035)	(0.038)	(0.030)	(0.030)	(0.023)	(0.051)	(0.057)
Hospitals	3,498	3,498	3,598	3,598	3,610	3,610	3,061	3,059

This table shows the baseline static allocation results of Table 4 in comparison to the same model run as a fixed effects Poisson regression. To make the models analogous, the Poisson regressand is the count of patients, not its logarithm. Standard errors are bootstrapped with 300 replications and are clustered at the market level.
The baseline sample is used; hospital counts can be smaller for the Poisson models because they exclude markets with only one hospital.

Table A10 - Sensitivity of Allocation Results to Risk Adjustment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Static Allocation				Dynamic Allocation			
Condition	AMI	HF	Pneu	Hip/Knee	AMI	HF	Pneu	Hip/Knee
Panel A - Risk-Adjusted Survival								
Baseline Risk-Adjustment	17.496	15.360	5.140		1.533	0.774	1.220	
	(0.995)	(1.320)	(0.777)		(0.379)	(0.501)	(0.354)	
Age/Race/Sex Only	16.898	14.798	3.209		1.565	0.654	1.065	
	(0.847)	(1.377)	(0.764)		(0.337)	(0.508)	(0.361)	
No Risk Adjustment	14.896	13.763	3.625		1.520	0.765	1.315	
	(0.571)	(1.192)	(0.713)		(0.241)	(0.420)	(0.332)	
Hospitals	2,890	4,023	4,325		2,890	4,023	4,325	
Panel B - Risk-Adjusted Readmission								
Baseline Risk-Adjustment	-9.162	-10.346	0.499	-21.037	-1.428	-2.300	-1.138	-1.112
	(1.621)	(1.782)	(1.575)	(2.027)	(0.611)	(0.651)	(0.679)	(0.836)
Age/Race/Sex Only	-10.358	-6.466	3.152	-20.023	-1.556	-1.621	-0.547	-1.048
	(1.212)	(1.457)	(1.280)	(1.859)	(0.475)	(0.528)	(0.482)	(0.767)
No Risk Adjustment	-10.909	-5.596	2.753	-20.710	-1.601	-1.514	-0.585	-1.140
	(1.096)	(1.441)	(1.272)	(1.819)	(0.435)	(0.511)	(0.465)	(0.732)
Hospitals	2,322	3,904	4,264	2,632	2,322	3,904	4,264	2,632

This table shows the sensitivity of the allocation results of Table 4 to the risk-adjustment procedure. In each panel, we repeat the baseline allocation results with full risk adjustment, followed by identical specifications under two alternative measures. First, we risk-adjust using only age-race-sex interactions. Second, we perform no risk-adjustment. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A11 - Static Allocation Restricted to Patients Treated in Market of Residence (Choice Model Subsample)

	(1)	(2)	(4)	(5)	(7)	(8)	(10)	(11)
Condition	AMI		Heart Failure		Pneumonia		Hip/Knee	
Sample	Baseline	Choice	Baseline	Choice	Baseline	Choice	Baseline	Choice
Share Baseline Patients in Sample	1.00	0.87	1.00	0.89	1.00	0.90	1.00	0.84
Risk-Adjusted Survival	17.496	16.769	15.360	15.184	5.140	5.402		
	(0.995)	(0.988)	(1.320)	(1.341)	(0.777)	(0.808)		
Hospitals	2,890	2,889	4,023	4,023	4,325	4,324		
Risk-Adjusted Readmission	-9.162	-9.447	-10.346	-10.561	0.499	-0.233	-21.037	-20.794
	(1.621)	(1.600)	(1.782)	(1.823)	(1.575)	(1.590)	(2.027)	(2.014)
Hospitals	2,322	2,322	3,904	3,904	4,264	4,263	2,632	2,632
Process of Care Z-Score	0.319	0.317	0.332	0.332	0.211	0.210		
	(0.026)	(0.026)	(0.016)	(0.016)	(0.015)	(0.016)		
Hospitals	2,398	2,397	3,666	3,662	3,920	3,918		
Patient Survey Z-Score	-0.321	-0.316	-0.252	-0.261	-0.210	-0.217	0.057	0.041
	(0.052)	(0.051)	(0.038)	(0.039)	(0.030)	(0.030)	(0.051)	(0.048)
Hospitals	3,498	3,480	3,598	3,594	3,610	3,608	3,061	3,046

This table repeats the static allocation analysis of Table 4 and shows how it is affected by restricting to patients who were treated in their market of residence, which is the patients who were included in the choice model. For each condition, the left column (Baseline) repeats our baseline results. The right column (Choice) runs the same regression but only counts patients residing in the hospital's market. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A12 - Choice Model of Patient Allocation - Raw Logit Coefficients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Condition	AMI	AMI	AMI	AMI	HF	HF	HF	HF
Mean miles to chosen hospital	12.48	12.67	12.65	12.45	8.27	8.27	8.30	8.35
SD miles to chosen hospital	20.06	20.27	20.27	20.31	13.25	13.25	13.27	13.33
Distance	-0.111	-0.128	-0.137	-0.099	-0.159	-0.163	-0.160	-0.159
	(0.006)	(0.007)	(0.007)	(0.005)	(0.006)	(0.007)	(0.007)	(0.007)
Distance^2	0.00019	0.00032	0.00034	0.00012	0.00018	0.00026	0.00013	0.00013
	(0.00003)	(0.00003)	(0.00003)	(0.00001)	(0.00001)	(0.00002)	(0.00001)	(0.00001)
Risk-Adjusted Survival	19.004				16.041			
	(1.147)				(1.728)			
Risk-Adjusted Readmission		-13.626				-16.491		
		(2.007)				(1.808)		
Process of Care Z-Score			0.568				0.353	
			(0.036)				(0.028)	
Patient Survey Z-Score				-0.031				0.015
				(0.037)				(0.032)
Patients	165,005	158,086	158,032	167,429	275,671	274,667	270,773	266,915
Observations	2,869,091	2,321,684	2,427,869	3,359,387	6,241,586	6,103,120	5,811,375	5,532,403
Avg Hospital Choices/Patient	17.39	14.69	15.36	20.06	22.64	22.22	21.46	20.73

This table shows the raw logit coefficients (log odds ratios) from the models of Table 5. See that table for sample restrictions. Standard errors are analytic and clustered at the market level.

Table A12 Continued - Choice Model of Patient Allocation - Raw Logit Coefficients

	(9)	(10)	(11)	(12)	(13)	(14)
Condition	Pneu	Pneu	Pneu	Pneu	Hip/Knee	Hip/Knee
Mean miles to chosen hospital	7.49	7.49	7.50	7.54	13.16	13.09
SD miles to chosen hospital	11.92	11.91	11.77	11.77	18.85	18.80
Distance	-0.178	-0.178	-0.181	-0.185	-0.105	-0.101
	(0.006)	(0.006)	(0.006)	(0.007)	(0.004)	(0.004)
Distance^2	0.00013	0.00013	0.00013	0.00013	0.00017	0.00012
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00002)	(0.00001)
Risk-Adjusted Survival	6.647					
	(0.962)					
Risk-Adjusted Readmission		-7.927			-24.091	
		(1.979)			(2.570)	
Process of Care Z-Score			0.238			
			(0.018)			
Patient Survey Z-Score				-0.007		0.157
				(0.028)		(0.039)
Patients	317,904	317,374	309,623	298,185	222,673	224,451
Observations	7,766,357	7,666,146	6,997,264	6,233,133	3,422,903	4,017,558
Avg Hospital Choices/Patient	24.43	24.15	22.60	20.90	15.37	17.90

This table shows the raw logit coefficients (log odds ratios) from the models of Table 5. See that table for sample restrictions. Standard errors are analytic and clustered at the market level.

Table A13 - Static and Dynamic Allocation over Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	AMI					HF				
Measure \ Year	1996	1999	2002	2005	2008	1996	1999	2002	2005	2008
Static Allocation	11.463	13.759	14.201	16.082	17.496	2.741	2.389	1.118	9.222	15.360
	(0.838)	(1.008)	(1.058)	(0.815)	(0.824)	(1.117)	(1.423)	(1.580)	(1.369)	(1.348)
Dynamic Allocation	0.536	0.875	0.787	1.103	1.533	0.609	-0.095	0.580	0.839	0.774
	(0.341)	(0.475)	(0.422)	(0.322)	(0.380)	(0.525)	(0.544)	(0.596)	(0.420)	(0.483)
Hospitals	3,708	3,558	3,487	3,267	2,890	4,361	4,254	4,176	4,130	4,023
	Pneumonia									
Measure \ Year	1996	1999	2002	2005	2008					
Static Allocation	-2.423	-2.723	-2.690	3.763	5.140					
	(0.750)	(0.778)	(0.845)	(0.800)	(0.768)					
Dynamic Allocation	-0.143	-0.608	0.181	-0.124	1.220					
	(0.314)	(0.365)	(0.288)	(0.331)	(0.363)					
Hospitals	4,565	4,496	4,403	4,375	4,325					

This table shows the allocation relationships with respect to risk-adjusted survival over time. Risk-adjusted survival is calculated from a regression of survival on hospital fixed effects and patient risk-adjusters. A separate regression is run for each of the year groups 1994-1996, 1997-1999, 2000-2002, 2003-2005, and 2006-2008. The empirical-Bayes-adjusted hospital effects are considered the hospital's quality metric for the year group's terminal year, the static allocation analysis is based on patient allocation in the terminal year, and the dynamic allocation is based on growth between the terminal year and the subsequent 2 years. Thus, for example, the "1996" analysis measures quality over 1994-1996, static allocation in 1996, and dynamic allocation based on growth between 1996 and 1998. Standard errors are analytic and clustered at the market level.

Table A14 - Allocation for Non-ED Non-Transfer Patients across Conditions

	(1)	(2)	(3)	(4)	(5)	(6)
	Static Allocation			Dynamic Allocation		
Measure \ Condition	AMI	HF	Pneu	AMI	HF	Pneu
Risk-Adjusted Survival	14.840	19.870	3.440	5.381	-0.125	0.637
	(1.215)	(2.093)	(1.159)	(1.071)	(0.993)	(0.720)
Hospitals	2,879	4,023	4,325	2,320	3,924	4,246
Risk-Adjusted Readmission	-6.086	-17.273	-3.285	-6.180	-0.828	0.330
	(3.831)	(2.469)	(1.900)	(1.695)	(1.324)	(1.173)
Hospitals	2,302	3,903	4,264	1,953	3,811	4,191
Process of Care Z-Score	0.317	0.173	0.015	0.151	0.038	0.023
	(0.046)	(0.018)	(0.017)	(0.028)	(0.015)	(0.012)
Hospitals	2,377	3,665	3,920	1,944	3,541	3,819
Patient Survey Z-Score	-0.039	-0.057	0.001	-0.103	0.001	-0.001
	(0.037)	(0.027)	(0.023)	(0.035)	(0.020)	(0.018)
Hospitals	3,498	3,598	3,610	2,653	3,472	3,513

This table repeats the analysis of Table 4 but considers hospital size and growth counting only non-ED non-transfer patients (the omitted category of patients from the analysis of Table 9). Static allocation uses the Poisson model (see Appendix Table A9) and the baseline allocation sample. Dynamic allocation uses the subset of hospitals with at least one non-ED non-transfer patient in 2008. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A15 - Sensitivity of ED and Non-ED Transfer Patient Static Allocation to Poisson Regression Model

	(1)	(2)	(3)	(4)	(5)	(6)
Condition	AMI		Heart Failure		Pneumonia	
Source of Admission	ED	Transfer	ED	Transfer	ED	Transfer
Risk-Adjusted Survival						
Baseline Static Allocation	14.377	29.264	15.254	24.036	7.720	5.658
	(0.878)	(2.184)	(1.693)	(2.566)	(1.387)	(1.658)
P-value of test for equality	0.000		0.001		0.222	
Hospitals	2,742	1,408	3,329	1,563	3,513	1,601
Poisson Static Allocation	14.489	42.532	15.727	50.673	7.168	14.049
	(1.022)	(2.609)	(1.586)	(4.664)	(0.983)	(2.941)
P-value of test for equality	0.000		0.000		0.009	
Hospitals	2,881	2,881	4,023	4,011	4,325	4,275
Risk-Adjusted Readmission						
Baseline Static Allocation	-7.903	-13.315	-10.690	-16.527	0.204	2.008
	(1.295)	(3.904)	(1.775)	(3.483)	(2.314)	(3.107)
P-value of test for equality	0.109		0.062		0.585	
Hospitals	2,276	1,361	3,277	1,550	3,488	1,592
Poisson Static Allocation	-8.128	-25.550	-11.265	-37.988	-1.647	1.089
	(1.730)	(5.921)	(2.329)	(6.744)	(2.021)	(6.252)
P-value of test for equality	0.001		0.000		0.653	
Hospitals	2,304	2,304	3,903	3,892	4,264	4,214
Process of Care Z-Score						
Baseline Static Allocation	0.272	0.874	0.370	0.325	0.307	0.078
	(0.023)	(0.062)	(0.024)	(0.036)	(0.030)	(0.035)
P-value of test for equality	0.000		0.233		0.000	
Hospitals	2,369	1,376	3,245	1,520	3,389	1,548
Poisson Static Allocation	0.326	1.179	0.377	0.754	0.262	0.214
	(0.021)	(0.090)	(0.025)	(0.058)	(0.018)	(0.043)
P-value of test for equality	0.000		0.000		0.261	
Hospitals	2,379	2,379	3,665	3,653	3,920	3,869
Patient Survey Z-Score						
Baseline Static Allocation	-0.234	0.232	-0.227	0.045	-0.218	-0.087
	(0.049)	(0.067)	(0.041)	(0.046)	(0.041)	(0.039)
P-value of test for equality	0.000		0.000		0.008	
Hospitals	3,116	1,438	3,214	1,496	3,257	1,484
Poisson Static Allocation	-0.157	-0.034	-0.141	-0.090	-0.137	-0.203
	(0.035)	(0.072)	(0.032)	(0.060)	(0.028)	(0.057)
P-value of test for equality	0.051		0.349		0.170	
Hospitals	3,498	3,498	3,598	3,586	3,610	3,559

This table shows the static allocation results for ED and non-ED transferred patients using Poisson regression (as in Table 9) and linear regression. The left hand side of these regressions considers hospital size counting only ED patients in the odd-numbered columns and only non-ED transferred patients in the even-numbered columns. To make the linear and Poisson models analogous, the Poisson regressand is the count of patients, not its logarithm. Both approaches include market fixed effects. Hospital counts can be smaller for the Poisson models because they exclude markets with only one hospital. In addition, in the Poisson models, hospital counts may differ between ED and non-ED transfers for the same condition and quality measure because the counts also exclude markets with no variation in the outcome (e.g. all zeroes). Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A16 - Choice Model of Patient Allocation for ED and Non-ED Transfer Patients across Conditions

	(1)	(2)	(3)	(4)	(5)	(6)
Condition	AMI		Heart Failure		Pneumonia	
Source of Admission	ED	Transfer	ED	Transfer	ED	Transfer
Share of patients in 2008	0.79	0.14	0.76	0.02	0.77	0.01
Median miles to chosen hospital	4.7	29.8	4.4	26.3	4.4	21.1
Mean miles to chosen hospital	8.2	36.8	7.1	32.3	7.1	27.7
Share treated at nearest hospital	0.57	0.04	0.58	0.09	0.60	0.15
Risk-Adjusted Survival						
MRS(1 pp risk-adjusted survival, miles)	-0.744	-16.061	-0.623	-16.642	-0.317	-4.014
	(0.075)	(1.423)	(0.103)	(1.601)	(0.054)	(0.777)
P-value of test for equality	0.000		0.000		0.000	
Patients	129,889	23,185	209,094	6,130	244,358	4,097
Risk-Adjusted Readmission						
MRS(1 pp risk-adjusted readmission, miles)	0.527	10.818	0.734	12.774	0.410	1.303
	(0.086)	(2.097)	(0.104)	(2.193)	(0.102)	(1.713)
P-value of test for equality	0.000		0.000		0.598	
Patients	124,707	22,947	208,842	6,115	244,181	4,088
Process of Care Z-Score						
MRS(1 SD process of care, miles)	-2.024	-36.431	-1.712	-25.332	-1.392	-6.321
	(0.191)	(3.468)	(0.212)	(2.521)	(0.122)	(1.545)
P-value of test for equality	0.000		0.000		0.001	
Patients	124,989	22,913	208,503	6,083	243,368	4,025
Patient Survey Z-Score						
MRS(1 SD patient survey, miles)	-0.052	0.090	-0.179	2.178	-0.024	4.896
	(0.217)	(3.591)	(0.172)	(2.092)	(0.144)	(1.682)
P-value of test for equality	0.968		0.250		0.002	
Patients	131,603	23,281	207,472	6,059	241,158	3,948
Distance MRSs were evaluated at	12.48	12.48	8.27	8.27	7.49	7.49

These regressions repeat the conditional logit choice models of Table 5 but are restricted to ED patients (in odd columns) and non-ED transfer patients (in even columns). Standard errors are analytic and clustered at the market level.

This table reports the marginal rates of substitution (MRSs) of quality for distance derived from the conditional logit model (see equation 6). For the survival and readmission rates, the MRS given by equation (6) is divided by 100 to put it into percentage point terms. MRSs for a condition are evaluated at the same distance that was used for the condition in Table 5.

The sample is ED patients (odd columns) or non-ED transfer patients (even columns) with the condition in 2008 who stayed in their market of residence for treatment. The choice set for a patient is all hospitals in his market with the quality measure available that treated at least one patient in 2008.

Table A17 - Static Allocation with Patients Mechanically Allocated to Nearest Hospital

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Condition	AMI		HF		Pneumonia		Hip/Knee	
Measure \ Method	Baseline	Mech	Baseline	Mech	Baseline	Mech	Baseline	Mech
Share truly going to closest	0.44		0.52		0.56		0.38	
Risk-Adjusted Survival	17.496	2.309	15.360	2.507	5.140	0.685		
	(0.995)	(0.577)	(1.320)	(1.115)	(0.777)	(0.653)		
Hospitals	2,890	2,888	4,023	4,021	4,325	4,322		
Risk-Adjusted Readmission	-9.162	-0.470	-10.346	-0.660	0.499	-0.328	-21.037	-2.783
	(1.621)	(1.111)	(1.782)	(1.305)	(1.575)	(1.130)	(2.027)	(1.140)
Hospitals	2,322	2,320	3,904	3,902	4,264	4,261	2,632	2,630
Process of Care Z-Score	0.319	0.018	0.332	0.158	0.211	0.099		
	(0.026)	(0.015)	(0.016)	(0.012)	(0.015)	(0.010)		
Hospitals	2,398	2,396	3,666	3,664	3,920	3,918		
Patient Survey Z-Score	-0.321	-0.203	-0.252	-0.238	-0.210	-0.182	0.057	-0.101
	(0.052)	(0.024)	(0.038)	(0.023)	(0.030)	(0.021)	(0.051)	(0.018)
Hospitals	3,498	3,496	3,598	3,595	3,610	3,607	3,061	3,058

This table shows our baseline static allocation results from Table 4 in comparison to an alternative allocation constructed by mechanically assigning each patient to his closest hospital. Only hospitals that treated at least one patient with the condition in 2008 are eligible for mechanical assignment. Distance is measured from the ZIP code centroid of the patient's residence to the ZIP code centroid of the hospital. The sample for each regression is all hospitals with the relevant quality measure and at least one mechanically allocated patient in 2008. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A18 - Dynamic Allocation with Patients Mechanically Allocated to Nearest Hospital

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Condition	AMI		HF		Pneumonia		Hip/Knee	
Measure \ Method	Baseline	Mech	Baseline	Mech	Baseline	Mech	Baseline	Mech
Share truly going to closest	0.44		0.52		0.56		0.38	
Risk-Adjusted Survival	1.533	-0.297	0.774	0.378	1.220	0.842		
	(0.379)	(0.250)	(0.501)	(0.497)	(0.354)	(0.356)		
Hospitals	2,890	2,888	4,023	4,021	4,325	4,322		
Risk-Adjusted Readmission	-1.428	-0.076	-2.300	-1.358	-1.138	-0.988	-1.112	-0.378
	(0.611)	(0.585)	(0.651)	(0.565)	(0.679)	(0.562)	(0.836)	(0.747)
Hospitals	2,322	2,320	3,904	3,902	4,264	4,261	2,632	2,630
Process of Care Z-Score	0.048	0.015	0.043	0.027	0.026	0.019		
	(0.010)	(0.010)	(0.009)	(0.008)	(0.009)	(0.009)		
Hospitals	2,398	2,396	3,666	3,664	3,920	3,918		
Patient Survey Z-Score	-0.065	-0.041	-0.003	0.000	0.007	-0.004	0.037	0.014
	(0.015)	(0.012)	(0.011)	(0.009)	(0.011)	(0.008)	(0.022)	(0.016)
Hospitals	3,498	3,496	3,598	3,595	3,610	3,607	3,061	3,058

This table shows our baseline dynamic allocation results from Table 4 in comparison to an alternative allocation constructed by mechanically assigning each patient to his closest hospital. Only hospitals that treated at least one patient with the condition in 2008 are eligible for mechanical assignment. Distance is measured from the ZIP code centroid of the patient's residence to the ZIP code centroid of the hospital. The sample for each regression is all hospitals with the relevant quality measure and at least one mechanically allocated patient in 2008. Standard errors are bootstrapped with 300 replications and are clustered at the market level.

Table A19 - Correlation with CMS Quality Measures

Measure \ Condition	(1) AMI	(2) HF	(3) Pneu
Risk-Adjusted Mortality	0.79	0.73	0.82
	[2,802]	[3,788]	[3,994]
Risk-Adjusted Readmission	0.66	0.67	0.71
	[2,254]	[3,681]	[3,924]

Each cell shows the correlation between our 2008 empirical-Bayes-adjusted quality measure and the CMS 2008 risk-standardized quality measure. We produce our risk-adjusted survival measure as risk-adjusted mortality to match the CMS measure. Hospitals used to calculate correlation in brackets.