MONASH University

MONASH BUSINESS SCHOOL

CENTRE FOR HEALTH ECONOMICS

WORKING PAPERS

Assessing the Quality of Public Services: For-profits, Chains, and Concentration in the Hospital Market

Discussion Paper no. 2023-01

Johannes S. Kunz, Carol Propper, Kevin E. Staub and Rainer Winkelmann

Keywords: Hospital Readmissions, Affordable Care Act, Hospital Quality, Competition, Hospital Chains

JEL Classification: H51, I11, I18

Johannes S. Kunz: Monash University (email: johannes.kunz@monash.edu); Carol Propper: Imperial College London (email: c.propper@imperial.ac.uk); Kevin E. Staub: University of Melbourne (email: kevin.staub@unimelb.edu.au); Rainer Winkelmann: University of Zurich (email: rainer.winkelmann@econ.uzh.ch).

Assessing the Quality of Public Services: For-profits, Chains, and Concentration in the Hospital Market*

Johannes S. Kunz — Monash University

Carol Propper — Imperial College London

Kevin E. Staub — University of Melbourne

Rainer Winkelmann — University of Zurich

November 7, 2022

Abstract

We examine variation in US hospital quality across ownership, chain membership, and market concentration. We use a new measure of quality derived from the penalties imposed on hospitals under the flagship Hospital Readmissions Reduction Program. We document a robust and sizable negative for-profit quality gap: for-profit hospitals are consistently of lower quality. We find that a substantial part of the gap is related to being located in less competitive markets. This reduction occurs most for hospitals that are part of large national chains. For such hospitals we find no quality gap in fully competitive markets.

Keywords: Hospital Readmissions; Affordable Care Act; Hospital Quality; Competition; Hospital Chains

JEL classifications: H51,I1,I11,I18

**Addresses for correspondence*: J. Kunz, Monash Business School, Centre for Health Economics, 900 Dandenong Road, 3145 Caulfield East, Vic, Australia, johannes.kunz@monash.edu.au; C. Propper, Imperial College London, Department of Economics and Public Policy, 414 City and Guilds Building, South Kensington Campus, United Kingdom, c.propper@imperial.ac.uk; K. Staub, The University of Melbourne, Department of Economics, 111 Barry Street, 3010 VIC, Australia, kevin.staub@unimelb.edu.au; R. Winkelmann, University of Zürich, Department of Economics, Zürichbergstrasse 14, CH-8039 Zürich, Switzerland, rainer.winkelmann@econ.uzh.ch.
Acknowledgments: We thank participants of the Econometrics Workshop (Melbourne), International Panel Data Conference (Thessaloniki), Australian Health Economics Society Conference (Sydney), iHEA (Basel), Econometrics Workshop (VU Wellington), Australasian Workshop on Econometrics and Health Economics (Darwin) as well as seminar participants at CINCH (Essen), Leibniz University Hannover, Monash University and the University of Zurich, for helpful comments. We also thank Daniel Avdic, Michael Darden, Joseph Doyle, Denzil Fiebig, Martin Gaynor, Martin Hackmann, Annika Herr, Joe Hirschberg, Andrew Jones, Jonathan Kolstad, Maarten Lindeboom, Jenny Lye, Janina Nemitz, Edward Norton, Arndt Reichert, Peter Sivey, Chris Skeels, Jonathan Skinner, Eric Sun, Ou Yang, and Jenny Williams for helpful comments.
Funding: Staub acknowledges funding from the Australian Research Council through grant DE170100644. Winkelmann acknowledges financial support by the Swiss National Science Foundation.
Conflict of interests: None of the authors has any conflict of interest to declare.

1 Introduction

Historically, the US hospital industry has been characterized by a large role for non-profit organizations. However, over time the share of patients treated in for-profit hospitals and hospitals that are part of systems (chains) has increased considerably.[1] From 1993 to 2017, the share of for-profits rose from 18% to over 26%, and chain hospitals experienced a similar increase. Thus, as the industry evolves, the long-standing question of whether for-profits and non-profits, in a market increasingly dominated by chains, behave differently has become highly salient (e.g. Beaulieu et al., 2020; Capps et al., 2020).

A large body of theoretical literature suggests that for-profits and non-profits have different objectives (a recent example focusing on public service providers such as hospitals is Besley and Malcomson, 2018). Early empirical studies showed mixed evidence for quality differences between for-profit and other hospitals.[2] More recent research has indicated that there exists a sizable raw quality gap between profit-oriented and other hospitals (e.g., Aswani et al., 2018; Gupta, 2017; Herrin et al., 2015; Jindal et al., 2018; Paul et al., 2020). While chains in the hospital sector have been less studied, Eliason et al. (2019) found that the quality of dialysis care decreased in hospitals after acquisition by a chain. But to date no study has devoted systematic attention to the interaction between a hospital's performance ownership, chain status, and market structure.

Our main contribution is to establish a new empirical fact that hospitals that are part of large for-profit chains exhibit significantly lower quality, which varies by local competition. This mirrors findings from earlier research that showed an important interaction between ownership and market structure: that is when for-profits and non-profits operate in competitive environments, they behaved similarly (Duggan, 2002; McClellan and Staiger, 2000; Sloan et al., 2001). This finding has been important. For example, much of the literature which has empirically examined the relationship between market structure and hospital quality has drawn on this, arguing that it can be assumed that non-profits behave as for-profits (e.g. Chandra et al., 2016; Zhang et al., 2016). But little is known about whether chain status affects these empirical regularities. Thus, given the large changes in the industry (chain hospitals now account for roughly 75% of hospitals, AHA, 2019), it is timely to return to the question of whether

[1] System is defined as either a multi-hospital or a diversified single hospital system. A multi-hospital system is two or more hospitals owned, leased, sponsored, or contract managed by a central organization (AHA, 2019); we will refer to them in the following as chains.

[2] Earlier literature which found a lack of difference between for-profit and non-profits focused largely on readmissions or mortality (Gaynor and Town, 2011), often based on acute myocardial infarction. Eggleston et al. (2008) provide a systematic review of 31 observational studies and find that studies representative of the US tend to report lower quality among for-profits than private non-profits. Most adverse effects are located in for-profits, but the size and significance depend on the populations and samples considered (cf. also, Baltagi and Yen, 2014; Colla et al., 2016; Cooper et al., 2018; Dranove and Satterthwaite, 2000; McClellan and Staiger, 2000; Paul et al., 2020; Picone et al., 2002).

non-profits have systematically different quality from for-profit hospitals and to focus directly on how this varies with both competition and with chain status.[3]

This paper provides rich descriptive evidence on this. We focus on publicly available measures observable to all players in the market for the period 2012-2016. Our preferred measure of quality is non-time varying and is the latent risk of being fined under a flagship program designed to incentivize hospital quality. This is the national Hospital Readmission Reduction Program (HRRP), one of the largest and most successful schemes linking financial penalties to hospital provider performance. It imposed financial penalties on hospitals that had annual (risk-adjusted) readmission rates for three emergency conditions above a threshold: acute myocardial infarction [AMI], heart failure [HF], and pneumonia [PN], for which over 70% of patients are admitted through the emergency room (Chandra et al., 2016, hereafter, 'emergency conditions' refers to these three conditions unless noted otherwise).

When considering hospital behavior, measures of quality that relate to incentivized metrics such as those in HRRP should reflect heterogeneity in strategic decisions (Mehta, 2019). This may differ by for-profit status. While penalties affect the trade-off between treatment cost and readmission probability, zero—or very few—readmissions may not be optimal from the point of view of hospital management. If the costs of avoiding readmissions are high, hospitals engaging in optimizing behavior will tolerate some penalties up to the point where marginal costs are equalized. Thus, systematic differences in the penalty likelihood between hospitals can be indicative of differences in the trade-off, or the way it is evaluated, between for-profits and non-profits. Competition may crowd out such strategic flexibility (Duggan, 2000; Gaynor et al., 2015; Moscelli et al., 2021).

Chain status means that hospitals are subject to central as well as local oversight. It is not immediately apparent how this might affect the difference between chain and non-chain hospitals with respect to the associations of for-profit status, local market competition, and quality. The market power derived from chain status may mean that chain hospitals can have lower quality than non-chain hospitals. However, this does not offer guidance as to the for-profit quality gap *among* chains. The negative quality gap between for and non-profits in the earlier literature was found to be driven out by competition. But again, it is not apparent the same will hold for chains. The greater national market power that chain hospitals have may insulate them from this local competition.[4] Thus, whether the chain's national market power allows chain for-profit hospitals to have lower quality than non-profit chains in highly competitive *local* markets remains an open empirical question in this—and other—markets (Benkard et al., 2021).

[3]The acceleration in the expansion of chains during the Covid pandemic has attracted attention from federal lawmakers pushing for greater oversight of hospitals (Times, 2021).

[4]Hausman and Lavetti (2021) find that the effects of local competition and chain level competition on prices can even have opposing signs in physician markets, though they do not study quality.

Our analyses show the following. For-profit hospitals have lower quality. This holds whether we examine our measure of the underlying penalty propensity, published measures of excess readmission rates, rates adjusted for risk, or (the unincentivized) measure of patient satisfaction.[5] These averages hide relevant heterogeneity. Local market concentration differences are associated with much of the negative quality difference between for-profits and non-profits. We estimate that 45% of the raw quality gap in our preferred quality measure can be attributed to a lack of competition at the mean market concentration. On average, chain status is not associated with the negative for-profit quality gap or the interaction with market structure. But for those hospitals which are part of a large chain (containing over 20 hospitals nationally),[6] the interaction with local market structure mutes this negative association between for-profit status and quality. Large chain for-profits and non-profits have very similar quality when operating in competitive local markets but have large quality gaps in more monopolistic local markets.

Patient selection could be one factor that could explain our findings. If sicker patients are treated in for-profit chain hospitals, they might appear to have lower observed quality, which, however might not correspond to 'true' quality differences. We select our measures of quality and undertake several specification tests to mitigate against patient selection by patients or hospitals. We focus on treatment for emergency patients. We use the published risk-adjusted measures and further control for patient risk, local area characteristics, and small samples. We examine within-market variation that covers similar populations. Additionally, selection by hospitals is unlikely in our setting due to the prerogative to accept emergency patients and the fact that hospitals do not observe penalty cut-offs for quality in the HRRP program (Gupta, 2021).[7] Finally, if for-profits were cream-skimming, this would bias our results away from finding higher quality in non-profits: we find the opposite.

Another potential explanation of our results is the sorting of hospitals into different markets. We address this by (1) controlling for the opening and closing of hospitals within the local market area in our risk adjustment measures, (2) focusing only on hospitals that did not change ownership in the sample period, and (3) using pre-sample measures of market structure. We also repeat our analysis constraining the sample to only hospitals which have not changed ownership in 10 years. These selections and adjustments ensure constant ownership and markets unaffected by potential acquisitions or market sorting. Our results are robust to all these checks. We also show that market concentration

[5]Patient satisfaction is technically part of an incentivized payment, but its weight is minimal.

[6]This is the top tercile of the size distribution of chains nationally.

[7]Further support for lack of patient selection in our context is given by Staiger (2018), who shows that comparing risk-adjusted estimates with those derived from (1) randomly or quasi-randomly experiments, (2) direct observation of surgical technical quality, (3) rich and complete data to control for comprehensive clinical information, and concludes that "standard risk-adjusted estimates of hospital performance compare favorably to credibly causal estimates." Again, in our setting, Doyle et al. (2019) conclude there to be little evidence of patient selection (see also Doyle et al., 2017, 2015).

does not predict the association between for-profits and chain hospitals. Finally, it seems reasonable to assume that entry decisions would not be driven primarily by the HRRP quality measures we focus on.

In sum, our findings show that in the hospital market, for-profit status and being part of a large chain are associated with lower quality and that there are important interactions with local market concentration. For hospitals that are part of a large chain, the negative for-profit quality gap is eliminated at low levels of market concentration. While we cannot distinguish between static and dynamic effects, our results indicate that anti-trust decision-makers may need to pay attention to both ownership and market structure when considering hospital merger and acquisition cases.

Our findings relate to a recent literature that has started to document large adverse effects of outsourcing to private firms on the provision of quality. In the healthcare context, Knutsson and Tyrefors (2022) show that quasi-random assignment to private ambulances is associated with lower quality. Gupta et al. (2021) show that the recent rise in private-equity owners of nursing homes results in lower quality. Our finding of market structure heterogeneity in quality also echos the analysis of price by Cooper et al. (2019) who find that market concentration is strongly associated with price in hospital markets, with similar findings in Craig et al. (2021).[8] Our results also complement a recent study of quality provision in nursing homes (Hackmann, 2019), which finds that financial reimbursements are more important than competition in improving patient outcomes. Finally, we find the heterogeneity by chain status and competition related to a small but growing literature on chain status and private equity in the healthcare market. For the dialysis market, Eliason et al. (2019) found that chain status lowers quality, but quality does not react to local competition. However, that paper does not distinguish between the for-profit and non-profit status of the firms. For the nursing home market, Gandhi et al. (2020) find that private equity-owned nursing homes' quality provision can be more sensitive to competition, and in the retail pharmacy market, Janssen and Zhang (2020) find some evidence that quality improves with chain affiliation. Finally, the issue of market power has recently attracted renewed attention in many contexts (Benkard et al., 2021).

2 Estimation and data

2.1 Measures of quality

We begin by constructing measures of quality. As noted above, our preferred measures are derived from the HRRP. This has several benefits. First, 'excess readmissions' are the measure the government uses

[8] Note that in our setting Medicare sets the prices (Cubanski et al., 2015).

to incentivize higher quality. Second, the penalties for having excess readmissions can be considerable: the Centers for Medicaid and Medicare Services [CMS] expected to recover $526 million in 2017, or 0.3% of overall Medicare payments going to hospitals. The penalty per excess readmission is about five to six times the Medicare base payment for a hospital stay for the particular condition (MedPAC-Report, 2018), and therefore much greater than the revenue to be gained from the readmission (Gupta, 2021). Thus, these quality measures are likely a choice variable for hospitals (Garthwaite et al., 2020). Third, the usefulness of these performance metrics is further underscored by them being predictive of market-shares (Chandra et al., 2016) and, at the regional level, are correlated with deaths from COVID-19 (Kunz and Propper, 2022).

The HRRP (detailed in Appendix A) takes the raw admission rates for three incentivized emergency conditions (AMI, HF, PN) and adjusts these for potential differences in patient populations to derive the excess readmission ratio [ERR]. This is the probability of readmissions relative to the average hospital with a similar case-mix (known as *peer-bench-marking*, Zhang et al., 2016). The penalty is applied if the ERR exceeds 1 in any of the three conditions. This cut-off is the policy-relevant discontinuity introduced into the hospital's cost function (Gupta, 2021).

These measures have been criticized because that risk-adjustment is only performed on age, sex, and other health conditions (co-morbidities) of the patients and does not take into account the differences in socioeconomic characteristics of the patients these hospitals serve (e.g., Aswani et al., 2018; Gu et al., 2014). Further, these measures are likely affected by regression to the mean (Joshi et al., 2019) and are over-reliant on small disturbances (Chandra et al., 2016). To address all these points, instead of using the raw measures directly, we perform an additional risk adjustment by extracting latent quality as measured by fixed effect for each hospital, and each condition (such approaches are often called *profiling* Normand et al., 2016). This approach adjusts for additional covariates and deals econometrically with the issue of small numbers of observations for each hospital (five years in our setting, more details on the method can be found in Appendix B). We use OLS for the readmission rate and the ERR and probit for the penalty status of the form:

$$E(y_{it}^c|\alpha_i^c, x_{it}) = \alpha_i^c + x_{it}'\beta \quad (1)$$

$$E(Penalty_{it}^c = 1|\alpha_i^c, x_{it}) = \Phi(\alpha_i^c + x_{it}'\beta), \quad (2)$$

where y_{it}^c denotes the different outcomes we consider: the patient survey, raw 30-day mortality, raw 30-day readmission rate, or the adjusted ERR. $Penalty_{it}^c$ is an indicator of hospital i being penalized in year t for exceeding readmissions in emergency condition c, i.e. $ERR_{it}^c > 1$, and $\Phi(\cdot)$ denotes the standard normal CDF. The vector x_{it} contains time-varying covariates at the hospital, Hospital Referral Region [HRR], and county level (detailed in section 2.3 below), as well as year indicators to control for common time shocks. We cluster standard errors at the hospital level.

Our interest lies in the fixed effects, α_i^c. We estimate several variants of the models: first, we pool the fixed effects across conditions resulting in one fixed effect per hospital, $\alpha_i^c = \alpha_i$ for all i (3,917 fixed effects); second, interacted condition×hospital fixed effects, α_i^c (8,713 fixed effects); third, estimating the regressions separately by condition.[9]

To deal with large and imprecise outliers, researchers often shrink the estimated fixed effects towards their mean using Empirical Bayes.[10] We follow this literature for the OLS fixed effects estimates in equation (1).[11] For the penalty status measure—equation (2)—we use a bias-reduced fixed effect probit approach (Kunz et al., 2021). This method's bias-reduction shrinks the α_i^c during estimation. Intuitively, it achieves this by adding a particular penalty term to the probit score function, which results in the fixed effects being shrunk towards zero. While serving the same purpose as Empirical Bayes for the linear model, this method also removes incidental parameter bias. It avoids the perfect prediction problem that affects the estimation of these types of nonlinear models. In our context, the perfect prediction problem means that any information in the covariates would be discarded when hospitals are always or never fined, as the corresponding predicted fixed effect (α_i^c) would be infinite. For such cases, shrinkage after estimation, such as with Empirical Bayes, would obviously not be possible as no estimates are obtained. In our application, the problem of hospitals being either always or never fined is particularly relevant since roughly 50 percent of hospitals did not change their penalty status during our five-year sample period.[12] In addition, Empirical Bayes often involves assumptions such as homogeneity, which cannot be satisfied with binary outcomes (Frederiksen et al., 2020). In contrast, the bias-reduction approach we pursue avoids any additional assumptions used for Empirical Bayes: the method automatically shrinks the fixed effects towards the conservative benchmark of the marginal hospital with neither positive nor negative relative quality.

The predicted fixed effects from Eq. (2), $\hat{\alpha}_i^c$, are used to calculate the marginal penalty propensities, $MPP_i^c \equiv \Phi(\hat{\alpha}_i^c)$ (and $MPP_i \equiv \Phi(\hat{\alpha}_i^c)$ respectively), which can be thought of as the propensity to

[9]Each approach leads to qualitatively similar conclusions.

[10]This approach has become ubiquitous in the literature on hospital quality (Chandra et al., 2016; Propper et al., 2008) and in other pay-for-performance metrics (Bonhomme and Weidner, 2019; Chetty et al., 2014). Although the HRRP aims to account for small counts by requiring at least 25 discharges in the respective emergency condition and corrects the excess readmission ratio using Empirical Bayes when estimating hospital fixed effects over five years and using them in secondary analysis, the inherent estimating error needs to be accounted for.

[11]Other approaches include random effects, using the residual of the quality equation, or fitting a normal distribution around the fixed effects estimates (see MacKenzie et al., 2015, for an overview of shrinkage approaches). Appendix Table C4, which contains results for a linear probability model with *ex-post* Empirical Bayes, shows that our results for (2) are robust to the employed methodology.

[12]One natural reason for the high level of persistence in the dependent variable over time is that a three-year reference window is used to determine the penalty status. We further account for this overlap when computing standard errors by clustering at the hospital level.

be fined if the hospital would otherwise be marginal ($x_{it}'\hat{\beta} = 0$). These estimated MPP_i^c are model-consistent as they respect the [0;1] bounds without any *ad hoc* adjustments. More details of the estimator are in Appendix B.

2.2 Assessing quality differences

Our analysis assesses whether there are differences in quality between for-profit and other hospitals and how this interacts with (local) market structure and chain status.[13] Using our preferred measure of hospital quality, we estimate OLS regressions of the form:

$$MPP_i^c = \gamma_1 \text{For-profit}_i + \gamma_2 \text{For-profit}_i \times \text{HHI}_{i_{hrr}} + z_i'\beta + \delta^c + \delta_{i_{hrr}} + \varepsilon_i, \tag{3}$$

where z_i contains (almost) constant measures of hospitals' teaching status, as well as urban, size, and county-level covariates. δ^c denote emergency condition fixed effects and $\delta_{i_{hrr}}$, hospital referral region fixed effects. To test whether observed differences by for-profit status (For-profit$_i$) change with market conditions, we interact the for-profit indicator with a measure of market concentration the Herfindahl-Hirschmann index, HHI. We repeat this analysis replacing for-profit and non-profit in (3) with chain/non-chain. Finally, to test the impact of being part of a hospital chain, we stratify our analysis by whether the hospital is part of a (large) chain (being in the top tercile).

2.3 Data

Our data are from the administrative Hospital Compare dataset provided by the Centers for Medicare and Medicaid Services [CMS], which gives information on penalties for 2012–2016. Reporting is delayed by one year, so the data relating to the three-year aggregates of readmissions during 2011–2015. For each of the 3,197 included hospitals, each of the three emergency conditions, and every year, we know whether or not a penalty was issued (there are 8,713 hospital×condition and 41,095 hospital×condition×year observations). As noted above, using profiling (1) and (2), we extract from the outcomes the latent hospital quality that is not driven by aggregate changes over time and local area characteristics beyond the hospital's control. We follow the extant literature (Chandra et al., 2016; Gu et al., 2014) to select relevant covariates. In equations (1) and (2), we use the following HRR- and county-level controls. Measures defined at the HRR-level are taken from the Dartmouth Atlas of Health Care. We use the number of ambulatory-care-sensitive conditions (ACSC), measuring the accessibility of local primary health care (Gu et al., 2014), and changes in the number of hospitals in the

[13] We keep government hospitals in our analysis but present estimates excluding them in Appendix Table C5.

region (Chandra et al., 2016), via changes in the hospital market (i.e., opening and closing of hospitals in the HRR).[14] At the county level, we use the Federal Information Processing Standard to add community characteristics, such as the poverty rate (which is correlated with potential changes in the Medicaid population) and the median household income, which have been discussed as determinants of readmission rates outside the control of the hospital (Herrin et al., 2015). Other controls are the number of discharges in the emergency conditions, adjusting for the different sizes of the hospitals and total population in the county served, to prevent over-weighting large hospitals or those in large catchment areas.[15]

Most hospital-level variables display little or no variation over time, and while it is essential to control for these (Normand et al., 2016), they can only be included in the second step estimation. We, therefore, add the following additional hospital-level controls not included in equations (1) and (2) to equation (3): an urban/rural indicator variable, teaching status of the hospital, size (number of beds), for-profit, and chain status from the corresponding final rule impact files. The covariates further account for racial and ethnic, educational, and age composition in the hospital environment. The latter captures potential differences in the Medicare population. To measure local market (HRR-level) concentration, we construct a standard measure of competition, the HHI, based on the number of beds. We define variants of this measure to examine sensitivity: (1) HHI based on discharges in the respective emergency conditions; (2) treating hospitals that belong to the same chains as one hospital (Hausman and Lavetti, 2021); (3) calculating the HHI from discharges before the HRRP (in 2008) to address concerns over endogeneity of location (our preferred specification); and finally, for hospitals organized in chains we calculate the nationally chain-wide HHI via (4) the simple average and (5) the discharges-weighted average of all local HHIs faced by its hospitals.[16]

In addition to our preferred quality measure, we also examine four other quality measures. These are a measure of patient satisfaction (the share that would recommend the hospital), the 30-day mortality rate; the 30-day raw admission rate; and the risk-adjusted excess readmission rate (i.e., the basis

[14]Here, as is common in the literature (see Chandra et al., 2016), we define market entry and exit as changes in the hospital market, which need not be literal entries or exits.

[15]In a robustness test below, we also assess whether additionally weighting by size.

[16]It is important to note that the coefficient on HHI in a regression of HHI on quality should not be interpreted causally as it is an equilibrium outcome (Miller et al., 2022). Here, we use the HHI measure merely to assess the heterogeneity of the quality associations. As noted above, to ensure the measure is constant with respect to equilibrium adjustments, we account for market changes in our risk adjustment, use pre-period HHI, and only use hospitals that do not change status to assess this relationship.

for our penalty propensity measure).[17] The variables and data sources are in Appendix D, and the descriptive statistics are in Table C1.

2.4 Association of quality measures and spatial distribution

We begin by graphically showing our extracted quality fixed effects (the risk-adjustment regressions are presented in the Appendix C3). Figure 1 shows the association of our quality measure with the other less incentivized hospital-wide quality metrics (for this, we pool across the three condition-specific hospital estimates to obtain a single hospital-level quality measure). The three graphs show that the rank ($Rank(\alpha_i)$) in the hospital quality measure (grouped into 100 bins) correlates, respectively, with the overall readmission rate,[18] the average hospital mortality (in the same condition) and patient satisfaction. While these three measures are not risk-adjusted, all three correlate positively with our quality measure; the readmission rate and patient satisfaction, very strongly so. The strong correspondence between our objective and subjective measures (the patient survey responses) is notable as *ex-ante* it is not obvious that patients would value the same quality dimensions as those related to readmissions or even mortality. These results suggest that our penalty propensity measures a common hospital quality. In the Appendix, we show that our penalty status measure is strongly correlated across the three different penalized conditions, further suggesting that the measure picks up overall hospital quality rather than department-specific quality (cf. Appendix Figure C2).

Next we turn to MPP_i, our penalty propensity measure of hospital quality. Figure 2, Panel A, shows its substantial spatial heterogeneity across the 307 HRRs (see also Finkelstein et al., 2016). The estimated propensity of being penalized varies from 15% to 95%, with an interquartile penalty-propensity difference of nearly 15 percentage points (p.p.). Panel B shows the distribution of for-profit hospital shares. Panel C shows that of chain shares.[19] Panel D shows the distribution of market concentration. Overall, the maps show considerable dispersion in all three measures and no obvious mapping in geographical space between quality, for-profit status, and market concentration. Further, a comparison of Panels A and B shows that for-profits are not necessarily more likely to be located in

[17] These three dimensions make up 66% of the publicized Hospital Compare metric, which is a five-star rating based on latent factor analysis. The metric has more than 20% missing observations in most years; it is also extracted from different mixes of indicators (prone to strategic reporting) based on availability for other hospitals and years. These issues compromise its comparability, and therefore, we do not use them here as a measure of quality in our analysis.

[18] This partially addresses concerns whether hospitals used observational stays to reduce readmissions in targeted conditions. For heart failure, this hypothesis has also been dispelled more directly in Albritton et al. (2018).

[19] Roughly 50% of the HRRs (164) contain both (at least one) for-profit and non-profit hospitals. For chains and lone-standing hospitals, 70% (226) share an HRR. These numbers are lower estimates of the mixing within HRR as they exclude hospitals that changed their status in the observation period.

higher-penalty areas.[20]

Figure 3 examines whether for-profit or chain status is associated with local market (HRR) concentration. The left panel shows the association between the share of for-profits in a local market and bins of the local market HHI based on 2008 discharges (pre-policy and our measurement period). The right-hand panel shows the same association for the share of chains. There is no association between the share of for-profits in a local market and market concentration. The association between the share of chains and market concentration is positive (chains locate more in more monopolistic markets) but is far from being economically or statistically significant at conventional levels. These provide evidence that (lagged) market concentration is not correlated with either share of for-profit or chain status.

3 Results

3.1 Readmission performance by hospital ownership

Table 1 presents the quality differences between for-profits and other hospitals for the five quality measures. Columns (1)-(3) show the unconditional means and raw differences pooled across years and conditions. Quality is unconditionally lower in for-profits on all measures. These differences are statistically significant for all measures except the 30-day mortality rate.[21] So, for example, the third row of the table shows for-profits have a statistically significant 0.3 p.p. higher rate of readmissions, which corresponds to around 10% of the total reduction in readmissions over the sample period, which fell from 21.5% to 17.8% (and was considered a success, e.g. Zuckerman et al., 2016). The fourth row presents the raw CMS measure, which adjusts for gender, age, and co-morbidities, and compares the adjusted rate to expected readmissions using comparable hospitals. Column (3) shows a 1.3 p.p. for-profit difference, indicating that the quality gap is not driven by differential coding of co-morbidities (one strategy by which hospitals reduce their penalty risk, Ody et al., 2019).[22] The final row shows the quality difference measured by our preferred penalty status measure. Column (3) shows that for-

[20] Table C2 shows descriptive bivariate correlations of for-profit ownership and regional characteristics. For example, for-profits are less likely to locate in rural areas or where the share of unemployed people is high, reversely they are more likely to locate where there is a larger share of older people (65 and older) and where there are more discharges for ambulatory care sensitive conditions (ASCS), a common measure of low quality of local health care provision outside the hospital.

[21] This is the noisiest of our measures as it is a relatively low probability event compared to the other measures we use here.

[22] Note that in our setting and time period, Buxbaum et al. (2019) find “changes in the coding of inpatient pneumonia admissions do not explain readmission reduction following the HRRP”; they also compare explicitly for-profit and other hospitals and find no differential coding behavior between the two types of hospitals.

profits have a 7.1 p.p. higher probability of being penalized.[23] Using these measures for profit-hospitals appear to provide lower quality services.

However, these measures have been criticized for insufficiently accounting for risk and market-level factors. Columns (4) and (5), therefore, use fixed effects estimated from equation (1) and MPP_i and MPP_i^c from equation (2), as the outcome measures, regressed upon the for-profit indicator and additional time-constant covariates. As discussed above, these measures are risk-adjusted for a large number of time-varying covariates and include shrinkage. Column (4) pools the three target conditions, MPP_i. The for-profit gap in all measures (except, again the 30-day mortality) remains large and statistically significant, although smaller than without the additional adjustments. The results imply a 0.1 p.p. higher readmissions rate, a 0.7 p.p. higher ERR, and a 2.3 p.p. higher marginal penalty propensity. Using the expected \$526 million penalties as a benchmark (MedPAC-Report, 2018), this amounts to 12 million per year expected to be paid in excess by for-profits. Column (5) shows the results estimating one fixed effect for each hospital×condition, MPP_i^c, which allows for inherent differences across the three conditions. The for-profit gap is slightly larger than in Column (4). For example, the difference is 3 p.p. in the marginal penalty propensity, corresponding to a 6.3% increase from the mean of 47.4%.[24] That all these measures, adjusted or not, point in the same direction makes it unlikely that this conclusion is driven by hospital coding choices.

3.2 Quality, ownership, and concentration

Since Table 1 shows qualitatively similar results across the different quality measures, our subsequent analyses use only our penalty propensity measure as it contains the most correction for patient selection and measurement error. Table 2 examines the interaction between for-profit status, chain status, and market concentration. It shows the differences in quality associated with competition by profit (Panel A) and chain status (Panel B).

Column (1) uses MPP_i^c, and all time-constant covariates, the coefficient in Panel A is, therefore, the same as that in the last row of Table 1-Column (5), as a benchmark. The marginal probability of incurring a penalty (i.e., having lower quality) is 3 p.p. higher for for-profit hospitals. As noted above, a small number of hospitals changed their for-profit status during the observation period. Column (2)

[23] Consequently, the significant gap in readmissions translates into penalties, which suggests that there is little bunching at the penalty cut-off. This is perhaps a result of the peer bench-marking embedded in the ERR and penalty status, which is unobserved by the individual hospital (Gupta, 2017; Zhang et al., 2016). Another way how hospitals could prevent penalties is by bunching below the admissions cut-off. CMS requires a minimum of 25-discharges (over three years) to be eligible for penalties. We show that being above or below the 25-discharge cut-off does not covary with for-profit status, market concentration, or chain status in Appendix Table C5.

[24] Appendix Table C3 presents the covariates in the regressions.

omits these hospitals from the estimation sample.[25] The for-profit gap remains the same (3.1 p.p.). Column (3) additionally controls for HRR fixed effects. The for-profit gap increases to 5 p.p. Thus, the within-HRR gap is larger than the between-HRR gap.

In Columns (4)-(7), we test the hypothesis that competition may drive out the for-profit differential by interacting the for-profit indicator with the market-level HHI for the respective penalized condition while continuing to condition on HRR fixed effects (which absorb the baseline level of competition). Column (4), Panel A, shows that the for-profit gap drops back to 3.3 p.p. for hospitals in markets with zero concentration. A one standard deviation change in HHI based on bed numbers ($SD(HHI) = 0.149$) increases the for-profit differential in the penalty propensity to 5.1 p.p. (shown in the 'For-profit + interaction' row) with a corresponding p-value of 0.026. These results are robust to other measures of HHI, based on discharges in respective emergency conditions only (Column 5), treating chains as one competitor in the local market (Column 6), and using market concentration before the period our data covers (Column 7). Using this last (and our preferred) specification to benchmark, these differences would imply that 47% of the mean raw quality difference can be attributed to the lack of competition in the average market.[26]

Panel B repeats this analysis focusing on the difference between hospitals that are part of a chain and those that are not. It shows that, in contrast to for-profit status, chain status *per se* is not associated with a significant difference in quality, nor is there an interaction with a market concentration in any of the specifications.

Table 3 presents several robustness tests. These all use market concentration measured in 2008 to reduce dynamic equilibrium-adjustment concerns. Panel A examines the effect of for-profit status. Column (1) repeats Column (7) of Table 2, the baseline specification. Column (2) examines potential non-linearities in local market concentration, replacing the continuous measure with a binary indicator variable for being above the median. This shows that being in an above-median HHI region essentially doubles the for-profit marginal penalty propensity relative to other hospitals. Next, we assess whether monopoly markets are driving the results by dropping fully-monopolistic HHR markets (Column 3) and dropping hospitals operating in markets in the highest 10th percentile of the HHI distribution (Column 4). Excluding these highly concentrated markets yields similar—even slightly larger—results.

Our period coincided with the insurance expansions mandated by the Affordable Care Act. This differentially impacted markets, and hospitals within markets, by improving hospital finances and

[25]We additionally verified that there are no significant pre-acquisition differences in the risk-adjusted quality among hospitals that changed ownership; results are available upon request.

[26]That is, the quality difference at the average market concentration (HHI=0.233) versus a fully competitive market (HHI=0) relative to the raw mean quality difference (7 p.p.): $0.47 = (0.235 \times 0.14)/0.07$.

changing their mix of patients and services. To address whether this policy change is driving our results, we include as an additional covariate the hospitals' county-level share of uninsured in 2011 (see Davis et al., 2020).[27] The results are unaffected by this additional control variable (Column 5). Our main analyses are at the condition-specific level to allow hospitals to respond to market power in that condition. However, as noted above, our quality measures are correlated and so pick up overall hospital quality, and decision-making about quality will be at the hospital level as well as at the condition level (as suggested by Figure C2). We therefore repeat our analysis using the average quality at the hospital level in Column (6) of Table 3. This shows little difference from the baseline analysis that is at the condition level. Finally, although we believe the HRR is the appropriate market level, hospitals may respond to (even) more local competition. To examine this, we re-estimate within the smaller health service area (HSA) markets (by replacing the HRR fixed effects with HSA fixed effects) (Column 7). The results are very similar to those of Table 2. For-profit hospitals are of lower quality and this quality is lower in more concentrated markets.

Panel B shows that, across all these robustness checks, chain hospitals, on average, do not differ from non-chain hospitals in terms of quality, nor do they differ in response to local market concentration.

3.3 Heterogeneity by chain status and size

We define a hospital to be part of a chain if it has any formal affiliation with another hospital. This classification, therefore, includes both large national as well as small local chains, and hospitals with a loose affiliation as well as those with much tighter chain-level managerial control. But it is possible that the quality of for-profit hospitals and the interaction with local market concentration varies within chain status, particularly with the size of the chain.

To examine this, Table 4 presents the estimated for-profit quality difference for various sub-samples defined by chain status. As before, all analyses use market concentration measured in 2008. Column (1) shows again our previous baseline estimates for all hospitals. Column (2) examines only non-chain hospitals. It shows that among non-chain hospitals, there is no significant for-profit difference in quality, nor does it significantly vary with market concentration. Column (3) shows that if we focus only on-chain hospitals, there is a sizable and significant for-profit gap and an interaction with the market concentration. The two terms are also jointly significant (p=0.006).

Columns (4) and (5) split chain hospitals into those belonging to nationally small chains and large chains, where small and large are defined by not belonging or belonging to the highest tercile of the

[27] We do not observe the share of Medicare or Medicaid patients in the local market in our data; but, to address this, our specifications include the share of county population living in poverty and the share of the population older than 65 in our standard set of controls as noted above.

chain size distribution (which equates to more than 20 hospitals) respectively. Column (4) shows that for-profits that are part of small chains have lower quality, but there is no interaction with market concentration. In contrast, Column (5) shows that for hospitals that are part of large chains, the for-profit gap is completely crowded out in fully competitive markets but increases strongly the more concentrated the local market is. Column (6) explores whether national market power matters, replacing the local market measure with the average HHI across all markets in the chain operates. Column (7) repeats this using discharges-weighted average HHI. The results are qualitatively the same as those using the local concentration measure: the for-profit gap is significantly larger in markets with greater market power.

3.4 Further robustness checks

Appendix Table C4 presents a series of checks on the precise definition of the quality measure. Columns (2)-(4) show that our results hold for the percentile ranking in our quality measure (which does not depend on the magnitude of the penalty propensity) and for a definition of being in the high-quality tail (defined as an indicator for being in either the top 10% or 25% hospitals).[28] Column (5) uses the raw penalty (i.e., without our adjustments for the socioeconomic environment and small samples). The results are similar to our preferred specification but are less precise, showing that the various fixed-effect profiling approaches are beneficial. The final three columns examine the Empirical Bayes OLS penalty measure, raw readmissions, and the published ERR measure. The use of these different measures also does not affect our qualitative conclusions regarding the for-profit-concentration gap among all hospitals and large chain hospitals.

Appendix Table C5 shows various further tests of robustness. Panel A is for all hospitals and Panel B for hospitals which are part of a large chain. In Column (2) we weight hospitals by their size; in Column (3) we drop government hospitals from the analysis. The results are robust to these changes. To further assuage endogeneity concerns, we restrict the sample to have constant ownership for the ten years before the end of our sample period (Column 4) and use only data for the first two years after HRRP enactment (Column 5). Our qualitative conclusions are unaffected, although some precision is lost due to the reduced sample sizes. In Column (6), we use a fractional response model (Papke and Wooldridge, 2008) for the second stage of our preferred estimator, as the MPP is bounded between 0-1. The results are robust to this change. Finally, in Column (7), we examine whether there is strategic behavior by for-profit hospitals or chain hospitals by examining whether the hospital is below the 25 discharges that determine whether the HRRP penalties can be applied (adjusted for our standard set of potential confounders). The results show no indication of strategic behavior on this measure. The

[28]See also the densities of the quality measure for both BRGLM-penalty and Ebayes-OLS-ERR approaches shown in Figure C1).

full results for all chain and non-chain sub-samples (analogous to Table 2) are presented in Table C6. This shows robustness across all the specifications.

4 Discussion and conclusion

The differences in the behavior of for-profits and non-profits in public service provision are well documented. Examples include technology adoption (Horwitz et al., 2018), the provision of free care (Garthwaite et al., 2018), and the need for financial incentives to motivate workers (Besley and Ghatak, 2005). Non-profits have historically played a large role in the US hospital industry, but the role of for-profits, and hospital chains, has been steadily increasing. As the ownership and market structures of the hospital market have changed, it is essential to understand how these changed conditions are associated with differences in hospital quality.

To address this, we examine differences in quality between for-profit and non-profit hospitals using the distribution of latent heterogeneity in quality derived from the underlying penalty risk of being fined under a flagship program designed to incentivize hospital quality.[29] This quality measure is strongly correlated across diagnostic conditions within a hospital, suggesting a common hospital-wide quality, which is further supported by the strong association of our quality measure with a hospital's overall readmission rates in non-incentivized and non-emergency conditions, mortality rates in incentivized conditions, and patient satisfaction.

We use our measure to examine quality differences by for-profit status, chain status, market structure, and the interaction between these. Our findings show for-profit hospitals are around 5% more likely to be penalized than other hospitals. There is no such quality gap by chain status. We find a substantial gradient in the negative for-profit gap in quality by market structure, but the chain-non-chain quality gap does not differ by market concentration. However, there is an interaction between for-profit status, chain status, and market structure. We find that the for-profit gap is largely driven by chain hospitals. For small chains, the quality gap is large (0.049) but not statistically significantly changing in HHI, while for large chains, the quality gap is strongly increasing in HHI.

Our finding that competition in local markets is associated with smaller for-profit gap chimes with earlier findings in the literature. Our finding that (large) chain hospitals are responsive to competition

[29] Many attempts to incentivize the provision of public services share with the HRRP a similar structure of being based on annually published metrics and using nonlinear policy cut-offs (e.g., penalty/no-penalty), which may frequently change as the basis for reward or penalization of providers. While our findings are specific to the HRRP, our approach to the extraction of a medium-run measure of quality from an observed binary signal of performance could be applied in similar settings to understand the factors associated with quality (Mehta, 2019).

contrasts with Eliason et al. (2019), who find that competition does not affect chain hospital quality in the dialysis market. Instead, it is chain status per se that drives lower quality. There are (at least) two possible reasons for this difference. First, dialysis patients have little choice over where they are treated, meaning competition may play little role. Second, performance for dialysis outcomes is not part of the measures published in patient-oriented quality metrics such as Hospital Compare. In contrast, the penalty rates we examine were heavily publicized. Thus it seems probable that hospitals are more likely to compete for patients on HRRP signals than on dialysis outcomes. Poor performance on these metrics might be expected to be related to loss of business (Chandra et al., 2016). An increase in competition will exacerbate this, which may be particularly the case for those hospitals at the lower end of the quality distribution (Jones et al., 2017). We show that hospitals in the large chain have lower quality if a lack of competition in their local market allows for flexibility in quality. But if they operate in highly competitive markets, the gap between for-profit and non-profit chain hospitals disappears.

An alternative explanation for our findings is that hospitals respond to the publication of performance metrics by selecting healthy patients. They do this more in competitive markets where they are more at risk of losing business. Chain hospitals may implement chain-wide methods to do this and be more efficient than hospitals not part of a chain. We have addressed endogenous sorting of patients by using readmission rates published by CMS for emergency admission, which prior research has shown are not driven by patient selection (Doyle et al., 2019) and by further risk-adjusting them. Therefore we think it is unlikely that our results are only driven by patient selection. And if for-profits were more likely to select healthier patients, a tendency which might plausibly be hypothesized to be even stronger among those in large chains, our estimated for-profit-chain gap would reflect a lower bound on the true gap. We have also addressed concerns about the differential location by hospital type and our results are robust to this.

Finally, while our results are not causal, our findings document that when located in markets in which they can exploit their market power, for-profits, particularly those that are parts of a large chain, have lower quality. This finding holds across different outcomes that are more or less risk-adjusted or incentivised and a very large number of checks and sensitivity analyses. Thus the correlation between lower quality and lower competition should form part of any regulatory assessment of proposed mergers and acquisition behavior in the hospital industry, especially in markets with high market concentration.

References

Abrams, D. S., M. Bertrand, and S. Mullainathan (2012). Do judges vary in their treatment of race? *Journal of Legal Studies 41*(2), 347–383.

AHA (2019). Fast facts on u.s. hospitals, 2019.

Albritton, J., T. W. Belnap, and L. A. Savitz (2018). The effect of the hospital readmissions reduction program on readmission and observation stay rates for heart failure. *Health Affairs 37*(10), 1632–1939.

Aswani, M. S., M. L. Kilgore, D. J. Becker, D. T. Redden, B. Sen, and J. Blackburn (2018). Differential impact of hospital and community factors on medicare readmission penalties. *Health Services Research 53*(6), 4416–4436.

Baltagi, B. H. and Y.-F. Yen (2014). Hospital treatment rates and spillover effects: Does ownership matter? *Regional Science and Urban Economics 49*, 193–202.

Beaulieu, N. D., L. S. Dafny, B. E. Landon, J. B. Dalton, I. Kuye, and J. M. McWilliams (2020). Changes in quality of care after hospital mergers and acquisitions. *New England Journal of Medicine 382*, 51–59.

Benkard, C. L., A. Yurukoglu, and A. L. Zhang (2021). Concentration in product markets. Technical report, National Bureau of Economic Research.

Bernheim, S. M., C. S. Parzynski, L. Horwitz, Z. Lin, M. J. Araas, J. S. Ross, E. E. Drye, L. G. Suter, S.-L. T. Normand, and H. M. Krumholz (2016). Accounting for patients' socioeconomic status does not change hospital readmission rates. *Health Affairs 35*(8), 1461–1470.

Besley, T. and M. Ghatak (2005). Competition and incentives with motivated agents. *American Economic Review 95*(3), 616–636.

Besley, T. and J. M. Malcomson (2018). Competition in public service provision: The role of not-for-profit providers. *Journal of Public Economics 162*, 158–172.

Bonhomme, S. and M. Weidner (2019). Posterior average effects. *arXiv preprint arXiv:1906.06360*.

Buxbaum, J. D., P. K. Lindenauer, C. R. Cooke, U. Nuliyalu, and A. M. Ryan (2019). Changes in coding of pneumonia and impact on the hospital readmission reduction program. *Health services research 54*(6), 1326–1334.

Capps, C., D. W. Carlton, and G. David (2020). Antitrust treatment of nonprofits: Should hospitals receive special care? *Economic Inquiery 58*(3), 1183–1199.

Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson (2016). Health care exceptionalism? performance and allocation in the us health care sector. *American Economic Review 106*(8), 2110–2144.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review 104*(9), 2633–2679.

Colla, C., J. Bynum, A. Austin, and J. Skinner (2016). Hospital competition, quality, and expenditures in the us medicare population. *NBER Working Paper No. 22826*.

Cooper, Z., S. V. Craig, M. Gaynor, and J. Van Reenen (2019). The price ain't right? hospital prices and health spending on the privately insured. *Quarterly Journal of Economics 134*(1), 51–107.

Cooper, Z., S. Gibbons, and M. Skellern (2018). Does competition from private surgical centres improve public hospitals' performance? evidence from the english national health service. *Journal of Public Economics 166*, 63–80.

Craig, S. V., M. Grennan, and A. Swanson (2021). Mergers and marginal costs: New evidence on hospital buyer power. *RAND Journal of Economics 52*(1), 151–178.

Cubanski, J., C. Swoope, C. Boccuti, G. Jacobson, G. Casillas, S. Griffin, and T. Neuman (2015). A primer on medicare: Key facts about the medicare program and the people it covers. Technical report, Kaiser Family Foundation's Program on Medicare Policy.

Davis, W., A. D. Gordan, and R. Tchernis (2020). Measuring the spatial distribution of health rankings in the united states. *National Bureau of Economic Research w27259, May*.

Desai, N. R., J. S. Ross, J. Y. Kwon, J. Herrin, K. Dharmarajan, S. M. Bernheim, H. M. Krumholz, and L. I. Horwitz (2016). Association between hospital penalty status under the hospital readmission reduction program and readmission rates for target and nontarget conditions. *Journal of the American Medical Association 316*(24), 2647–2656.

Doyle, J. J., J. Graves, and J. Gruber (2019). Evaluating measures of hospital quality: Evidence from ambulance referral patterns. *Review of Economics and Statistics 101*(5), 841–852.

Doyle, J. J., J. A. Graves, and J. Gruber (2017). Uncovering waste in us healthcare: Evidence from ambulance referral patterns. *Journal of Health Economics 54*, 25–39.

Doyle, J. J., J. A. Graves, J. Gruber, and S. A. Kleiner (2015). Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy 123*(1), 170–214.

Dranove, D. and M. A. Satterthwaite (2000). The industrial organization of health care markets. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, Volume 1, Part B, Chapter 20, pp. 1093–1139. Elsevier.

Duggan, M. (2002). Hospital market structure and the behavior of not-for-profit hospitals. *RAND Journal of Economics 33*(3), 433–446.

Duggan, M. G. (2000). Hospital ownership and public medical spending. *Quarterly Journal of Economics 115*(4), 1343–1373.

Eggleston, K., Y.-C. Shen, J. Lau, C. H. Schmid, and J. Chan (2008). Hospital ownership and quality of care: what explains the different results in the literature? *Health Economics 17*(12), 1345–1362.

Eliason, P. J., B. Heebsh, R. C. McDevitt, and J. W. Roberts (2019). How acquisitions affect firm behavior and performance: Evidence from the dialysis industry. *Quarterly Journal of Economics 135*(1), 221–267.

Finkelstein, A., M. Gentzkow, and H. Williams (2016). Sources of geographic variation in health care: Evidence from patient migration. *Quarterly Journal of Economics 131*(4), 1681–1726.

Frederiksen, A., L. B. Kahn, and F. Lange (2020). Supervisors and performance management systems. *Journal of Political Economy 128*(6), 2123–2187.

Friedrich, B. U. and M. B. Hackmann (2021). The returns to nursing: Evidence from a parental-leave program. *Review of Economic Studies 88*(5), 2308–2343.

Gandhi, A., Y. Song, and P. Upadrashta (2020). Private equity, consumers, and competition: Evidence from the nursing home industry. *Consumers, and Competition: Evidence from the Nursing Home Industry (June 12, 2020)*.

Garthwaite, C., T. Gross, and M. J. Notowidigdo (2018). Hospitals as insurers of last resort. *American Economic Journal: Applied Economics 10*(1), 1–39.

Garthwaite, C., C. Ody, and A. Starc (2020). Endogenous quality investments in the us hospital market. Technical report, National Bureau of Economic Research.

Gaynor, M., K. Ho, and R. J. Town (2015). The industrial organization of health-care markets. *Journal of Economic Literature 53*(2), 235–84.

Gaynor, M. and R. J. Town (2011). Competition in health care market. In M. V. Pauly, T. G. Mcguire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 2, Chapter Chapter 9, pp. 499–637. Elsevier B.V.

Gu, Q., L. Koenig, J. Faerberg, C. R. Steinberg, C. Vaz, and M. P. Wheatley (2014). The medicare hospital readmissions reduction program: potential unintended consequences for hospitals serving vulnerable populations. *Health Services Research 49*(3), 818–837.

Gupta, A. (2017, Nov). Impacts of performance pay for hospitals: The readmissions reduction program. *Mimeo*.

Gupta, A. (2021). Impacts of performance pay for hospitals: The readmissions reduction program. *American Economic Review 111*(4), 1241–83.

Gupta, A., S. T. Howell, C. Yannelis, and A. Gupta (2021). Does private equity investment in healthcare benefit patients? evidence from nursing homes. Technical report, National Bureau of Economic Research No 28474.

Hackmann, M. B. (2019). Incentivizing better quality of care: The role of medicaid and competition in the nursing home industry. *American Economic Review 109*(5), 1684–1716.

Hausman, N. and K. Lavetti (2021). Physician practice organization and negotiated prices: evidence from state law changes. *American Economic Journal: Applied Economics 13*(2), 258–96.

Herrin, J., J. St. Andre, K. Kenward, M. S. Joshi, A.-M. J. Audet, and S. C. Hines (2015). Community factors and hospital readmission rates. *Health Services Research 50*(1), 20–39.

Horwitz, J. R., C. Hsuan, and A. Nichols (2018). The role of hospital and market characteristics in invasive cardiac service diffusion. *Review of Industrial Organization 53*(1), 81–115.

Janssen, A. and X. Zhang (2020). Retail pharmacies and drug diversion during the opioid epidemic. *Available at SSRN 3745414*.

Jencks, S. F., M. V. Williams, and E. A. Coleman (2009). Rehospitalizations among patients in the medicare fee-for-service program. *New England Journal of Medicine 360*(14), 1418–1428.

Jindal, R. P., D. K. Gauri, G. Singh, and S. Nicholson (2018). Factors influencing hospital readmission penalties: Are they really under hospitals' control? *Decision Support Systems 110*, 58–70.

Jones, D. B., C. Propper, and S. Smith (2017). Wolves in sheep's clothing: Is non-profit status used to signal quality? *Journal of Health Economics 55*, 108–120.

Joshi, S., T. Nuckols, J. Escarce, P. Huckfeldt, I. Popescu, and N. Sood (2019). Regression to the mean in the medicare hospital readmissions reduction program. *JAMA Internal Medicine 179*(9), 1167–1173.

Joynt, K. E. and A. K. Jha (2013). Characteristics of hospitals receiving penalties under the hospital readmissions reduction program. *Journal of the American Medical Association: Research Letters 309*(4), 342–343.

Knutsson, D. and B. Tyrefors (2022). The quality and efficiency of public and private firms: Evidence from ambulance services. *The Quarterly Journal of Economics 137*(4), 2213–2262.

Kosmidis, I. and D. Firth (2009). Bias reduction in exponential family nonlinear models. *Biometrika 96*(4), 793–804.

Kunz, J. S. and C. Propper (2022). Jue insight: Is hospital quality predictive of pandemic deaths? evidence from us counties. *Journal of Urban Economics*, 103472.

Kunz, J. S., K. E. Staub, and R. Winkelmann (2021). Predicting individual effects in fixed effects panel probit models. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 184*(3), 1109–1145.

MacKenzie, T. A., G. L. Grunkemeier, G. K. Grunwald, A. J. O'Malley, C. Bohn, Y. Wu, and D. J. Malenka (2015). A primer on using shrinkage to compare in-hospital mortality between centers. *Annals of Thoracic Surgery 99*(3), 757–761.

Maddala, G. S. (1983). *Qualitative and limited dependent variable models in econometrics*. Cambridge: Cambridge University Press.

McClellan, M. B. and D. O. Staiger (2000). Comparing hospital quality at for-profit and not-for-profit hospitals. In *The changing hospital industry: comparing for-profit and Not-for-profit institutions*, pp. 93–112. University of Chicago Press.

McIlvennan, C. K., Z. J. Eapen, and L. A. Allen (2015). Hospital readmissions reduction program. *Circulation 131*(20), 1796–1803.

MedPAC-Report (2018). Mandated report: the effects of the hospital readmissions reduction program. In *Report to Congress: Medicare and the Health Care Delivery System*, pp. 14–15.

Mehta, N. (2019). Measuring quality for use in incentive schemes: The case of "shrinkage" estimators. *Quantitative Economics 10*, 1537–1577.

Mellor, J., M. Daly, and M. Smith (2017). Does it pay to penalize hospitals for excess readmissions? intended and unintended consequences of medicare's hospital readmissions reductions program. *Health Economics 26*, 1037–1051.

Miller, N., S. Berry, F. Scott Morton, J. Baker, T. Bresnahan, M. Gaynor, R. Gilbert, G. Hay, G. Jin, B. Kobayashi, et al. (2022). On the misuse of regressions of price on the hhi in merger review. *Journal of Antitrust Enforcement 10*(2), 248–259.

Moscelli, G., H. Gravelle, and L. Siciliani (2021). Hospital competition and quality for non-emergency patients in the english nhs. *RAND Journal of Economics 52*(2), 382–414.

Normand, S.-L. T., A. S. Ash, S. E. Fienberg, T. A. Stukel, J. Utts, and T. A. Louis (2016). League tables for hospital comparisons. *Annual Review of Statistics and Its Application 3*, 21–50.

Ody, C., L. Msall, L. S. Dafny, D. C. Grabowski, and D. M. Cutler (2019). Decreases in readmissions credited to medicare's program to reduce hospital readmissions have been overstated. *Health Affairs 38*(1), 36–43.

Papke, L. E. and J. M. Wooldridge (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics 145*(1), 121–133.

Paul, J. A., B. Quosigk, and L. MacDonald (2020). Does hospital status affect performance? *Nonprofit and Voluntary Sector Quarterly 49*(2), 229–251.

Picone, G., S.-Y. Chou, and F. Sloan (2002). Are for-profit hospital conversions harmful to patients and to medicare? *RAND Journal of Economics 33*(3), 507–523.

Propper, C., S. Burgess, and D. Gossage (2008). Competition and quality: Evidence from the nhs internal market 1991-9. *Economic Journal 118*(525), 138–170.

Sloan, F. A., G. A. Picone, D. H. Taylor, and S.-Y. Chou (2001). Hospital ownership and cost and quality of care: is there a dime's worth of difference? *Journal of Health Economics 20*(1), 1–21.

Song, H., E. Andreyeva, and G. David (2021). Time is the wisest counselor of all: The value of provider–patient engagement length in home healthcare. *Management Science*.

Staiger, D. (2018). What health care teaches us about measuring productivity in higher education. In *Productivity in Higher Education*, pp. 17–29. University of Chicago Press.

Times, T. N. Y. (2021, 6). Buoyed by federal covid aid, big hospital chains buy up competitors.

Wasfy, J. H., C. M. Zigler, C. Choirat, Y. Wang, F. Dominici, and R. W. Yeh (2017). Readmission rates after passage of the hospital readmissions reduction programa pre–post analysisreadmission rates after passage of hospital readmissions reduction program. *Annals of Internal Medicine 166*(5), 324–331.

Zhang, D. J., I. Gurvich, J. A. V. Mieghem, E. Park, R. S. Young, and M. V. Williams (2016). Hospital readmissions reduction program: An economic and operational analysis. *Management Science 62*(11), 3351–3371.

Zuckerman, R. B., S. H. Sheingold, E. J. Orav, J. Ruhter, and A. M. Epstein (2016). Readmissions, observation, and the hospital readmissions reduction program. *New England Journal of Medicine 374*(16), 1543–1551.

Tables and Figures

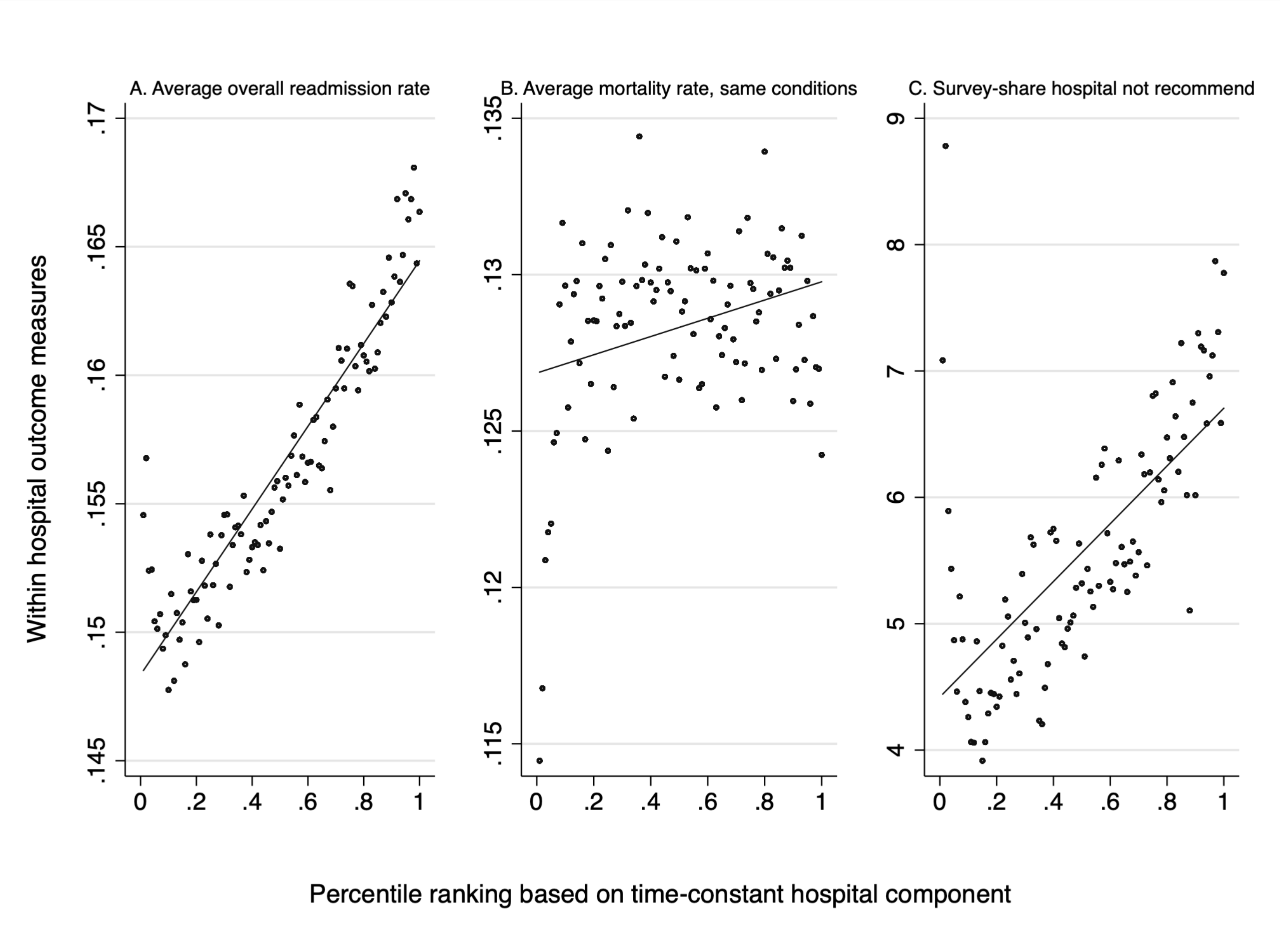
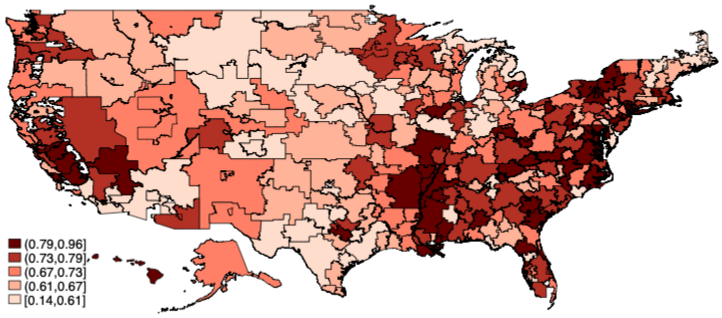


Figure 1: Within hospital readmission penalty propensity across diagnosis related groups, binned

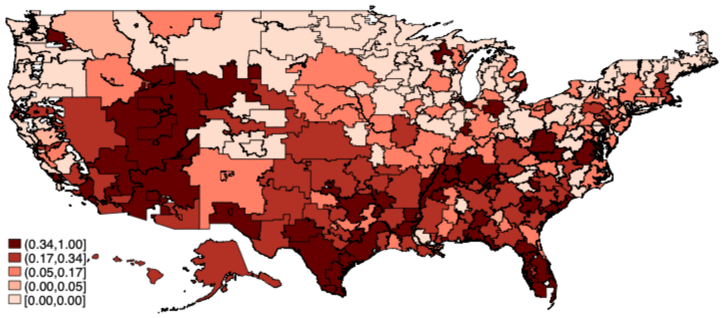
Note: Figure plots bins of hospital fixed effects and corresponding averages of other quality measures: overall readmission rate, mortality rate in respective emergency condition, patient survey responses. The x-axes are identical, and based on the pooled regression model (Column 4, Table 1). These estimated fixed effects are ranked and binned into 100 equal sized percentile ranks. Within each rank the y-axis quality measures are averaged. In Panel A - overall readmission rate in the hospital (b = .0161(se 0.0001) R^2: 0.856), B- the average mortality across time and the 3 diagnosis groups (b = 0.002(se 0.00017) R^2: 0.080), and C- the survey based share of people answering they would not recommend this hospital (b = 2.285(se 0.045) R^2: 0.447).

Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

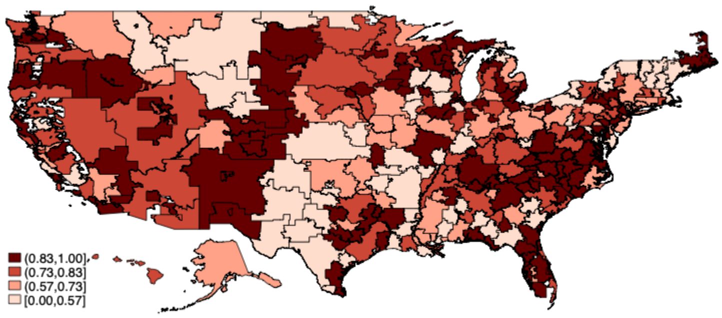
A. Average marginal penalty propensity (darker worse quality)



B. For-profit share (darker the higher the for-profit share)



C. Chain share (darker the higher the chain share)



D. Average HHI-beds (darker the more concentrated)

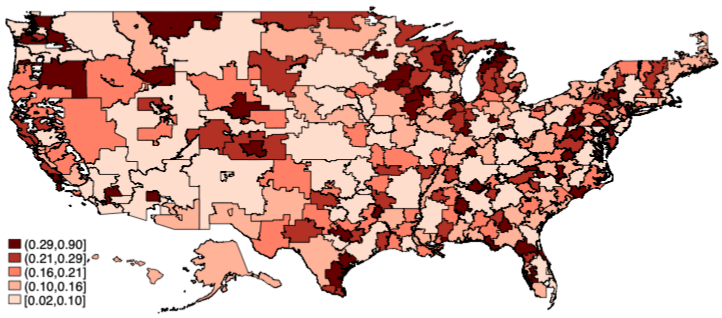


Figure 2: Average hospital adjusted readmission penalty propensities, for-profit share, and average HHI across hospital referral regions

Note: Figure plots, in Panel A average hospital penalty-propensity (for the marginal hospital) across the map of hospital referral regions (HHR), based on averages of marginal penalty propensity using the interactive fixed effects of Table 1, Column 5. Panel B depicts the share of for-profit hospitals in HRRs, C the share of chain hospitals, and D the average HHI based on the number of beds.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

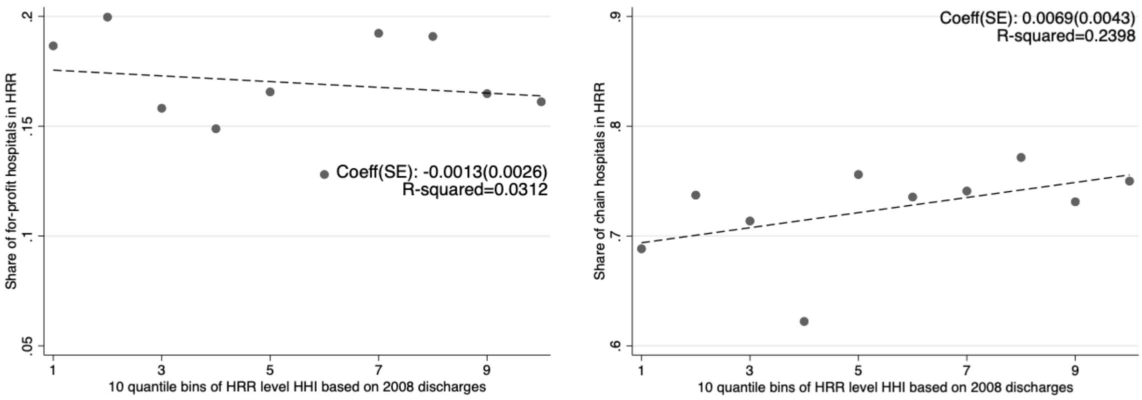


Figure 3: Association for-profit and chain penetration by local competition

Note: Figure plots in Panel a. the share of for-profit hospitals and the share of chain/chain hospitals in Panel b. over 10 quantile bins of the HHI (2008) distribution of the local HRR market, regression coefficient and R^2 of the linear fits are presented in the figure.

Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table 1: For-profit differential in hospital performance

	Unconditional Means			Fixed Effects: Covariate and Bias adjusted	
	For-profit	Other	Difference	Hospital-level	×condition
	(1)	(2)	(3)	(4)	(5)
Share would not recommend this hospital	7.327	5.213	2.114	1.471	
	(0.160)	(0.060)	(0.170)	(0.132)	
30-day mortality rate×100	12.920	12.878	0.043	0.039	0.067
	(0.051)	(0.023)	(0.055)	(0.044)	(0.069)
30-day raw readmission rate	0.199	0.196	0.003	0.001	0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Risk-adjusted (excess) RR	1.012	0.999	0.013	0.007	0.008
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Penalty	0.545	0.474	0.071	0.023	0.030
	(0.012)	(0.006)	(0.013)	(0.010)	(0.010)
Observations[a]	7,997	33,098	41,095	3,197	8,713
Hospital and area characteristics				✓	✓
Diagnosis indicators					✓

Notes: Table presents various measures of readmission performance: survey share would not recommend the hospital, the 30-day mortality rate (× 100), Raw readmission rate - RR, excess readmission ratio - ERR (CMS-risk-adjusted based on age, gender and co-morbidities), and an indicator for penalty status, whether ERR>1, by ownership status (for-profit and other hospitals). Column (1) presents averages (and standard error, clustered on hospital-level, in parentheses) across hospital×condition(AMI,HF,PN)×year(2011-2015) for For-profit and Column (2) for other hospitals. In Column (3), we present the unconditional mean difference of Column (1) and (2), for RR as well as ERR, and the marginal effect from a Probit model. Columns (4) and (5), present the coefficient of a regression of the For-profit indicator on fixed effects estimated in an auxiliary regression and includes time constant covariates (urban setting, teaching status, chain status, etc.). The auxiliary fixed effects are estimated in a pooled model one fixed effect for each hospital irrespective of condition (Column 4) and one fixed effect by hospital×condition (Column 5). For the mortality rate, raw readmission rate and the ERR, we first run an OLS regression on time-varying covariates (number of discharges, median household income in the county, etc.) and extract the fixed effects, which are then shrunken via Bayes-adjustment (Chandra et al., 2016), for the Penalty status we use an analogous procedure via bias-reduced Fixed Effects Probit model that embeds a shrinkage (Kunz et al., 2021). [a]Note that Observations refer to ERR as well as Penalty, for the RR six hospitals did not have a RR, about 1,000 have no mortality rate by condition (i.e. not reported due to small counts), and the survey does not vary by condition hence several observations are lower by roughly a third. Table C1 presents descriptive statistics by penalty status. More on the construction and definition of the variables used can be found in Appendix Table D1.

Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table 2: Marginal penalty propensity by for-profit, competition, chain status

Dependent variable: Marginal penalty propensity MPP_i^c							
			Local hospital market				
		Constant	HRR	Competition: HHI based on			
	All	Ownership	FE	-beds	-discharges	-chain	2008
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: For-profit status							
For-profit hospital (Yes/No)	0.030	0.031	0.052	0.033	0.035	0.020	0.036
	(0.010)	(0.011)	(0.009)	(0.015)	(0.014)	(0.017)	(0.013)
× HHI-measure				0.180	0.138	0.168	0.140
				(0.107)	(0.082)	(0.071)	(0.071)
For-profit + Interaction[a]				0.051	0.052	0.042	0.055
				p=.026	p=.016	p=.001	p=.005
Panel B: Chain status							
Chain hospital (Yes/No)	0.006	0.009	0.005	0.010	0.011	0.007	0.008
	(0.009)	(0.009)	(0.008)	(0.014)	(0.013)	(0.016)	(0.012)
× HHI-measure				-0.046	-0.051	-0.010	-0.034
				(0.088)	(0.074)	(0.068)	(0.067)
For-profit + Interaction[a]				0.006	0.005	0.006	0.003
				p=.646	p=.538	p=.957	p=.666
Observations	8,713	8,072	8,072	8,072	8,072	8,072	7,987
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
Area demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects			✓	✓	✓	✓	✓

Notes: Table presents OLS coefficients (standard errors clustered on hospital level in parentheses). Columns (1) repeats Column (5) from Table 1, Column (2) restricts the sample to those did not change ownership status, Column (3) conditions on HRR fixed effects. Columns (4)-(7) interact for-profit status with various measures of local competition, via HHI, (4) based on the concentration in the number of beds, (5) based on discharges, (6) based on discharges but on the chain level rather than individual hospital-level, and (7) uses lagged HHI on chain level in 2008. For-profit + Interaction[a], gives the interaction with the one standard deviation of the respective HHI measures, and p-value for the corresponding joint F-test. Panel B: repeats the analysis for the chain status instead of the for-profit indicator.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table 3: ROBUSTNESS AND SPECIFICATION TESTS

Dependent variable: Marginal penalty propensity MPP_i^c	Base	Above median HHI	Non-monopol markets	Non-mono-polistic	County share uninsured	Overall hospital	HSA FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: For-profit status							
For-profit hospital (Yes/No)	0.036	0.037	0.036	0.030	0.036	0.034	0.029
	(0.013)	(0.013)	(0.013)	(0.016)	(0.013)	(0.014)	(0.020)
× HHI-measure (2008)	0.140	0.036	0.140	0.229	0.140	0.127	0.248
	(0.071)	(0.018)	(0.071)	(0.126)	(0.071)	(0.075)	(0.103)
For-profit + Interaction[a]	0.055	0.073	0.055	0.062	0.055	0.051	0.063
	p=.005	p=.000	p=.005	p=.023	p=.005	p=.015	p=.002
Panel B: Chain status							
Chain hospital (Yes/No)	0.008	0.002	0.008	0.007	0.007	0.010	0.001
	(0.012)	(0.011)	(0.012)	(0.014)	(0.012)	(0.012)	(0.022)
× HHI-measure (2008)	-0.034	0.006	-0.034	-0.034	-0.035	-0.046	-0.035
	(0.067)	(0.016)	(0.067)	(0.120)	(0.068)	(0.068)	(0.136)
For-profit + Interaction[a]	0.003	0.008	0.003	0.003	0.002	0.003	-0.004
	p=.666	p=.530	p=.666	p=.808	p=.641	p=.554	p=.778
Observations	7,987	7,987	7,970	7,184	7,987	2,918	7,985
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects	✓	✓	✓	✓	✓	✓	

Notes: Table presents OLS coefficients (standard errors clustered on hospital level in parentheses), see Table 2 notes. Column (1) corresponds to Table 2 Column (7) - 2008 HHI measure, provided as reference. Column (2) replaced the interaction by using indicator for above median HHI, consequently For-profit + Interaction[a] here is a unit change (below to above median) in the HHI rather than the standard deviation. Column (3) drops fully monopolistic markets, and (4) hospitals in areas the top (10th) percentile of the HHI distribution. Column (5), includes the hospitals county share of uninsured people from the County Health Rankings (see Davis et al., 2020), (6) uses the overall hospital quality (instead of the condition-specific quality measures), (7) replaces the HRR fixed effects with the smaller HSA fixed effects where, HSA is a collection of ZIP codes whose residents receive most of their hospitalizations from the hospitals in that area.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table 4: MARGINAL PENALTY PROPENSITY BY FOR-PROFIT, COMPETITION, CHAIN STATUS: SUB-SAMPLES

Dependent variable: Marginal penalty propensity MPP_i^c

			Chain				
		Non-				Highest tercile	
	All	Chain	All	Small	Large	HHI^{chain}	wHHI^{chain}
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
For-profit hospital (Yes/No)	0.036	-0.059	0.044	0.049	0.009	-0.111	0.046
	(0.013)	(0.042)	(0.014)	(0.026)	(0.021)	(0.060)	(0.017)
× HHI-measure (2008)	0.140	0.216	0.145	0.121	0.239	1.201	0.637
	(0.071)	(0.306)	(0.078)	(0.209)	(0.101)	(0.398)	(0.394)
For-profit + Interaction[a]	0.055	-0.035	0.065	0.067	0.043	0.061	0.138
	p=.005	p=.571	p=.006	p=.376	p=.005	p=.001	p=.079
Observations	7,987	2,144	5,843	3,441	2,402	2,402	2,402
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects	✓	✓	✓	✓	✓		
HHI-chain main effect						✓	✓

Notes: See Table 2 - Column (7) Panel a - notes for details . Column (1) shows presents this for reference again. Column (2) restricts the sample to hospitals not organised in chain, and (3)-(6) to those in chain. Column (3) uses all chain organised hospitals, and (4) and (5) split them further by the highest tercile (33% - 20 or more hospitals in the chain). Column (6) and (7) replaces the HHI measure based on the average, and discharges weighted- HHI using only hospitals the chain faces. The full set of results across the different sub-samples is presented in Appendix Table C6.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

For online publication only.

A Institutional setting

Hospital readmissions have been identified as a major driver of health care costs (Jencks et al., 2009), but while costly at the aggregate level, discharging patients too early or not offering sufficient post-discharge care can be rational from the point of view of an individual hospital when reimbursements are based on diagnosis-related groups rather than actual costs. In an attempt to have hospitals internalize the costs of readmissions, the 2010 ACA established a financial penalty for hospitals whose Medicare readmission rates exceed a certain threshold in three common emergency conditions. Since the HRRP's introduction in October 2012, thousands of hospitals were fined and billions of dollars in fines were paid. The program has received considerable attention in economic (Gupta, 2017; Mellor et al., 2017; Zhang et al., 2016) and health services research (Bernheim et al., 2016; Desai et al., 2016; Joynt and Jha, 2013).

Beginning in 2009, hospital readmission rates have been published publicly on an annual basis by the Center for Medicare & Medicaid Services (CMS). Announced on March 23, 2010, starting on October 1, 2012 (for the financial year 2013), eligible hospitals were penalized by up to one per cent of their Medicare reimbursements if over the prior three-year period (i.e., from June 2008 to July 2011) there were higher-than-expected risk-standardized 30-day readmission rates for at least one of the three emergency conditions: *acute myocardial infarction* (AMI), *heart failure* (HF), or *pneumonia* (PN). All of these three are largely emergency conditions with a limited role for selection on hospitals quality (Chandra et al., 2016). It has been estimated that up to \$1 billion could be saved yearly by preventing these readmissions (McIlvennan et al., 2015).

Figure A1, depicts the changes in the HRRP over the sample period. The policy amount was repeatedly changed as shown by the maximum penalty cap.

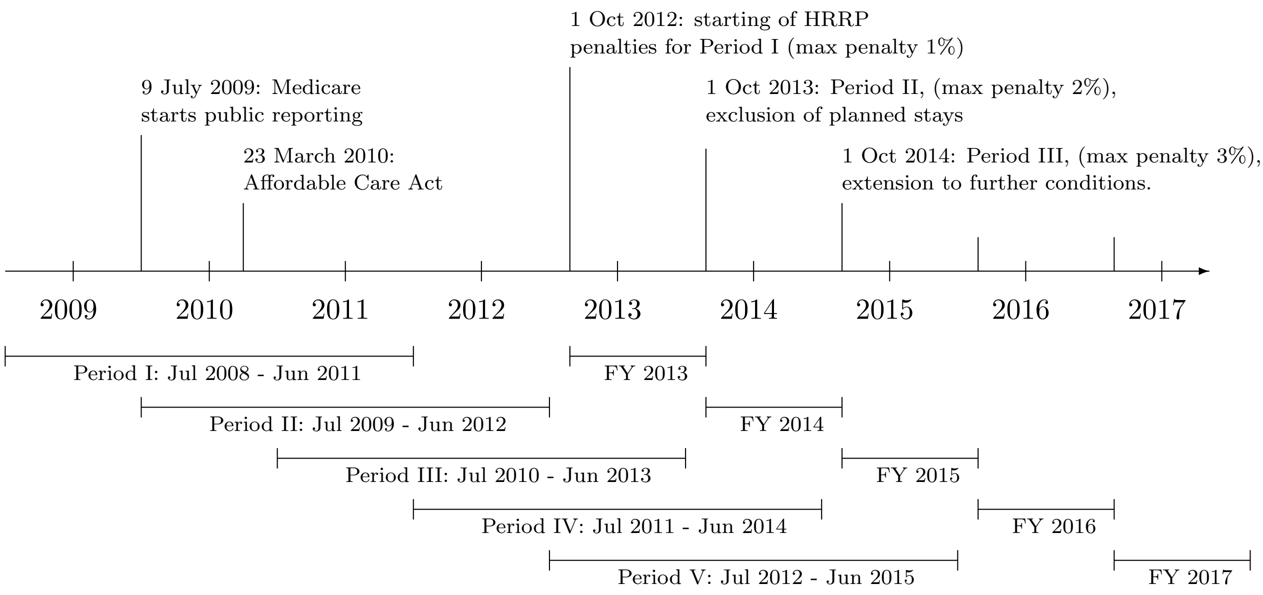


Figure A1: Event line and sample period of HRRP Policy
Source: Adapted from Figure 1 in Wasfy et al. (2017)

If a Medicare the patient was initially hospitalized—based on their primary discharge diagnosis—for one of the conditions and was readmitted to the same or another hospital within 30 days after release for the same or any other condition (all-cause), the patient counted as readmission for the initial hospital. Readmissions were then used to estimate the probability to be readmitted at a given hospital and compared to the average hospital with a similar case mix. Thus, risk-adjustment was based on the case mix, including age, sex, and co-existing conditions, but did not account for differences in socio-economic characteristics in the hospitals' environment. This changed after 2016, which is why we limit our sample period to before 2017.

Hospitals with at least 25 discharges for a diagnosis and part of the inpatient prospective payment system (IPPS) were eligible.[30] Their excess readmission ratio in emergency category c in year t measures the "total predicted readmissions [PRR] at a hospital [i] compared with the total expected readmission if the patients were treated at an average hospital with similar patients" (McIlvennan et al., 2015), i.e:

$$ERR_{it}^{c} = \frac{PRR_{it}^{c}}{E(RR_{it}^{c})}.$$

The resulting risk-adjusted *excess readmission ratio* ERR_{it}^{c} in emergency category c at year t measures the "total predicted readmissions at a hospital [i] compared with the total expected readmission if the patients were treated at an average hospital with similar patients" (McIlvennan et al., 2015). A penalty was imposed if this ratio of expected versus average was strictly greater than 1, ie. $ERR_{it}^{c} - 1 > 0$. The dollar amount of the penalty was calculated as 1 minus the "readmissions adjustment factor" —aggregate payments for these excess Medicare payments divided by aggregate Medicare payments for all discharges— multiplied by the hospital's base diagnosis related group payments, thus,

$$\text{Reimburstment adjustment} = 1 - \min\left[cap, \sum_{c=1}^{C} \max\{ERR_{it}^{c} - 1, 0\} \frac{Payment_{it}^{c}}{All\ payments_{it}}\right].$$

We focus on the extensive margin of a penalty—i.e. if $ERR_{it}^{c} - 1 > 0$—across any of the three conditions. The penalty cut-off is a policy-relevant discontinuity introduced into the hospitals' cost function (see, Gupta, 2017, for a detailed discussion).

Other interesting policy metrics are the conditional-on-penalty size of the readmission ratio and the second implied cut-off, which occurs at the upper end of the excess readmissions, where the payment amount does not increase further as it is capped, this *cap* maximum rate of penalty was one per cent in 2013, raised to two in 2014, and three in 2015, again complicating both the quality metrics over time.[31]

[30] Hospitals in Maryland are excepted from the policy due to a special Medicare agreement.

[31] In FY 2015, the conditions were extended to include *chronic obstructive pulmonary disease* (COPD), *elective hip or knee replacement*, and in FY 2017 to *coronary artery bypass graft* (CABG) and an extended pneumonia definition was put in place.

B Estimation details

We focus on the extensive margin of a penalty—i.e., if $ERR^c_{it} > 1$—across any, and separately for each, of the three conditions. This penalty cut-off is the policy-relevant discontinuity introduced into the hospitals' cost function and is thought to be more relevant than the incremental changes in the penalty amount (Gupta, 2017). Other advantages of assessing the penalty status as compared to its *amount* are that it is constant across years, independent from the relative reimbursements for that emergency condition and less likely to be affected by regression to the mean (Joshi et al., 2019).[32] Increases or decreases in the reimbursement amount might confound the time-invariant performance analysis. For instance, some hospitals might do very poorly, but since they have low relative reimbursements in that category, they could end up with the same penalty amount.

Furthermore, frequent policy changes complicate the assessment of penalty risk over time, for a detailed treatment of the policy see appendix A. In the HRRP the extensive margin of being penalized was constant over the sample period. In contrast, the intensive margin (penalty intensity) was adjusted frequently, see Appendix Figure A1. Further, the main policy interest/attention is the penalty/no-penalty status rather than the direct readmissions (Gupta, 2017; Zhang et al., 2016).

To assess the the penalty propensity, we use the baseline probit model, that is

$$P(y^c_{it} = 1 | x_{it}, \alpha^c_i) = \Phi(\alpha^c_i + x'_{it}\beta), \tag{4}$$

where y^c_{it} is an indicator of hospital i being penalized in year t for exceeding readmissions in emergency condition c. The interest lies in the fixed effects α^c_i, the quality measures for hospital i in condition c over the whole sample period. The vector x_{it} contains time-varying covariates at the hospital, hospital-referral-region (HRR) and county level.

Since there are only five available years, in the constant policy environment, standard maximum likelihood approaches for binary fixed effects panel data would disregard any hospital that does not change its penalty status. This would mean that a hospital in a very healthy environment could get the same ranking as one in a less advantaged area, if both are either never or always fined (this problem of *perfect prediction* is well known in the non-linear panel literature, see Maddala, 1983). In our data, this issue is quite severe, as there is a large share of hospitals that either is penalized in every period (outcome always equal to 1) or never (outcome always equal to 0).

In the case of a never-fined hospital, for instance, the implied maximum likelihood probability of a fine would be an exact 0; i.e., the event is deemed impossible. Given that there are only five-time periods it is unreasonable to think that not observing a penalty implies that no shock could ever occur to push the hospital over the penalty cutoff. Further, the maximum likelihood estimator disregards any environmental information for such hospitals, such that a hospital in a very healthy environment gets the same "ranking" as one in a less advantaged area if they are either never or always fined, which are rather unappealing model implications.

To overcome this issue and related issues discussed above, we estimate equation (4) using the bias-reduced (BR) generalized linear model estimator first suggested by Kosmidis and Firth (2009) and adjusted to the panel probit framework by Kunz et al. (2021). The model is discussed in detail in Kunz et al. (2021), briefly for the fixed effect panel probit model, the adjusted score function with respect to α_i is

$$\begin{aligned} s^{BRFE}(\alpha_i) &= \sum_{t=1}^{T} \left[y_{it} - \Phi(\eta_{it}) - \frac{1}{2} h_{it}\eta_{it} \frac{\Phi(\eta_{it})(1-\Phi(\eta_{it}))}{\phi(\eta_{it})} \right] \frac{\phi(\eta_{it})}{\Phi(\eta_{it})(1-\Phi(\eta_{it}))} \\ &= s^{ML}(\alpha_i) - \sum_{t=1}^{T} \frac{1}{2} h_{it}\eta_{it}, \end{aligned} \tag{5}$$

where $s^{ML}(\alpha_i)$ is that standard probit maximum likelihood score, η_{it} is the linear index, and h_{it} are the it-th

[32] Since our measure is based on the penalty, rather than the readmission ratio or penalty amount, regression to the mean is much less problematic. Almost half of the hospitals are either always or never fined, indicating a strong persistence in penalty status incompatible with regression to the mean. Our approach further circumvents regression towards the mean by explicitly accounting for the longitudinal dimension in estimation.

diagonal elements of the $NT \times NT$ projection matrix

$$H = W^{1/2}X(X'WX)^{-1}X'W^{1/2}, \tag{6}$$

with X the $NT \times (K+N)$ matrix of the K regressors and N individual-unit indicator vectors, and W is the $NT \times NT$ diagonal matrix with typical element $w_{it} = \phi(\eta_{it})^2/[\Phi(\eta_{it})(1-\Phi(\eta_{it}))]$. Kunz et al. (2021) show that the penalty that is added to the score function shrinks the fixed effects that are otherwise minus or plus infinity, towards 0, i.e. neither positive or negative quality relative to the marginal hospital. Intuitively, the values of the α_i's will use information on how often was a penalty observed, say one hospital is observed twice and another three times, both being always penalised. The three times observed hospital will be assigned a lower quality, *ceteris paribus*. Similarly, the α_i's will entail information on the local health environment, suppose two hospitals are observed equally often and both always penalised, if one is in a worse health environment it will get a better health rating as it perform better than what could be expected from its environment. As noted in the text, the estimator also deals with the small sample (short T) problem of estimating the hospital fixed effects with few observations and embeds an *ex-ante* shrinkage of the fixed effects towards 0, which would imply a latent hospital quality corresponding to a 50-50 propensity to be penalised net of covariates, as part of the estimation. Shrinkage estimation is an increasingly common tool to deal with these types of problems for metrics in health economics (Chandra et al., 2016; Propper et al., 2008) but also in other pay-for-performance metrics (Bonhomme and Weidner, 2019; Chetty et al., 2014). An advantage of this approach over the linear probability model is the model-consistent interpretation as the propensity (score) of penalty for a marginal hospital which respects the 0-1 bounds and bears a useful probability interpretation.

C Additional results

C.1 Descriptive statistics and auxiliary regression results

First, we present simple descriptive statistics of our sample Table C1. Table C2 shows which area characteristics predict ownership in our sample. The detailed fixed effect regression is presented in Table C3.

Table C1: Descriptives by penalty-status

	Sample means by penalty-status			Bivariate
	Never	Sometimes	Always	Correlation
	(1)	(2)	(3)	(4)
Hospital-level covariates				
For-profit hospital	0.146	0.203	0.226	0.072
	(0.004)	(0.003)	(0.004)	(0.005)
Hospital is part of a system	0.727	0.743	0.747	0.014
	(0.005)	(0.003)	(0.005)	(0.004)
Teaching Hospital	0.368	0.402	0.604	0.071
	(0.006)	(0.004)	(0.008)	(0.003)
Number of beds	209.878	206.669	255.656	0.000
	(1.873)	(1.169)	(2.361)	(0.000)
Average HHI-beds	0.143	0.124	0.116	-0.369
	(0.001)	(0.001)	(0.001)	(0.019)
HRR-level covariates				
Share of for-profit hospitals	0.178	0.202	0.199	0.115
	(0.002)	(0.001)	(0.002)	(0.011)
Number of hospitals per 100T capita	1.181	0.848	0.628	-0.045
	(0.017)	(0.008)	(0.009)	(0.001)
Hospital opening	0.273	0.271	0.260	-0.005
	(0.005)	(0.003)	(0.005)	(0.004)
Hospital closing	0.150	0.181	0.200	0.043
	(0.004)	(0.003)	(0.004)	(0.005)
Discharges for ACSC	28.493	31.834	34.632	0.010
	(0.087)	(0.058)	(0.092)	(0.000)
County-level covariates				
Share in poverty (all ages)	15.576	16.398	17.551	0.008
	(0.052)	(0.038)	(0.065)	(0.000)
Median HH income (in 10T$)	5.186	5.185	5.104	-0.005
	(0.013)	(0.009)	(0.016)	(0.001)
Population (in 100T)	6.498	8.603	11.132	0.002
	(0.147)	(0.115)	(0.208)	(0.000)
Unemployment rate	6.823	7.349	7.934	0.026
	(0.024)	(0.016)	(0.025)	(0.001)
Observations	9,772	22,428	8,895	41,095
Share	21.65%	54.57%	23.78%	

Notes: Table presents means (standard deviations). Columns (1)-(3) display averages of covariates by penalty status, Column (4) presents simple bivariate OLS coefficients of the variables with the penalty indicator. More on the construction and definition of the variables used can be found in Appendix Table D1.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table C2: Predicting ownership: For-profit, chain, and for-profit chain status

	For-profit	Chain	For-profit chain
	(1)	(2)	(3)
Percent ages 65 and older (2010, county)	0.004	-0.001	0.004
	(1.59)	(-0.32)	(1.82)
Rural area (county)	-0.010	-0.049	-0.012
	(-2.02)	(-8.44)	(-2.48)
Percent unemployed, in county	-0.009	0.003	-0.004
	(-2.66)	(0.70)	(-1.17)
Total population in 100'000, in county	0.000	0.000	-0.000
	(0.78)	(0.45)	(-0.31)
Household median income in 10'000$, in county	-0.018	-0.013	-0.015
	(-1.73)	(-1.15)	(-1.51)
Percent living in poverty, in county	-0.003	0.006	-0.002
	(-1.01)	(1.99)	(-0.96)
Discharges ACSCs per 1'000 enrollees, in HRR	0.003	-0.002	0.002
	(2.95)	(-2.58)	(2.07)
Average HHI in discharges	-0.058	0.047	-0.072
	(-0.95)	(0.69)	(-1.26)
Hospitals per head in HRR	-0.005	-0.009	-0.003
	(-0.69)	(-1.13)	(-0.55)
Percent less than high school (2010, county)	0.008	-0.006	0.006
	(3.73)	(-2.38)	(2.78)
Percent black non-Hispanic (2010, county)	0.001	0.000	0.001
	(1.84)	(0.28)	(1.17)
Percent Hispanic (2010, county)	0.003	-0.001	0.002
	(3.86)	(-1.17)	(3.26)
N	3,197	3,197	3,197
R^2	0.046	0.055	0.029

Notes: Table presents OLS coefficient estimates (and t-statistics in parentheses) from separate bi-variate regressions for each covariate. Columns (1) regresses For-profit status indicator on selected area level covariates, (2) chain status indicator, and (3) For-profit chain status indicator.

Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table C3 presents the detailed results from equation (2), and simple bivariate cross-sectional relationships, from Table 1 above (Column 3) for comparison. Column (1) shows coefficients of a simple bivariate linear regression of the penalty indicator on each of the variables indicated in the row names. A substantial HRR market-level variation is evident for, for example, the HRR-level covariates of ACSC discharges or closing hospitals. The (unconditional) county-level correlates of penalty risk are as expected: poverty, population and unemployment increase penalty risk, while median income reduces it.

Column (1) corresponds to Table 2 Column (3). Columns (2) and (3) to (4) and (5) respectively, present the results from our bias-reduced fixed effects Probit estimator and show coefficient estimates of selected covariates. Column (2) corresponds to a restricted (pooled) model, with separate sets of hospital fixed effects (α_i) and condition fixed effects (α^c). Column (3) corresponds to equation (2), which includes one fixed effect for each hospital-condition combination (α_i^c). The results show that most of the HRR and county-level associations become statistically insignificant and/or economically small once hospital fixed effects are conditioned upon. Only discharges for ACSC in the HRR seem to have a consistent effect on the penalty risk. Such discharges measure the HRR-level accessibility and effectiveness of local primary health care, which is outside the hospital's control but significantly affects their penalty risk, confirming cross-sectional evidence (Gu et al., 2014).

Table C3: ASSOCIATION OF PENALTY-RISK AND SELECTED COVARIATES

	Regressions		
	Bivariate OLS	BR probit with α_i FE	BR probit with α_i^c FE
	(1)	(2)	(3)
HRR-level covariates			
Hospital opening	-0.005	0.007	0.008
	(0.004)	(0.022)	(0.024)
Hospital closing	0.043	-0.047	-0.046
	(0.005)	(0.018)	(0.020)
Discharges for ACSC	0.010	0.025	0.026
	(0.000)	(0.004)	(0.005)
County-level covariates			
Share in poverty (all ages)	0.008	-0.004	-0.004
	(0.000)	(0.006)	(0.006)
Median HH income (in 10T$)	-0.005	-0.028	-0.025
	(0.001)	(0.041)	(0.045)
Population (in 100T)	0.002	0.036	0.031
	(0.000)	(0.037)	(0.041)
Unemployment rate	0.026	-0.011	-0.012
	(0.001)	(0.012)	(0.013)
Observations	41,095	41,095	41,095
Hospital characteristics		✓	✓
Year fixed effects		✓	✓
Diagnosis indicators		✓	
Hospital-condition fixed effects		3,197	8,173

Notes: Table presents bivariate-OLS coefficients (robust standard errors), and BR panel probit regressions (clustered standard errors in brackets). Columns (1) presents association (bivariate regression coefficient) of mean penalty on indicated characteristics, identical to the last column of Table C1. Columns (2) and (3), show the first-step of our main results, the BR probit regressions using individual fixed effects, in (2) pooled across conditions (one fixed effect per hospital), (3) one fixed effect for each hospital-condition pair. More on the construction and definition of the variables used can be found in Appendix Table D1.

Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

C.2 Assessing quality measure

Next, we present the density plots separately by condition and for the BR probit and OLS estimation approach, both exhibit a long high-quality tail, and a comparable between approximative-OLS and model-consistent BR probit.

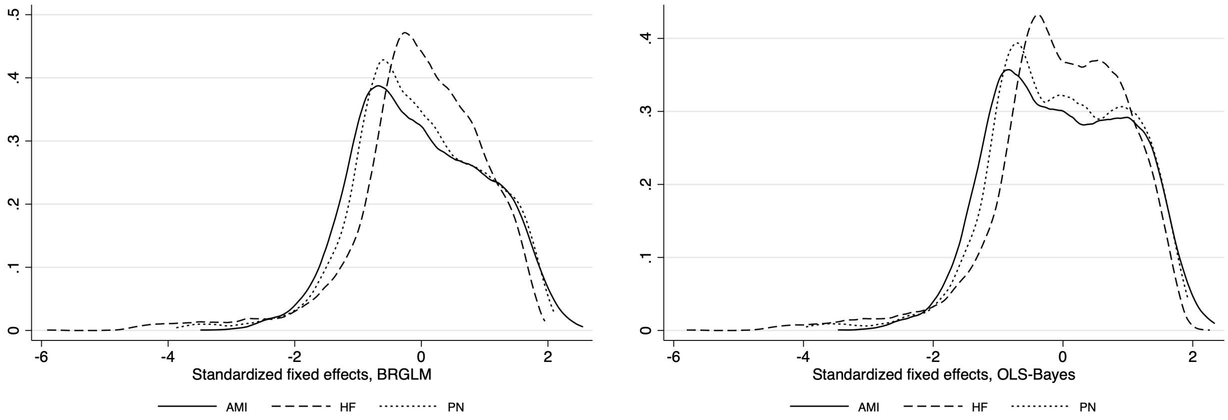


Figure C1: Density plot of separately estimated fixed effects, left Penalty-BR, right RR-OLS-Bayes

Note: Figure shows density plots of the hospital fixed effects across diagnosis conditions (AMI, HF, PN). We use the estimated fixed effects from eq. (2), ie. MPP_i^c from the separated model. Left graph is based on Penalty-BR probit model and right graph on the Penalty-OLS model.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Figure C2 shows the within-hospital correlation between the marginal penalty propensity across conditions. We bin these into twenty groups using 5%-penalty-probability increments for the different conditions; see below for the uni-variate-density, non-binned version, and fixed effects based on OLS. The strong correlation across conditions is consistent with a hospital-wide (or at least emergency-department-wide) component to quality, rather than with behaviour in which there is a trade-off between lowering risk in one condition at the expense of tolerating a higher risk in another.[33]

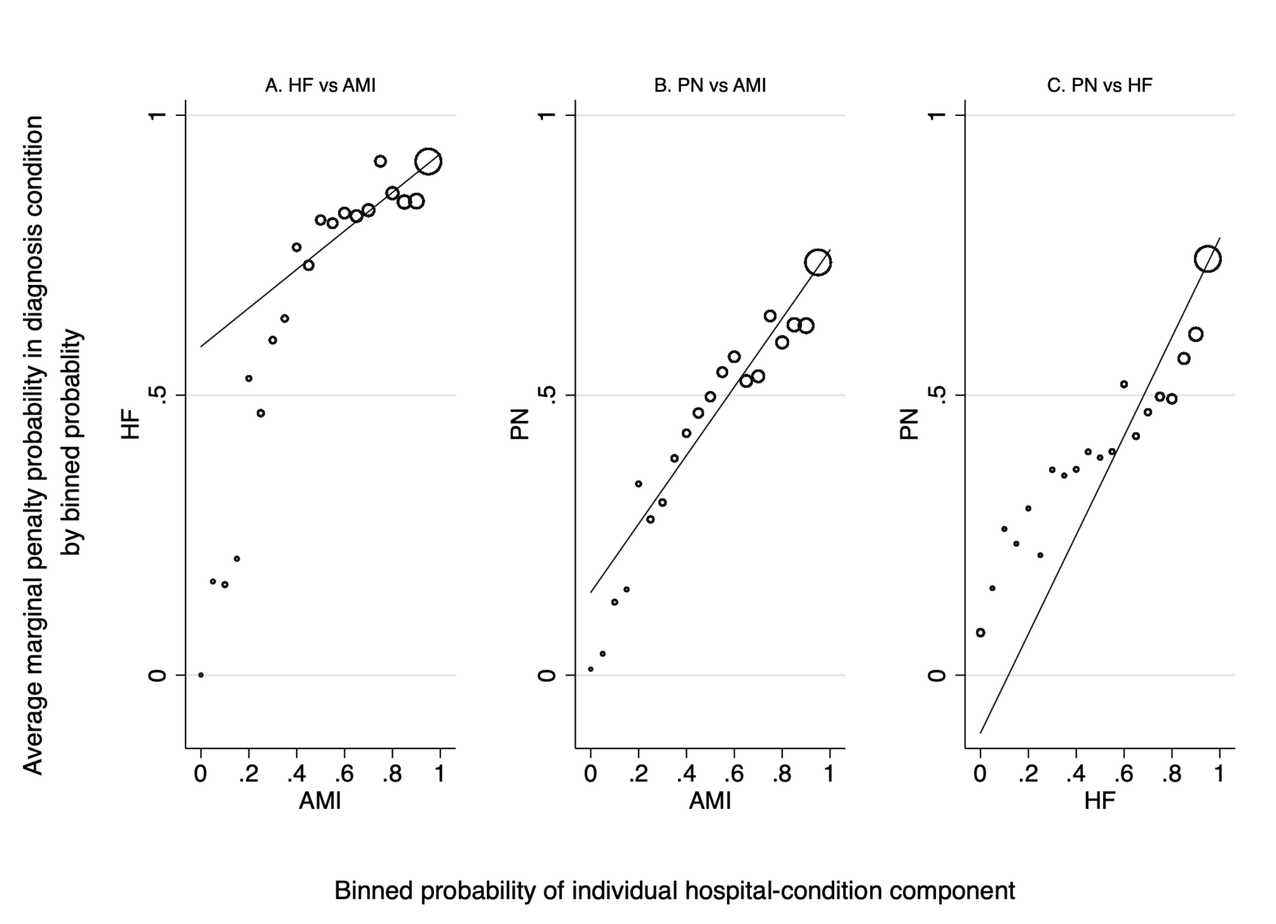


Figure C2: Within hospital readmission penalty propensity across diagnosis related groups, binned

Note: Figure plots correlation of hospital fixed effects across diagnosis conditions (AMI, HF, PN). We use the estimated fixed effects from eq.(2), ie. MPP_i^c from the interacted model. We separate these into 20 bins of 5 percent increments. Then we calculate the corresponding average penalty propensity in the condition given on the y-axis. For example, the first dot in Panel A, corresponds to the 5 percent best quality hospitals in AMI. Hospitals in this bin have an average propensity for a penalty in HF of almost 0, the same hospitals in panel B, are slightly higher but still close to 0 for PN. The size of the circles show the number of hospitals in this category. Figure C3 presents the non-binned raw α version as well as those based on analogous OLS estimates.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

[33] We tested whether there is evidence for additional condition-specific variation or whether all or most of the quality is constant within the hospital using a randomization test Abrams et al. (2012) that accounts for the uncertainty in the estimated fixed effects (results available on request). We find that PN differs significantly from HF but both are too similar to AMI to be distinguished, see also Kunz and Propper (2022).

Here, we assess the correlation in quality between conditions. Across each, there are highly related and the long high-quality tail is evident not only in the uni-variant comparison above but also across conditions. Again using the shrunken BR fixed effects probit or OLS with post-estimation shrinkage gives comparable results, yet, the long quality tail is much clearer when using the marginal penalty propensity (main text). Also, many of the top bins in this plot have a marginal propensity of 100%, which would need to be converted using the OLS results but is implicit in the BR in the main text.

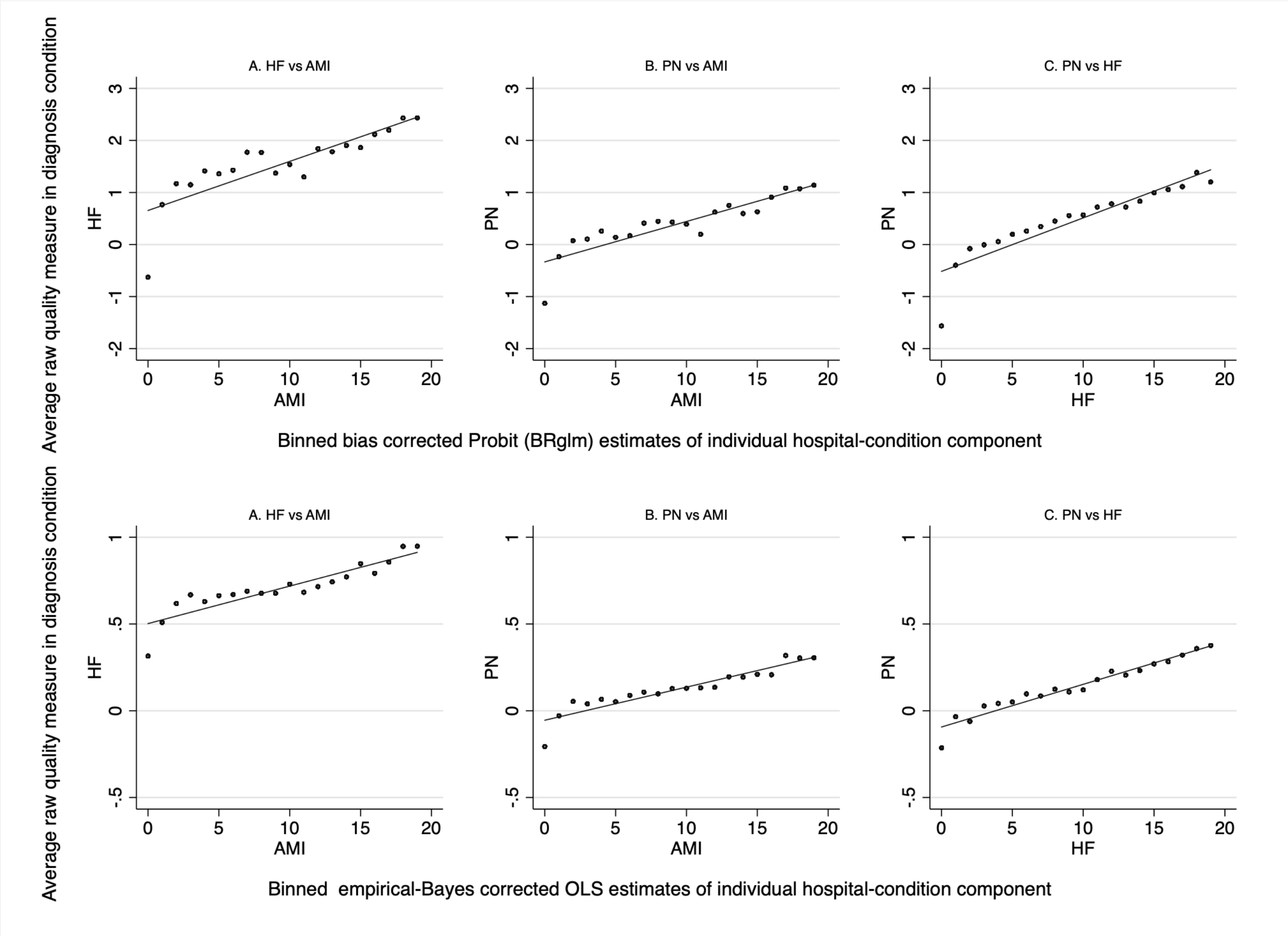


Figure C3: Robustness to Figure C2, equal sized bins of raw alpha and comparison with OLS fixed effects

Note: Figure presents the unrestricted fixed effects from regression eq (2) (top row) and analogues OLS regressions with post-estimation empirical Bayes shrinkage (bottom row). The fixed effects are clustered in 20 equal sized bins and the average performance (in the y-axis condition) is plotted.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table C4: ALTERNATIVE MEASURES

	Base		High quality tail		Penalty		Readmissions	
	$\Phi(\alpha_i)$	Rank(α_i)	10%	25%	Raw	Ebayes-OLS	Raw	ERR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel a: All hospitals								
For-profit hospital (Yes/No)	0.036	0.047	-0.004	-0.030	0.069	0.063	0.002	0.010
	(0.013)	(0.014)	(0.017)	(0.021)	(0.018)	(0.016)	(0.001)	(0.003)
× HHI-discharges (2008)	0.140	0.129	-0.282	-0.222	0.070	0.145	0.007	0.043
	(0.071)	(0.076)	(0.091)	(0.122)	(0.100)	(0.092)	(0.003)	(0.017)
For-profit + Interaction*	0.055	0.064	-0.042	-0.060	0.078	0.083	0.003	0.016
	p =.005	p =.009	p =.000	p =.020	p =.118	p =.011	p =.003	p =.000
Panel b: Large chain hospitals (highest tercile)								
For-profit hospital (Yes/No)	0.009	0.018	0.019	-0.011	0.020	0.036	0.001	0.006
	(0.021)	(0.022)	(0.028)	(0.034)	(0.029)	(0.026)	(0.001)	(0.005)
× HHI-discharges (2008)	0.239	0.218	-0.458	-0.325	0.227	0.245	0.008	0.053
	(0.101)	(0.109)	(0.148)	(0.207)	(0.145)	(0.133)	(0.005)	(0.024)
For-profit + Interaction*	0.042	0.049	-0.045	-0.057	0.052	0.070	0.002	0.014
	p =.005	p =.013	p =.001	p =.068	p =.055	p =.018	p =.039	p =.005
Observations^a	7,987	7,987	7,987	7,987	7,987	7,987	7,983	7,987
Observations^b	2,402	2,402	2,402	2,402	2402	2,402	2,402	2,402
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Notes: See Table 2, Column (1)-Panel a. replicates Column (7), Panel b is defined analogous to Tables 4's Col(5). In Column (2), re replace the marginal penalty propensity by the ranking of the hospitals latent quality. Columns (3) and (4) assess the high quality tail. Column (5)-(8) vary the measures we use and the various adjustments, (5) assesses raw penalty, (7) raw readmission rate, (8) excess readmission rate. In (6) we add an additional approach using Ebayes-adjusted Fixed Effects from an OLS regression a more standard approach in the literature that is relatively ad-hoc however.

Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table C5: Further robustness checks

		Weighting	Sub-samples			Method	Selection
	Base	by hospital size	No Government	10 year constant ownership	Only first 2 years	Fractional Response	Too few discharges
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel a: All hospitals							
For-profit hospital (Yes/No)	0.036	0.054	0.044	0.055	0.074	0.125	0.034
	(0.013)	(0.019)	(0.014)	(0.015)	(0.015)	(0.032)	(0.013)
× HHI-discharges (2008)	0.140	0.239	0.153	0.114	0.053	0.383	-0.051
	(0.071)	(0.094)	(0.087)	(0.079)	(0.075)	(0.183)	(0.052)
For-profit + Interaction*	0.055	0.087	0.065	0.071	0.082	0.071	0.027
	p =.005	p =.000	p =.011	p =.015	p =.055	p =.002	p =.684
Panel b: Large chain hospitals (highest tercile)							
For-profit hospital (Yes/No)	0.009	0.032	0.023	0.031	0.054	0.045	0.046
	(0.021)	(0.027)	(0.023)	(0.024)	(0.023)	(0.053)	(0.020)
× HHI-discharges (2008)	0.239	0.338	0.157	0.275	0.091	0.698	-0.121
	(0.101)	(0.110)	(0.130)	(0.114)	(0.117)	(0.271)	(0.083)
For-profit + Interaction*	0.042	0.079	0.045	0.070	0.066	0.057	0.029
	p =.005	p =.000	p =.116	p =.002	p =.168	p =.002	p =.278
Observations^a	7,987	7,987	6,729	7,352	7,812	7,987	8,620
Observations^b	2,402	2,402	2,264	2,111	2,364	2,402	2,638
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects	✓	✓	✓	✓	✓	✓	✓

Notes: See notes in Table 2, Column (1) repeats Column (7) for simplicity. Column (2) weights hospital×condition by thier first period number of discharges, which downweights small hospitals, Column (3)-(5) consider sub-samples: (3) drops government hospitals, (4) restricts the sample to those that did not have an ownership change between 2007 and 2016, and (5) estimates quality fixed-effects using only the first 2 years, which reduces the precision but leaves the main results intact. In (6) we re-estimate model (1) using fractional response model and in (7) we assess potential sample selction due to the 25-discharge rule, which we find not to be correlated with the core variables in our setup. Panel b. replicates the analysis in the subsample of large national chain hospitals (highest tercile, 20 and larger).

Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table C6: MARGINAL PENALTY PROPENSITY BY FOR-PROFIT, COMPETITION, CHAIN STATUS: SUB-SAMPLES

Dependent variables: Marginal hospitals penalty propensity $\Phi(\alpha_i^c)$								
			Local hospital market					Average
				Competition: HHI based on				
	All	Constant Ownership	HRR FE	-beds	-discharges	-chain	2008	Chain-level HHI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Sub-sample - non-chain hospitals								
For-profit hospital (Yes/No)	-0.069	-0.069	-0.043	-0.059	-0.060	-0.070	-0.059	
	(0.028)	(0.029)	(0.029)	(0.052)	(0.045)	(0.052)	(0.042)	
× HHI-measure				0.171	0.161	0.176	0.216	
				(0.399)	(0.257)	(0.247)	(0.306)	
For-profit + Interaction[a]				-0.043	-0.042	-0.049	-0.035	
				p=.751	p=.650	p=.604	p=.571	
Observations	2,318	2,153	2,153	2,153	2,153	2,153	2,144	
Panel B: Sub-sample - Chain hospitals								
For-profit hospital (Yes/No)	0.046	0.046	0.061	0.037	0.039	0.031	0.044	0.017
	(0.011)	(0.011)	(0.010)	(0.017)	(0.016)	(0.019)	(0.014)	(0.027)
× HHI-measure				0.230	0.192	0.159	0.145	0.334
				(0.120)	(0.099)	(0.079)	(0.078)	(0.192)
For-profit + Interaction[a]				0.061	0.062	0.053	0.065	0.065
				p=.013	p=.008	p=.003	p=.006	p=.035
Observations	6,395	5,919	5,919	5,919	5,919	5,919	5,843	5,843
Panel C: Sub-sample - Chain hospitals, less than 20 hospitals in the system								
For-profit hospital (Yes/No)	-0.002	-0.007	0.060	0.051	0.049	0.040	0.049	0.053
	(0.021)	(0.024)	(0.020)	(0.029)	(0.028)	(0.036)	(0.026)	(0.030)
× HHI-measure				0.094	0.113	0.120	0.121	0.067
				(0.250)	(0.217)	(0.200)	(0.209)	(0.218)
For-profit + Interaction[a]				0.061	0.062	0.057	0.067	0.063
				p=.527	p=.413	p=.349	p=.376	p=.541
Observations	3,746	3,508	3,508	3,508	3,508	3,508	3,441	3,441
Panel D: Sub-sample - Chain hospitals, more than 20 hospitals in the system								
For-profit hospital (Yes/No)	0.053	0.056	0.036	0.005	0.007	-0.009	0.009	-0.122
	(0.014)	(0.015)	(0.016)	(0.024)	(0.022)	(0.027)	(0.021)	(0.051)
× HHI-measure				0.304	0.268	0.236	0.239	1.151
				(0.176)	(0.145)	(0.104)	(0.101)	(0.341)
For-profit + Interaction[a]				0.035	0.038	0.023	0.042	0.039
				p=.050	p=.034	p=.007	p=.005	p=.000
Observations	2,649	2,411	2,411	2,411	2,411	2,411	2,402	2,402
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects			✓	✓	✓	✓	✓	✓

Notes: See Table 2 in combination with Table 4 notes for details, Panels A - D present various sub samples, based on chain status.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

D Data sources

FIPS crosswalk

We start by performing minor corrections on the file `CBSAtoCountycrosswalk_FY13.xls`[34] to the crosswalk between the county and FIPS State county, which is linkable to the hospital compare data (i.e. SAINT CLAIR we set equal to ST. CLAIR). Note, that island states such as AMERICAN SAMOA are dropped because we could not merge them to county nor HRR information.

County information

Using the FIPS indicators, we compiled for the years 2011-2015:
`Rural_Atlas_Update14/Jobs.csv` and `Rural_Atlas_Update14/People.csv` files[35] from which we get the variables such as: yearly unemployment rate, and yearly total population/100,000.

Next, we use the file `SAIPESNC_05APR17_15_02_58_98.csv`,[36] which provides yearly measures of all ages in poverty (in percent) and the median household income (in dollars/10,000). We then merge them via the FIPS crosswalk, all hospitals which could not be merged are included in the regressions with a missing indicator for the county.

Hospital Referral Region information

We use zip code crosswalks:[37] `ZipHsaHrr10.xls-ZipHsaHrr14.xls` from the *Dartmouth Atlas*, which allows us to connect the Zip codes to HRRs. We use the one year lagged values as hospital data is published with a lag. We calculate the number of hospitals for each year and define two indicators, one if there are more hospitals (in HRR) than in the previous year, and one if there were fewer. Note, that we can not distinguish, whether these are openings/closings of hospitals or a result of mergers or separations.

We use the number of Discharges for Ambulatory Care Sensitive Conditions from the selected medical discharge rates files:[38] `2010_med_discharges_hrr.Xls-2014_med_discharges_hrr.xls` where we subtract the conditions that are equal to our outcome measures (BacterialPneumoniaDischargesp and CongestiveHeartFailureDischar) form the total discharges (DischargesforAmbulatoryCareS). We then merge them via the zip code crosswalk, all hospitals which could not be merged are included in the regressions with a missing indicator for HRR.

Hospital Compare data

Our main data set is provided by the Centers for Medicare & Medicaid Services. More specifically, we use from the Acute Inpatient PPS:[39]

```
- FY 2012 Final Rule- IPPS Impact File PUF-August 15, 2011_1.txt
- FY 2013 Final Rule CN - IPPS Impact File PUF-March 2013.txt
```

[34] downloaded from http://www.nber.org/ssa-fips-state-county-crosswalk/ (accessed 26.03.17).

[35] downloaded from https://www.ers.usda.gov/data-products/atlas-of-rural-and-small-town-america/download-the-data/ (accessed 26.03.17).

[36] downloaded from https://www.census.gov/data-tools/demo/saipe/saipe.html (accessed 26.03.17).

[37] downloaded from http://www.dartmouthatlas.org/tools/downloads.aspx?tab=39 (accessed 26.03.17).

[38] downloaded from http://www.dartmouthatlas.org/tools/downloads.aspx?tab=41 (accessed 26.03.17).

[39] downloaded from https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/index.html (accessed 26.03.17).

- FY 2014 Final Rule IPPS Impact PUF-CN1-IFC-Jan 2014.txt
- FY 2015 IPPS Final Rule Impact PUF-(CN data).txt
- FY 2016 Correction Notice Impact PUF - (CN data).txt

The construction of these variables is taken from Gu et al. (2014). First, we use the information on the number of hospital beds which we include as 2 indicators for 100-399 beds and more than 400. Second, we use the resident to bed or daily ratio (rday) is larger than 0.25 as indicators of major teaching hospitals and lower than 0.25 but larger than 0 as minor teaching hospitals. Also urban if either *urgeo* or *urspa* indicate an urban area. These covariates are almost always constant within the hospital, for the very few minor changes we set them to the first observed state, to make them time-consistent.

Next, we use Hospital Compare data archive:[40]

- `HOSArchive_Revised_Flatfiles_20121001/Hospital_Data.csv` and `READMISSION REDUCTION.csv`
- `HOSArchive_Revised_Flatfiles_20131001/Hospital_Data.csv` and `READMISSION REDUCTION.csv`
- `HOSArchive_Revised_Flatfiles_20141218/Hospital General Information.csv` and `READMISSION REDUCTION.csv`
- `HOSArchive_Revised_FlatFiles_20151210/Hospital General Information.csv` and `READMISSION REDUCTION.csv`
- `Hospital_Revised_Flatfiles/Hospital General Information.csv` and `READMISSION REDUCTION.csv`

from which we get for each health conditions' READM-30-AMI-HRRP, READM-30-HF-HRRP, READM-30-PN-HRRP the excess readmission ratio, which we define as a penalty if larger than 1, we drop the hospitals with missing information in this (our key) variable. We use for each condition its corresponding number of discharges. Further, across the three conditions we calculate the total number of discharges leaving-out the current condition's discharges. Finally, the hospital's ownership is defined for-profit, if neither governmental nor non-profit (as above very minor changes, which we made time-consistent by taking the maximum observed value).

Note, that missing values in the readmission variable corresponds to "too few discharges" (less than 25). In robustness, see appendix Table C5, we use this as the dependent variable. Finally, we summarise our definitions of the key variables in the following table.

[40]downloaded from https://data.medicare.gov/data/archives/hospital-compare (accessed 26.03.17).

Table D1: Variable definitions

Variable	N	Mean	SD	Min	Max	Definition and construction	Source	Level
Main panel model, eq. (2)								
Penalty	41,095	0.49	0.50	0.00	1.00	Indicator whether there was a penalty issued in the respective condition	CMS	Hospital-condition-year
Number of discharges	41,095	320.46	295.24	0.00	3667.00	Number of discharges in hospital year condition	CMS	Hospital-condition-year
Total discharges leave out	41,095	616.35	548.44	0.00	6966.00	Sum number across ER conditions, leave out own condition	CMS	Hospital-condition-year
Number of beds	41,095	218.04	189.73	1.00	1928.00	Total number of beds	IPPS	Hospital-year
Discharges ACSC	41,095	31.65	8.95	0.00	66.36	Discharges for Ambulatory Care Sensitive Conditions - Regional measure of primary health provision, GU	DartmouthAtlas	HRR-year
Opening hospital, in HRR	41,095	0.27	0.44	0.00	1.00	Positive change in the number of ER providers one year to next	Own calculation	HRR-year
Closing hospital, in HRR	41,095	0.18	0.38	0.00	1.00	Negative change in the number of ER providers one year to next	Own calculation	HRR-year
All ages in poverty per cent	41,095	16.45	5.67	0.00	55.10	Local poverty rate	SAIPESNC	County-year
Median hhincome 10T dollars	41,095	5.17	1.39	0.00	12.59	Local median household income 10T dollars	SAIPESNC	County-year
Total population by 100T	41,095	8.65	17.25	0.00	101.70	Local total population estimate	Rural Atlas	County-year
Unemployment rate	41,095	7.35	2.42	0.00	28.90	Local unemployment rate	Rural Atlas	County-year
Other quality metrics						All averaged in hospital-condition		
Readmission rate	3,194	0.20	0.01	0.16	0.27	Condition-specific 30-day readmission rate	CMS	Hospital-condition-year
Excess readmission ratio	3,197	1.00	0.05	0.82	1.30	Condition-specific 30-day readmission rate	CMS	Hospital-condition-year
Overall readmission rate	3,154	0.16	0.01	0.12	0.20	Overall 30-day readmission rate of hospital	Sacarny-webpage	Hospital-year
Mortality rate	3,178	0.13	0.01	0.08	0.18	Condition-specific 30-day mortality rate of hospital	Sacarny-webpage	Hospital-condition-year
Patient satisfaction	3,186	5.57	3.03	0.00	29.93	Survey measure: Share would not recommend hospital	HCAHPS	Hospital-year
Fixed effect model, eq. (3)								
Measures of fixed effects	3,197	0.50	0.29	0.00	1.00	Extracted fixed effects - various standardisations, main Phi(alpha), also percentile rank, …	Own calculation	Hospital-condition
For-Profit	3,197	0.20	0.40	0.00	1.00	For-profit 0/1 Indicator whether it is a for-profit hospital (voluntary and government in ref. category)	IPPS	Hospital
Changed Ownership	3,197	0.08	0.27	0.00	1.00	A change in ownership status over sample period	IPPS - Own calculation	Hospital
System hospital	3,197	0.72	0.45	0.00	1.00	0/1 Indicator whether hospital is part of a system as defined in AHA	DartmouthAtlas	Hospital
HHI-discharges	3,197	0.15	0.13	0.01	1.00	HHI based on discharges in ER condition : $Average_t(\sum_{Hrr}(dischargesincondition_t/AllDischarges)^2)$	Own calculation	HRR-condition
HHI-bed	3,197	0.14	0.11	0.02	1.00	HHI based on number of beds: $Average_t(\sum_{Hrr}(nrbeds_t/AllBeds)^2)$	Own calculation	HRR
Size	3,197	0.77	0.64	0.00	2.00	Hospital size 1/3 How many beds categories following Gu et al.: 0-,100-,400- IPPS Hospital constant	IPPS, Own calculation	Hospital
Teaching status	3,194	1.40	0.64	1.00	3.00	Whether the hospital has a high resident-to-bed or rday, 0, positive, or is larger than 0.25	IPPS, Own calculation	Hospital
Urban	3,197	0.73	0.44	0.00	1.00	Urban 0/1 Whether the hospital is located in urban area	IPPS	Hospital
Share over 65	3,193	13.66	3.54	5.57	43.38	Share of county population in 2010 that are over 65	ACS	County
Share black	3,193	12.38	13.89	0.02	85.44	… black (non-hispanic white reference)	ACS	County
Share asian	3,193	3.50	5.02	0.03	43.01	… asian	ACS	County
Share hispanic	3,193	13.88	15.90	0.36	95.74	… hispanic	ACS	County
Share high school only	3,193	29.66	7.46	8.30	51.68	… high school degree (less than HS, reference)	ACS	County
Share some college	3,193	21.25	3.59	8.67	33.76	… some college	ACS	County
Share associate degree	3,193	8.03	1.97	2.04	16.66	… associate degree	ACS	County
Share college and more	3,193	27.04	10.71	5.11	72.88	… college or more	ACS	County
Share uninsured	3,193	0.20	0.07	0.00	0.48	Share of county population in 2010	Health Ranking	County