Assessing the Quality of Public Services: Does Hospital Competition Crowd Out the For-Profit Quality Gap?*

Johannes S. Kunz — Monash University

Carol Propper — Imperial College London

Kevin E. Staub — University of Melbourne

Rainer Winkelmann — University of Zurich

June 4, 2021

Abstract

We examine variation in US hospital quality across ownership, market concentration and membership of a hospital system. 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 competition erases a substantial part of the gap. In particular, we find that the reduction of the gap occurs in hospitals that are part of large national chains. For such hospitals, we estimate that in a fully competitive market the gap is completely erased.

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

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 system hospitals experienced a similar increase. Thus, as the industry evolves, the long-standing question of whether for-profits and not-for-profits behave differently has become highly salient (e.g. Capps et al., 2017).

A large body of theoretical literature suggests that for-profits and not-for-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., 2019). While chains have been less studied, Eliason et al. (2019) found that the quality of dialysis care decreased in hospitals after acquisition by a chain.

None of these papers have devoted systematic attention to the interaction between ownership, market and chain status. However, the early empirical research indicated an important interaction with market structure: when for-profits and not-for-profits operated 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 not-for-profits behave as for-profits (e.g. Chandra et al., 2016; Zhang et al., 2016). Given the large changes in the industry since this research was conducted, it is timely to return to the question of whether not-for-profits have systematically different quality from for-profit hospitals and to focus directly on how this varies with both competition and chain status. Indeed, the acceleration in the expansion of chains during the Covid pandemic is attracting attention from federal lawmakers pushing for greater oversight of hospitals (NYTimes, 2021).

To address this, we examine differences in quality between for- and not-for-profit US hospitals using a

[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).

[2]Eggleston et al. (2008) provide a systematic review of 31 observational studies and find that studies representative of the US as a whole tend to report lower quality among for-profits than private non-profits and most adverse effects are located in for-profits, but, the size and significance depend on the populations and samples considered (cf., Baltagi and Yen, 2014; Colla et al., 2016; Cooper et al., 2018; Dranove and Satterthwaite, 2000; McClellan and Staiger, 2000; Paul et al., 2019; Picone et al., 2002).

measure derived from the risk of being fined under a flagship program designed to incentivize hospital quality. We exploit publicly available measures of quality for the period 2012-2016 from the national Hospital Readmission Reduction Program (HRRP). This is 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).[3]

Such publicly available measures are observable to all players in the market and may therefore be expected to affect the behavior of sellers and buyers.[4] Measures of quality that relate to incentivized metrics such as those in HRRP should reflect heterogeneity in strategic decisions (Mehta, 2019) across different types of hospitals. However, while penalties affect the trade-off between treatment cost and readmission probability, it is by no means clear whether zero, or very few, readmissions are optimal from the point of view of a hospital's management. If the costs of avoiding readmissions are high, hospitals engaging in optimizing behaviour 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 not-for-profits. Competition may crowd-out such strategic flexibility (Duggan, 2000; Gaynor et al., 2015).

Chain status means that hospitals are subject to central as well as local oversight. It is not immediately obvious how this might affect the difference between chain and non-chain hospitals with respect to quality and the association of for-profit status with local market competition. The market power derived from chain status may mean that chain hospitals can have lower quality (as in Eliason et al., 2019; Gupta et al., 2021). However, this does not mean the quality gap between for-profit and non-profit chains is necessarily smaller. Chain for-profits may put more weight on costs than quality compared to non-profits because they need to make a profit and have ownership by groups who often have short-term objectives. Raising quality may require longer-term investments.

Whether local competition disciplines quality among chains has been questioned by Eliason et al. (2019) for the dialysis market, though this is almost exclusively for-profit, however. Gupta et al. (2021) find more monopolistic local markets exacerbate the poor quality provision by private-equity nursing homes but differences are small. Given that the for-profit negative quality gap was found to be driven out by competition in earlier literature, and that chain hospitals have larger market power,

[3]Earlier literature which found a lack of difference between for-profit and not-for-profits focused largely on readmissions or mortality (Gaynor and Town, 2011), often using AMI, and are therefore directly comparable.

[4]It has been shown in several settings that patients use these measures when choosing hospitals (Chandra et al. 2016, Emmert and Schlesinger 2017, Varkevisser et al. 2012, and Saghafian and Hopp 2020).

one might expect chain for-profits to be insulated from local competition. Thus whether the chain's national market power allows for-profit hospitals lower quality provision in highly competitive local markets is an empirical question.[5]

Measuring hospital quality with measures derived from the HRRP has several benefits. First, 'excess readmissions' are the measure the government uses to incentivize higher quality. Second, the penalties for having excess readmissions can be large: 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, 2017). Thus, these measures of quality are likely to be a choice variable for hospitals (Garthwaite et al., 2020). Third, the metric is based on admissions for three primarily emergency admission conditions and the exact penalty cut-off is not observed by the hospitals (Gupta, 2017), limiting the scope for self-selection by patients and hospitals.[6] Additional evidence for the usefulness of these performance metrics is confirmed by them being predictive of market-shares (Chandra et al., 2016) and, at the regional level, being correlated with deaths from COVID-19 (Kunz and Propper, 2020).

However, potential sources of error remain when using either the excess-readmissions rates or the excess readmissions-based penalty status as a performance metric. First, readmissions are likely to be subject to idiosyncratic variation unrelated to quality (inherent in small count measures) and to be affected by demographic factors beyond those used in the published standard risk-adjusted measures. In the HRRP case, they may also exhibit regression to the mean (Joshi et al., 2019). Second, for published penalty rates, simply focusing on whether a hospital is fined or not as a measure of quality encounters the problem that many hospitals are either never or always fined: 45% of hospitals were never or always penalized within the first five-year period of the HRRP. While we examine the raw and adjusted readmission rates in our initial descriptive analysis to be comparable to other research, we use as our preferred measure a medium-run measure of the hospital's latent propensity to be penalised, which addresses these statistical issues. We derive the measure from hospitals' HRRP-penalty status (whether or not a hospital was fined) adjusted for low volumes and demographic factors in the local hospital market. This overcomes the problems outlined above and follows a large literature on extracting unobserved heterogeneity from incentivized performance metrics, such as teacher (fixed) effects in value-added models (Chetty et al., 2014). We show these emergency-condition-

[5] 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.

[6] The lack of self-selection in these measures is validated in Doyle et al. (2019). For the same argument see Garthwaite et al. (2020). Other approaches to correct for patient selection include Hull (2016) and Geweke et al. (2003).

specific measures of quality are also strongly associated with a set of non-incentivized and less directly incentivized measures of quality — the overall readmission rate in all emergency conditions, mortality in the target conditions, and patient satisfaction — which arguably capture more direct dimensions of quality than the propensity to be fined.

Regardless of whether we use the published measures of excess readmission rates, rates adjusted for risk (Gu et al. 2014, Herrin et al. 2015), or our measure of the underlying penalty propensity, we find that for-profit hospitals are of lower quality. System hospitals (roughly 75% of hospitals, AHA, 2019) do not have significantly different quality from non-system hospitals.

However, we find important interactions with market structure and system status. Local market concentration differences are associated with much of the negative quality difference between for-profits and not-for-profits. Roughly 35% of the quality gap can be attributed to competition in the hospital's market. System status per se does not alter the for-profit quality gap or the interaction with market structure very much. However, for those hospitals which are part of a large system (defined as containing over 20 hospitals nationally), the interaction with local market structure completely drives out any negative association between for-profit status and quality. In other words, for- and not-for-profit system hospitals have very similar quality when part of a large system and operating in competitive local markets, but the more monopolistic the local market the larger is the for-profit quality gap.

Our finding that market structure matters for quality echoes the analysis for price by Cooper et al. (2019), who find that market concentration is strongly associated with price in hospital markets. It contrasts with a study of quality provision in nursing homes (Hackmann, 2019), which finds that financial reimbursements are more important than competition in improving patient outcomes. Our findings of an interaction between system status and competition relate to a recent literature on chain status and private equity in the health care market. Eliason et al. (2019) found that chain status has a negative effect on quality, but one that does not react to local competition (but does not distinguish between ownership type). Gupta et al. (2021) find that private equity acquisitions lowered quality (increased mortality) and increased costs considerably in the nursing home market.

In sum, our findings show that for-profit status and being part of a large system are associated with lower quality and that there are important interactions with local market concentration. For hospitals that are part of a large system, the negative for-profit quality gap is completely eliminated at low levels of market concentration. Our findings indicate that if for-profits were located in the same markets as not-for-profits their penalty propensity would be even larger, suggesting either that for-profits (and particularly for-profits that are part of systems) choose to locate in more concentrated markets or they engage in behaviour that over time makes these markets more concentrated. In this paper, we cannot distinguish between static and dynamic effects, but our results indicate that anti-trust decision-makers need to pay attention to both ownership and market structure when considering hospital merger and acquisition cases.

2 Estimation and data

2.1 Measures of quality

We begin by constructing measures of quality. 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]. The penalty is applied if the ERR is above 1 in any of the three conditions. This cut-off is the policy-relevant discontinuity introduced into the hospital's cost function (Gupta, 2017). These measures have been criticised because that risk-adjustment is only performed on age, sex, and other health conditions (co-morbidities) of the patients and do not take into account the differences in socioeconomic characteristics these hospitals serve (e.g., Aswani et al., 2018; Gu et al., 2014), 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 this, we extract latent quality via fixed effect measures for each hospital and each condition, adjusting for covariates and dealing with the issue of small numbers of observations (details in Appendix B). To do this, we use fixed effects regressions for OLS for the readmission rate and the ERR and probit for the penalty-status:

$$\begin{aligned} E(y_{it}^c|\alpha_i^c, x_{it}) &= \alpha_i^c + x_{it}'\beta && (1) \\ E(Penalty_{it}^c = 1|\alpha_i^c, x_{it}') &= \Phi(\alpha_i^c + x_{it}'\beta), && (2) \end{aligned}$$

where y_{it}^c denotes the raw readmission rate or the ERR, and $Penalty_{it}^c$ 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, as well as year indicators to control for common time shocks. We cluster standard errors at the hospital level.

Our interest is 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); third, estimating the regressions separately by condition.[7] To deal with large and imprecise outliers, researchers often shrink

[7] Each approach leads to qualitatively similar conclusions.

the estimated fixed effects towards their mean using Empirical Bayes.[8] We follow this literature for the OLS fixed effects estimates in (Eq. 1).

For the penalty status measure (Eq. 2), we use a bias-reduced fixed effect probit approach (Kunz et al., 2018). This method's bias-reduction shrinks the α_i^c during estimation. This serves the same purpose as Empirical Bayes for the linear model, but it also removes incidental parameter bias and avoids the perfect prediction problem. The latter means that any information in the covariates would be discarded when hospitals are always or never fined. For such cases, shrinkage after estimation, such as with Empirical Bayes, would not be possible as no finite estimates of α_i^c are obtained. In addition, Empirical Bias often involves assumptions such as homogeneity which cannot be satisfied with binary outcomes (Frederiksen et al., 2019). In contrast, the bias-reduction approach we pursue avoids any additional assumptions used for Empirical Bayes:[9] the method automatically shrinks the fixed effects towards the conservative benchmark of the marginal hospital that has neither positive nor negative relative quality. The predicted the fixed effects from Eq. (2), $\hat{\alpha}_i^c$, are used to calculate the marginal penalty propensities, $\Phi(\hat{\alpha}_i^c)$, which can be thought of as the propensity to be fined if the hospital would otherwise be marginal ($x_{it}'\hat{\beta} = 0$). The estimated $\Phi(\hat{\alpha}_i^c)$ are model-consistent (they respect the 0-1 bounds without any *ad hoc* adjustments).

Figure 1 presents the association of these quality measures with other non-incentivized hospital-wide quality metrics (we use one fixed effect per hospital to obtain a single hospital-level quality measure). The three graphs show that the rank in the hospital quality measure (grouped into 100 bins) correlates, respectively, with the overall readmission rate,[10] the average hospital mortality (in the same condition), and patient satisfaction.[11] All three of these measures correlate positively with our quality measure; the readmission rate and patient satisfaction very strongly so. The strong correspondence between our objective measure and a subjective one (the patient survey responses) is notable. *Ex-ante* it is by no means clear that patients would value the same quality dimensions as those related to readmissions or even mortality. Taken together, these results suggest that the extracted penalty propensity measures

[8]This approach has become ubiquitous in the literature on hospital quality (Chandra et al., 2016; Hull, 2016; Propper et al., 2008) and in other pay-for-performance metrics (Bonhomme and Weidner, 2019; Chetty et al., 2014).

[9]These 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). However, Appendix Table C5, 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.

[10]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).

[11]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 different hospitals and years. These issues compromise its comparability.

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).

2.2 Assessing quality differences

Our analysis assesses whether there are differences in quality between for-profit and not-for-profit hospitals and how this interacts with (local) market structure and chain status.[12] Using our preferred measure of hospital quality, we estimate OLS regressions of the form:

$$\Phi(\hat{\alpha}_i^c) = \gamma_1 \text{For-profit}_i + \gamma_2 \text{For-profit}_i \times \text{HHI}_{i_{hrr}} + z_i' \beta + \delta_i^c + \delta_{i_{hrr}} + \varepsilon_i, \tag{3}$$

where z_i contains (almost) constant measures of hospitals' teaching status, urban, size, and county-level covariates. δ_i^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 HHI, the Herfindahl-Hirschmann index. We repeat this analysis replacing for-profit and not-for-profit in (3) with system/non-system. 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) system.

2.3 Data

Our data are from the administrative Hospital Compare dataset which gives information on penalties for 2012–2016. Reporting is delayed by one year, so the data relate to the three-year aggregates of readmissions during the years 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 (2) we extract from the penalty status indicator 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. 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 region (Chandra et al., 2016). At the county-level, we use the Federal Information Processing Standard to add community characteristics, such

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

as the poverty rate and the median household income, which have been discussed as determinants of readmission rates outside the control of the hospital (Herrin et al., 2015). Most hospital-level variables display little or no variation over time. Thus, they can only be included in the second step estimation (3). We use urban/rural, teaching status of the hospital, size (number of beds), for-profit, and system status from the corresponding final rule impact files. To measure local market (HRR-level) concentration, we construct a standard measure of competition, the HHI based on the number of beds. We assess robustness to this in three ways: using HHI based on discharges in the respective emergency conditions, treating hospitals that belong to the same chains as one hospital (a similar approach was used in physician markets, Hausman and Lavetti, 2021), and recalculating the HHI from discharges before the introduction/announcement of the HRRP (in 2008). Variables and data sources are in Appendix D, descriptive statistics in Table C1.

3 Results

3.1 Readmission performance by hospital ownership

Table 1 presents the differences between for- and not-for-profits in readmission rates, risk-adjusted ERR, and the penalty indicator (ERR>1). Columns (1)-(3) show the unconditional means and differences pooled across years and conditions. The first row shows for-profits have a statistically significant 0.3 percentage points (p.p.) higher rate of readmissions, which corresponds to around 10% of the total reduction in readmissions over the sample period (and was considered a success Zuckerman et al., 2016, which fell from 21.5% to 17.8%). The second row presents the CMS measure, which adjusts for gender, age, and comorbidities, 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 comorbidities (one strategy by which hospitals reduce their penalty risk, Ody et al., 2019). The third row shows the quality difference as measured by penalty status. Column (3) shows that for-profits have a 7.1 p.p. higher probability of being penalised than not-for-profits.[13] These raw correlations show that the penalty appropriately corresponds to the differences observed in the raw readmissions, and that for-profits have a lower quality in terms of readmissions and consequently a higher likelihood of being penalised.

[13]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 benchmarking embedded in the ERR and Penalty status, respectively, that 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 3 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 in Appendix Table C6.)

Columns (4) and (5) presents the medium-run fixed effects estimated from equation (1) and (2), so accounting for a large number of time-varying covariates and shrinkage. Column (4) pools the three target conditions. The for-profit gap in all three measures remains large and statistically significant, although smaller than in the unadjusted comparison. 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. Given that there are 20% for-profit hospitals and the share is growing this differential is likely to grow in importance.

Column (5) shows the results estimating one fixed effect for each hospital×condition, which allows for inherent differences across the three conditions. The negative 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%).[14]

3.2 Quality, ownership, and competition

Table 1 shows qualitatively similar results across different measures including our preferred measure of penalty propensity, so all subsequent analysis focuses on this measure. We begin by presenting the spatial distribution of quality. Figure 2, Panel A, shows substantial spatial heterogeneity in hospital quality (net of regional socioeconomic indicators as these are already accounted for in our measure) across HRRs (see also Finkelstein et al., 2016). The estimated propensity of being penalized at HRR level varies from 15% to 95%, with an interquartile penalty-propensity difference of nearly 15 p.p. In Panel B, we show that for-profits, perhaps surprisingly, are located in lower-penalty areas.

To get a sense of the magnitude of this sorting, we calculate a descriptive Blinder-Oaxaca decomposition. Using a simple expansion around the mean difference in penalty status, the for-profit [FP] gap can be written $(\bar{x}^{FP} - \bar{x}^{NFP})\beta^{FP} + \bar{x}^{FP}(\beta^{FP} - \beta^{NFP})$, where $\bar{x}$ contains either HRR fixed effects or the HRR-level concentration index. The first term captures the difference in sorting of hospitals into areas and the second the difference in response to a given level of market concentration. When including only HRR fixed effects in x the results show that if for-profits were located in the same areas as non-profits they would have an even higher penalty propensity of approximately 5.6p.p. (instead of the 3p.p. we find in Table 1).

Replacing the HRR fixed effects with the HRR-level HHI index, the decomposition suggests that for-profits do not appear to select on competition. The mean difference in the competition for- and not-for profit hospitals face is close to 0, making the first term close to zero. In contrast, the 'response' to

[14] Appendix Table C2 presents the covariates in the regressions and Table C3 each condition separately.

competition differs markedly: a large part of the FP-gap is explained by differential 'behaviour' with respect to competition ($\beta^{FP}_{HHI} - \beta^{NFP}_{HHI}$). This suggests that for-profits have similar quality to other hospitals when facing similar market concentration.

Table 2 shows the differences in quality associated with competition by profit (Panel A) and chain (system)-status (Panel B). Column (1) uses the full model in which there is a fixed effect for each hospital-condition. This corresponds to Table 1, Column (5), third row, but additionally controls for time-constant hospital and area characteristics. The Panel A result indicates that the marginal probability of incurring a penalty (i.e., having lower quality) is 3 p.p. higher for for-profit hospitals. A small number of hospitals changed their for-profit status during the observation period and Column (2) omits these hospitals from the estimation sample. The negative for-profit gap remains the same (3.1 p.p.). Column (3) additionally controls for the full set of HRR fixed effects. The for-profit gap increases to 5p.p. Thus, the within HRR gap is larger than the across HRR gap, mirroring the unconditional Blinder-Oaxaca decomposition results.

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, continuing to condition on HRR fixed effects to absorb the baseline level of competition. Column (4), Panel A, shows the gap between a for- and not-for-profit facing high levels of competition falls to 3.3 p.p. However, at the mean market concentration ($\overline{HHI} = 0.149$) the for-profit differential in the penalty propensity is 5.1 percentage-points (shown in the For-profit + interaction). These results are robust to other measures of HHI, based on discharges in respective emergency conditions (Column 5), treating chains as one competitor in the local market (6), and using a measure of market concentration before the policy enactment (Column 7).[15]

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 chain status *per se* is not associated with a significant difference in quality nor is there an interaction with market concentration.

Table 3 examines the for-profit quality difference analogously for various subsamples. Panel A shows non-system hospitals, Panel B system hospitals, and Panels C and D split these further into hospitals which are part of nationally small and large systems respectively. Panel A shows that among non-system hospitals there is no significant difference in quality between for- and not-for-profits and nor does it vary with concentration. Panel B shows that among system hospitals there is a sizable for-profit gap. Panel B, Column (3), shows that for-profit hospitals have a 6 p.p. higher latent propensity to be fined. Column (4), corroborated by the robustness tests in Columns (5)-(7), shows a strong interaction

[15]The results hold also for the percentile ranking, which does not depend on the magnitude difference in the penalty propensity, as well as an indicator for being in the top 10% hospitals, cf. Appendix Table C4.

with market concentration. A for-profit hospital that is part of a system located in a market with the mean HHI has a 6.1 p.p. higher penalty propensity. System status thus accounts for the for-profit gaps shown in Table 2, both in terms of level differences as well as the increases in the gap in more concentrated local markets.

Further decomposing system hospitals into those belonging to nationally small (Panel C) and large chains (more than 20 hospitals, Panel D) shows that within HRR both sets of for-profits have sizeably lower quality than not-for-profits within HRR (Column 3). Panel D, Column (4) shows that for hospitals that are part of large chains, the for-profit gap is completely crowded out in fully competitive markets (under full competition the interaction is 0), but increases strongly the more concentrated the local market is. Panel C shows that among smaller chain hospitals there remains a level difference between for- and not-for-profit hospitals and a more muted response to competitive pressure. Appendix Tables C5 and C6 show robustness to a large set of checks based on other measures derived from the HRRP.[16]

4 Discussion and conclusion

The differences in the behaviour of for- and not-for-profits in public service provision are well documented. Recent 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 of hospital chains, has been steadily increasing. As the ownership and market structures of the hospital market have changed it is important to understand how these are associated with differences in hospital quality.

To address this, we examine differences in quality between for- and not-for-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. This measure of quality is strongly correlated across diagnostic conditions within a hospital, suggesting a common hospital-wide quality, 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

[16] C5 shows consistency across all measures. The comparison with the raw-average penalty benchmark makes it clear that regional adjustment is necessary to meaningfully assess the association with market concentration. C6 Panel A shows that our main results are not affected by using OLS in the second step. Panel B drops government hospitals from the comparison. Panel C uses HHI-beds (which are constant within hospital) rather than HHI-discharges. Panel D adds an additional post-estimation shrinkage to the already shrunk BR estimates. Panel E shows there are similar but less pronounced tendencies among the top 25% of hospitals.

conditions, and patient satisfaction.

In the raw data, for-profit hospitals are around 7% more likely to be penalized than other hospitals. Our analysis shows that a large share of this difference remains even after extracting the medium-run quality component and adjusting for market (HRR) fixed effects. We find a substantial gradient in this negative for-profit gap in quality by market structure. We also find that chain status matters: both the quality gap and the interaction of for-profit status with the market structure are driven by hospitals that are part of large systems.

Our finding that competition in local markets is associated with a smaller for-profit gap chime with earlier findings in the literature. Our finding that chain hospitals are responsive to competition contrasts with Eliason et al. (2019), who find that competition does not affect chain hospital quality in the dialysis market: rather it is chain status per se that drives lower quality. In contrast to performance for dialysis outcomes, which are not used in the measures published in patient-oriented quality metrics such as Hospital Compare, the penalty rates we examine were heavily publicized and are among the main quality signals available to buyers and consumers of health care. It is thus more likely that hospitals compete for patients on the HRRP signals than on the dialysis outcomes and thus need to exert more effort when faced with rivals serving the same market. Moreover, since the measure of quality that we examine is associated with financial penalties, hospitals have greater incentives to respond. Poor performance on these metrics might be expected to be associated with loss of business (Chandra et al., 2016). An increase in competition will exacerbate this, and this may be particularly so for those hospitals at the lower end of the quality distribution Jones et al. (2017). We show that hospitals in large systems have lower quality if a lack of competition in their local market allows for flexibility in quality, but if operating in markets that are highly competitive the gap between for-profit and not-for-profit chain hospitals disappears.

An alternative explanation for our findings is that hospitals respond to publication of performance metrics by selecting healthy patients and do this more in competitive markets where they are more at risk of losing business. Chain hospitals may implement system-wide methods to do this and be more efficient at it than hospitals that are not part of a chain. However, the readmission rates published by CMS are for emergency admission and risk-adjusted to address this issue and we undertook further risk-adjustment in deriving our preferred quality measure. We, therefore, think that is it unlikely that our results are only driven by patient selection.

Finally, 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 change frequently, 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).

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.

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*.

Capps, C., D. W. Carlton, and G. David (2017). Antitrust treatment of nonprofits: Should hospitals receive special care? Technical report, NBER Working Paper No. 23131.

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.

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. A. Graves, and J. Gruber (2019). Evaluating measures of hospital quality. *Review of Economics and Statistics 101*(5), 841–852.

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.

Emmert, M. and M. Schlesinger (2017). Hospital quality reporting in the united states: does report card design and incorporation of patient narrative comments affect hospital choice? *Health services research 52*(3), 933–958.

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 (2019). Supervisors and performance management systems. *Journal of Political Economy Forthcoming*.

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.

Geweke, J., G. Gowrisankaran, and R. J. Town (2003). Bayesian inference for hospital quality in a selection model. *Econometrica 71*(4), 1215–1238.

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., 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.

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.

Hull, P. (2016). Estimating hospital quality with quasi-experimental data. *Job Market Paper*.

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.

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

Kunz, J. S. and C. Propper (2020). Association between covid-19 deaths and hospital quality in the usa. *mimeo*.

Kunz, J. S., K. E. Staub, and R. Winkelmann (2018). Predicting fixed effects in panel probit models. *HEDG Working Paper series 18/23*.

Kunz, J. S., K. E. Staub, and R. Winkelmann (2019). Predicting fixed effects in panel probit models. *Monash Business School 10/19*.

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.

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.

Paul, J. A., B. Quosigk, and L. MacDonald (2019). Does hospital status affect performance? *Nonprofit and Voluntary Sector Quarterly*.

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.

Saghafian, S. and W. J. Hopp (2020). Can public reporting cure healthcare? the role of quality transparency in improving patient–provider alignment. *Operations Research 68*(1), 71–92.

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.

New York Times, (2021, June). Buoyed by federal covid aid, big hospital chains buy up competitors.

Varkevisser, M., S. A. van der Geest, and F. T. Schut (2012). Do patients choose hospitals with high quality ratings? empirical evidence from the market for angioplasty in the netherlands. *Journal of Health Economics 31*(2), 371–378.

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

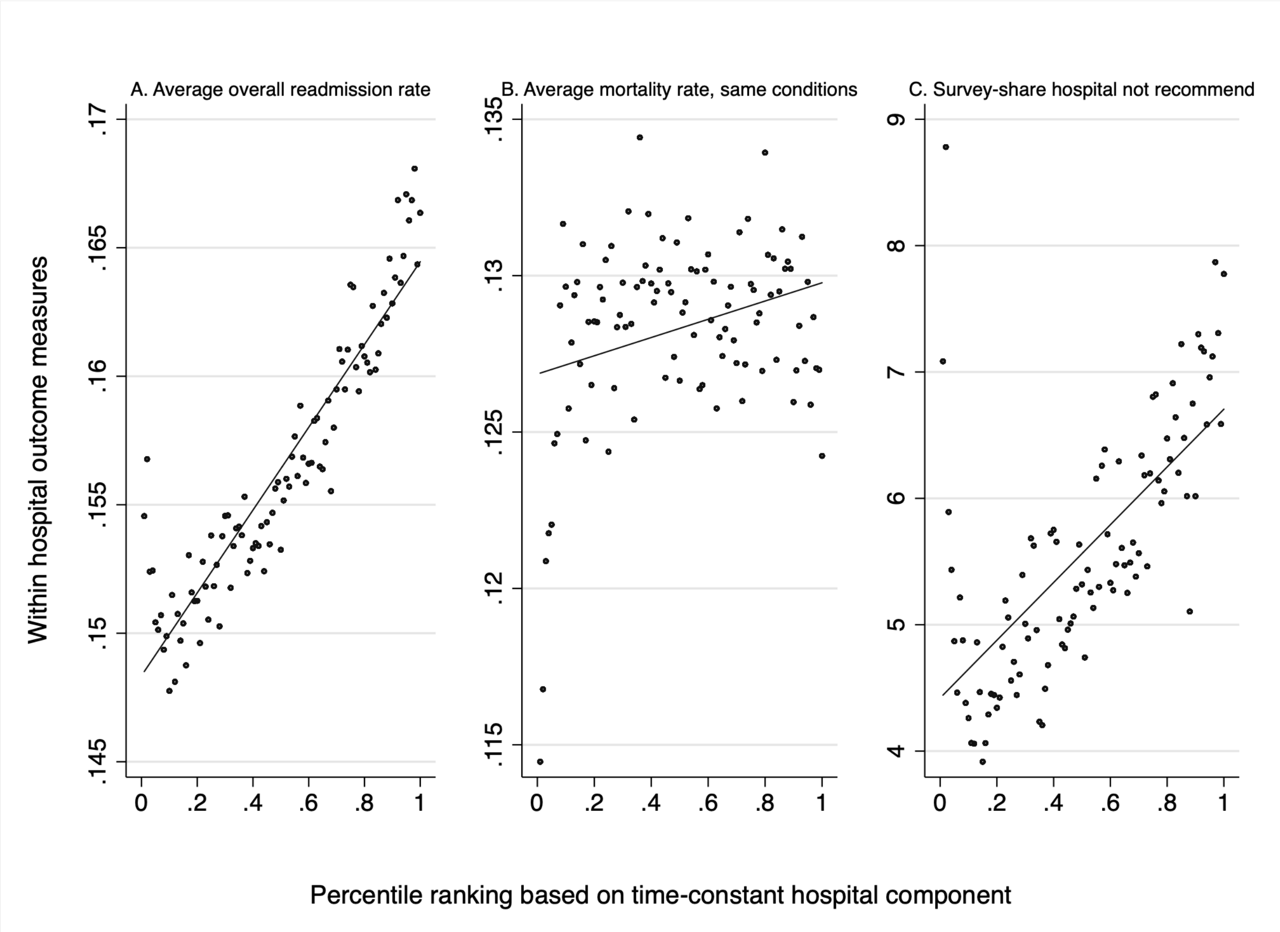

Figure 1: Within hospital readmission penalty propensity across diagnosis related groups, binned

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- the average mortality across time and the 3 diagnosis groups, and C- the survey based share of people answering they would not recommend this hospital.

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

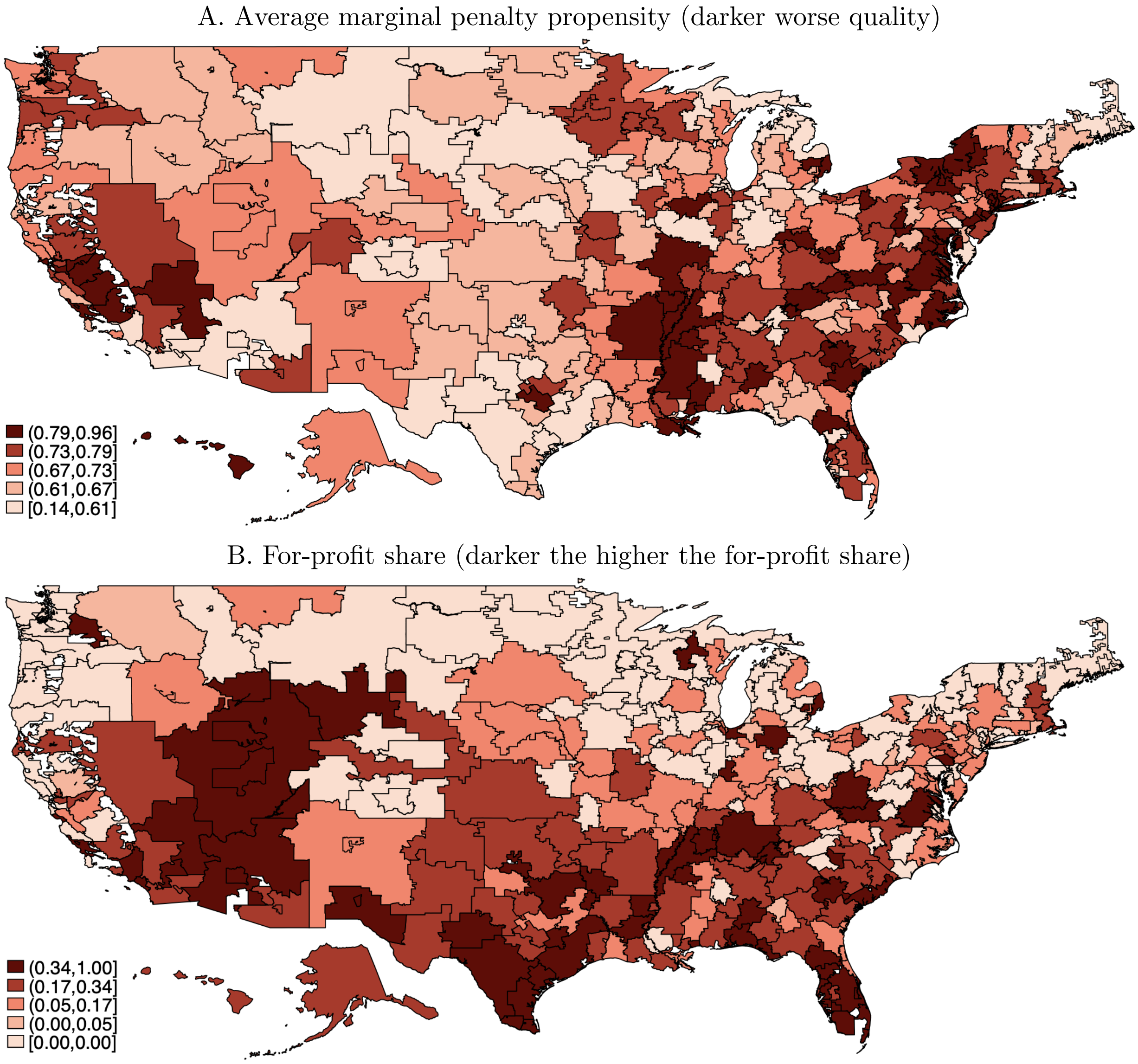

Figure 2: Average hospital adjusted readmission penalty propensities and for-profit share across hospital referral regions

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 6. Panel B depicts the share of For-profit hospitals in HRRs.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table 1: For-profit differential in readmission performance

	Unconditional Means			Fixed Effects: Covariate and Bias adjusted	
	For-Profit	Other	Difference	hospital-level	condition-level
	(1)	(2)	(3)	(4)	(5)
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
Time-varying characteristics				✓	✓
Time-constant characteristics				✓	✓
Diagnosis indicators					✓

Notes: Table presents various measures of readmission performance: Raw readmission rate - RR, excess readmission ratio - ERR (CMS-risk-adjusted based on age, gender and comorbidities), 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, system-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 raw readmission ratio and the ERR, we first run a OLS regression on time varying covariates (number of discharges, median household income in 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., 2019). [a]Note for the RR six hospitals did not have a RR but both ERR and Penalty status in the CMS data, the Observations refer to ERR as well as Penalty. Table C1 presents descriptive statistics by penalty status and C3 presents the full regression results for each condition separately, alongside analogous OLS regressions for the excess readmission ratio. 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 variables: Marginal hospitals penalty propensity $\Phi(\alpha_i^c)$							
			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*				0.051	0.052	0.042	0.055
				p=.026	p=.016	p=.001	p=.005
Panel B: System status							
System 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*				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	✓	✓	✓	✓	✓	✓	✓
County 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 system level rather than individual hospital-level, and (7) uses lagged HHI on system level in 2008. For-profit + Interaction, gives the interaction with the mean of the various HHI measures, and corresponding joint F-test p-value. Panel B: repeats the analysis for the system status instead of the For-profit indicator. Table C1 presents descriptive statistics by penalty status and C3 presents the full regression results for each condition separately, alongside analogous OLS regressions for the excess readmission ratio. 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 3: Marginal penalty propensity by For-profit, competition, chain status: Subsamples

Dependent variables: Marginal hospitals penalty propensity $\Phi(\alpha_i^c)$							
			Local hospital market				
	All	Constant Ownership	HRR FE	Competition: HHI based on -beds	-discharges	-chain	2008
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Sub-sample - non-system 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*				-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 - System hospitals							
For-profit hospital (Yes/No)	0.046	0.046	0.061	0.037	0.039	0.031	0.044
	(0.011)	(0.011)	(0.010)	(0.017)	(0.016)	(0.019)	(0.014)
× HHI-measure				0.230	0.192	0.159	0.145
				(0.120)	(0.099)	(0.079)	(0.078)
For-profit + Interaction*				0.061	0.062	0.053	0.065
				p=.013	p=.008	p=.003	p=.006
Observations	6,395	5,919	5,919	5,919	5,919	5,919	5,843
Panel C: Sub-sample - System 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
	(0.021)	(0.024)	(0.020)	(0.029)	(0.028)	(0.036)	(0.026)
× HHI-measure				0.094	0.113	0.120	0.121
				(0.250)	(0.217)	(0.200)	(0.209)
For-profit + Interaction*				0.061	0.062	0.057	0.067
				p= .527	p= .413	p= .349	p= .376
Observations	3,746	3,508	3,508	3,508	3,508	3,508	3,441
Panel D: Sub-sample - System 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
	(0.014)	(0.015)	(0.016)	(0.024)	(0.022)	(0.027)	(0.021)
× HHI-measure				0.304	0.268	0.236	0.239
				(0.176)	(0.145)	(0.104)	(0.101)
For-profit + Interaction*				0.035	0.038	0.023	0.042
				p=.050	p=.034	p=.007	p=.005
Observations	2,649	2,411	2,411	2,411	2,411	2,411	2,402
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects			✓	✓	✓	✓	✓

Notes: See Table 2 notes for details.
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 percent 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 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.

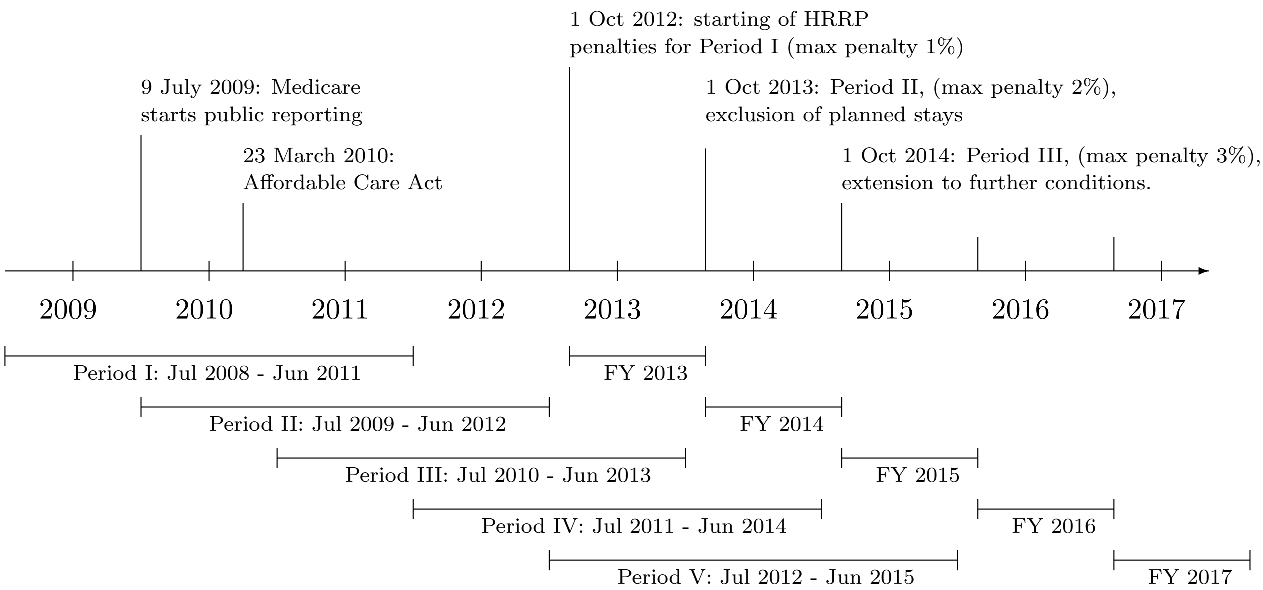

Figure A1: Event line and sample period of HRRP Policy
Source: Adapted from Figure 1 in Wasfy et al. (2017)

If a Medicare 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 a 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.[17] 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 percent in 2013, raised to two in 2014, and three in 2015, again complicating both the quality metrics over time.[18]

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

[18] 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 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).[19] 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 end up with the same penalty amount.

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 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 a large share of hospitals are penalized in every period or never. Since with only a short set of time periods it is unreasonable to assume that not observing a penalty implies that no shock could ever occur to push the hospital over the penalty cut-off and disregard any environmental information for such hospitals, this is a rather unappealing model implications for a measure of performance.

To address this, we estimate equation (2) using the bias-reduced fixed effects panel probit estimator as proposed by Kunz et al. (2018), which is based on a generalized linear model estimator first suggested by Kosmidis and Firth (2009). This estimator 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 (this value implies a marginal latent hospital quality corresponding to a 50-50 propensity to be penalized net of covariates). The estimator allows decomposing the time-variant from the time-invariant heterogeneity of *all* hospitals.

[19] 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.

C Additional results

C.1 Descriptive statistics and auxiliary regression results

First, we present simple descriptive statistics of our sample C1, then the detailed fixed effect regression C2, and C3 by condition and contrasting the estimates with the linear probability model.

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)
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) is the same as Col(1) in Table 1, for the time-invariant characteristics. Table C3 presents the full regression results for each condition separately, alongside analogous OLS regressions for the excess readmission ratio. 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 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.

Columns (2) and (3), correspond to Table 2 Panel C, Column (4) and (5), 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
eqrefbrfe, 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 propensity, confirming cross-sectional evidence (Gu et al., 2014).

Table C2: 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.

Table C3: DETAILED REGRESSION RESULTS OF PENALTY-RISK AND EXCESS READMISSION RATION, BY DIAGNOSIS CONDITION

Dependent variable: readmission penalty indicator	Penalty indicator (Yes/No) Probit-BR			Excess Readmission Ratio OLS		
	AMI	*HF*	*PN*	*AMI*	*HF*	*PN*
	(1)	(2)	(3)	(4)	(5)	(6)
Number of discharges	0.0008	0.0014	0.0002	0.0000	0.0001	0.0000
	(0.0005)	(0.0003)	(0.0002)	(0.0000)	(0.0000)	(0.0000)
Number of discharges, other ER conditions	0.0004	0.0000	0.0007	0.0000	0.0000	0.0000
	(0.0002)	(0.0002)	(0.0002)	(0.0000)	(0.0000)	(0.0000)
Number of beds	−0.0003	−0.0004	−0.0006	−0.0000	−0.0000	−0.0000
	(0.0008)	(0.0008)	(0.0007)	(0.0000)	(0.0000)	(0.0000)
Discarges ACSCs per 1'000 enrolles, in HRR	0.0302	0.0268	0.0205	0.0026	0.0016	0.0012
	(0.0093)	(0.0079)	(0.0073)	(0.0005)	(0.0004)	(0.0004)
Hospital opening, in HRR	−0.0215	−0.0010	0.0384	−0.0026	0.0007	0.0008
	(0.0456)	(0.0395)	(0.0388)	(0.0019)	(0.0016)	(0.0017)
Hospital closing, Hin HRRrr	−0.0683	−0.0331	−0.0428	−0.0036	−0.0012	−0.0028
	(0.0379)	(0.0321)	(0.0326)	(0.0016)	(0.0014)	(0.0014)
Percent living in poverty, in county	−0.0066	0.0008	−0.0055	−0.0006	−0.0000	−0.0008
	(0.0146)	(0.0110)	(0.0091)	(0.0006)	(0.0005)	(0.0004)
Household median income in 10'000$, in county	−0.1466	0.0574	−0.0200	−0.0054	−0.0032	−0.0032
	(0.1055)	(0.0793)	(0.0697)	(0.0051)	(0.0036)	(0.0034)
Total population in 100'000, in county	0.0187	0.0594	0.0313	−0.0010	0.0047	0.0034
	(0.0702)	(0.0665)	(0.0601)	(0.0033)	(0.0030)	(0.0030)
Percent unemployed, in county	−0.0267	−0.0130	−0.0069	−0.0019	−0.0011	−0.0011
	(0.0276)	(0.0221)	(0.0203)	(0.0013)	(0.0010)	(0.0010)
Number of observations	10,972	14,951	15,172	10,972	14,951	15,172
Number of hospitals	2,397	3,138	3,178	2,397	3,138	3,178
Share of concordant observations, 0	24.2	24.5	24.9			
Share of concordant observations, 1	23.4	23.4	21.2			
$\overline{\phi(x'\beta)}$	0.19	0.09	0.30			
Hospital fixed effects	✓	✓	✓	✓	✓	✓
Time fixed effects	✓	✓	✓	✓	✓	✓

Notes: See Table notes in 1. Here, regressions are separately run for each condition in Column (1)-(3), and as comparison (4)-(6) present OLS-fixed effect regressions using the excess readmission ratio as outcome, recall that penalty is defined in this ratio is larger than 1, thus large differences between the two sets of regressions might indicate gaming.
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.

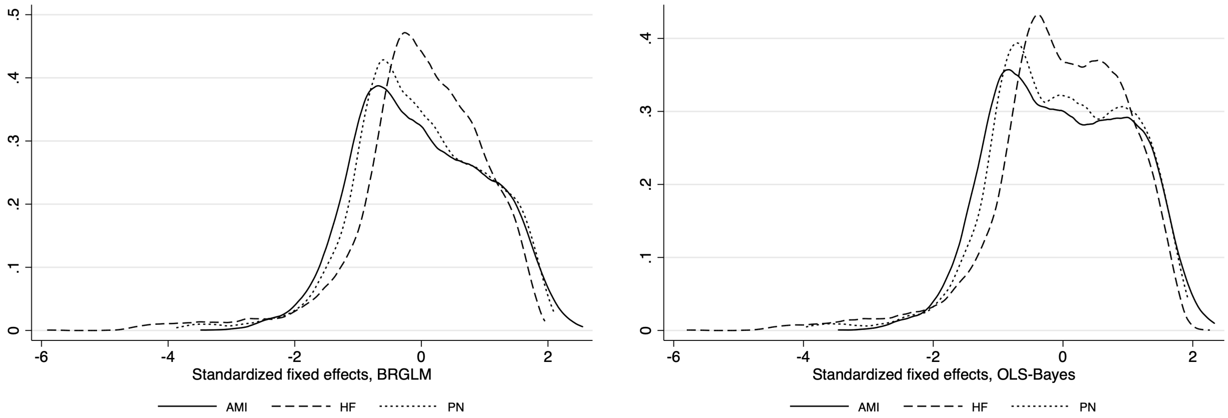

Figure C1: Density plot of separately estimated fixed effects, left Penalty-BR, right RR-OLS-Bayes

Note: Figure plots density plots of hospital fixed effects across diagnosis conditions (AMI, HF, PN). We use the estimated fixed effects from eq.(2), ie. $\Phi(\hat{\alpha}_i^c)$ 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.[20]

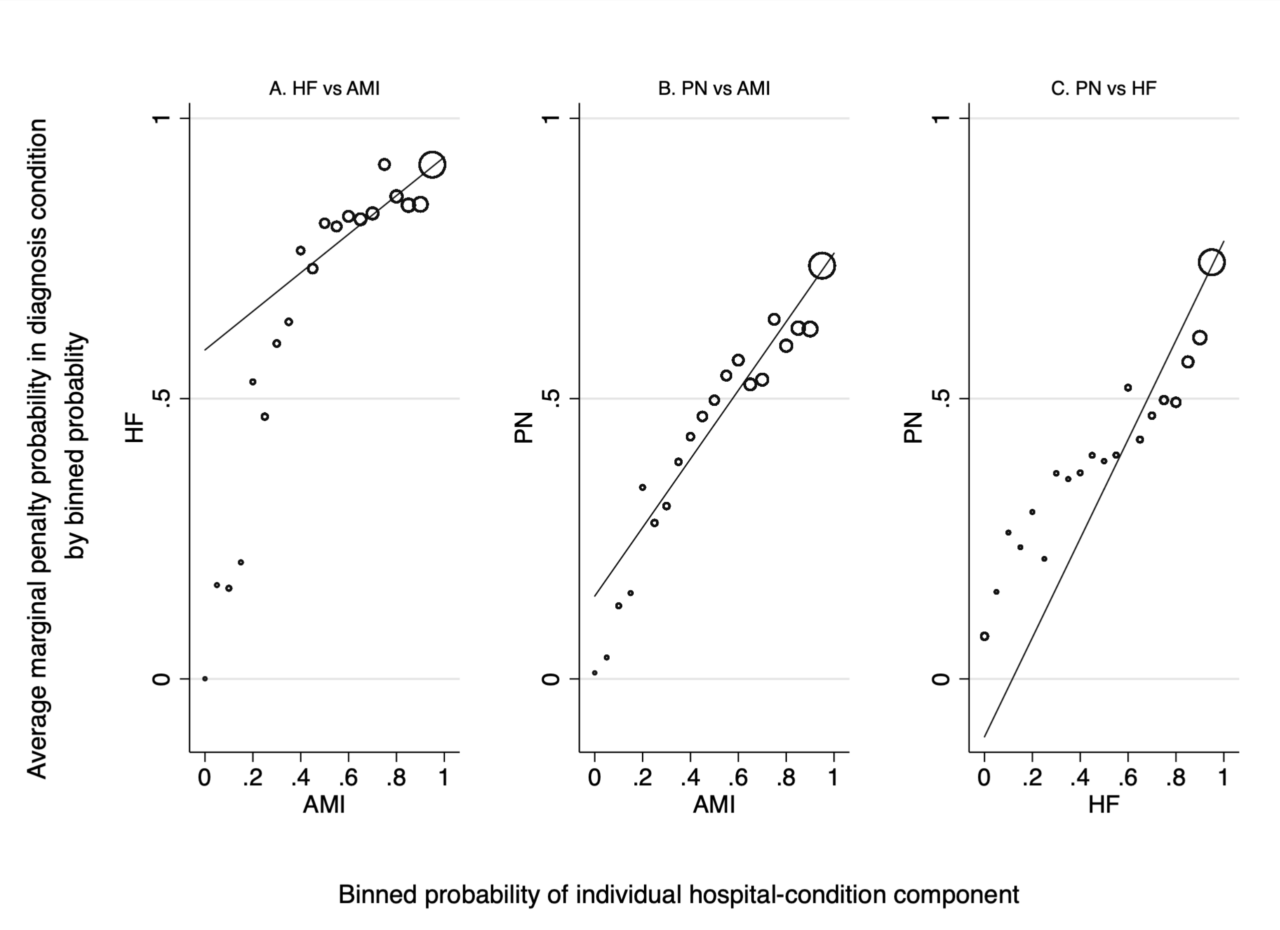

Figure C2: WITHIN HOSPITAL READMISSION PENALTY PROPENSITY ACROSS DIAGNOSIS RELATED GROUPS, BINNED

Note: Figure plots correlation of hospital fixed effects across diagnosis conditions (AMI, HF, PN). We use the estimated fixed effects from eq.(2), ie. $\Phi(\hat{\alpha}_i^c)$ 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.

[20]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 (2020).

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 actually 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.

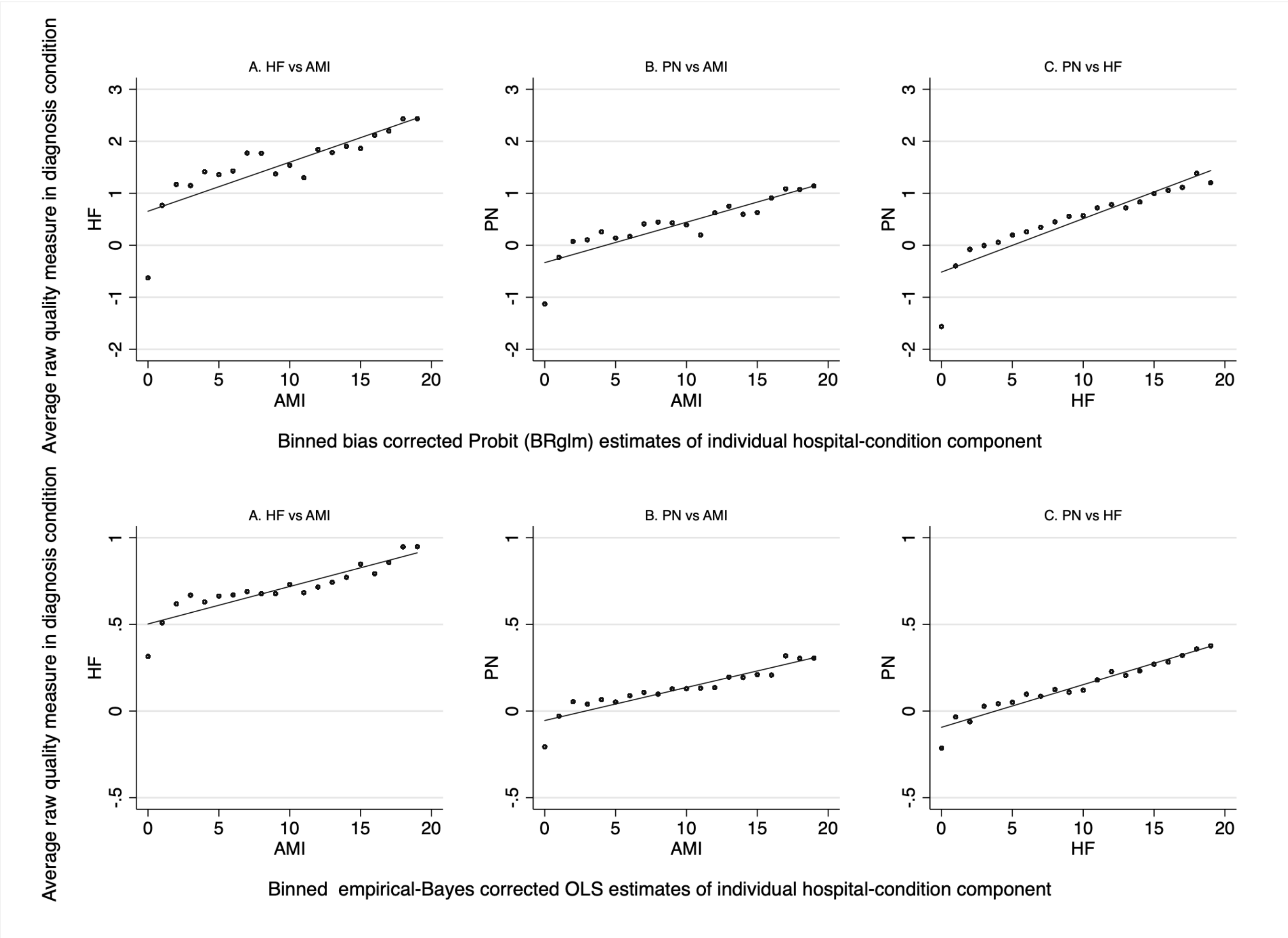

Figure C3: Robustness to Figure C2, equal sized bins of raw alpha and comparison with OLS fixed effects

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: Association between hospital quality and For-profit status, market concentration, and system status

Dependent variables: Various measures of hospitals time-invariant penalty heterogeneity (BR fixed effects)

				Local hospital market			
			Constant	HRR		Part of a system	
	Pooled	All	Ownership	FE	HHI	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Marginal hospitals penalty propensity $\Phi(\alpha_i)$							
For-profit hospital (Yes/No)	0.023	0.030	0.031	0.052	0.035	-0.056	0.037
	(0.010)	(0.010)	(0.011)	(0.009)	(0.014)	(0.044)	(0.016)
× HHI-discharges					0.129	0.106	0.195
					(0.074)	(0.221)	(0.092)
For-profit + Interaction*					0.054	-0.040	0.066
					p =.011	p =.788	p =.004
Panel B: Hospital percentile ranking of poor quality							
For-profit hospital (Yes/No)	0.043	0.036	0.038	0.061	0.046	-0.064	0.050
	(0.013)	(0.011)	(0.011)	(0.010)	(0.015)	(0.046)	(0.017)
× HHI-discharges					0.111	0.077	0.186
					(0.079)	(0.235)	(0.099)
For-profit + Interaction*					0.063	-0.053	0.078
					p =.023	p =.951	p =.007
Panel C: Indicator whether hospital is in the top 10% of marginal hospital quality							
For-profit hospital (Yes/No)	0.005	-0.010	-0.013	-0.036	0.001	0.053	0.003
	(0.013)	(0.011)	(0.012)	(0.011)	(0.018)	(0.063)	(0.019)
× HHI-discharges					-0.284	-0.257	-0.338
					(0.093)	(0.270)	(0.108)
For-profit + Interaction*					-0.041	0.015	-0.047
					p =.000	p =.347	p =.000
Observations	3,197	8,713	8,072	8,072	8,072	2,153	5,919
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects				✓	✓	✓	✓

Notes: Table presents OLS coefficients. Column (1) displays estimates based on pooled fixed effects and robust standard errors from Table 1(Col. 5). Columns (2)-(7) show resulting using the full set of fixed effects and cluster robust standard errors at the hospital level, corresponding to Table 1(Col. 6). All regressions contain hospital characteristics (teaching status 3-categories, size based on number of beds 3-categories, whether it belongs to a system, and whether its located in an urban area) as well as county-level demographics from the Census (share of people older 65, share of non-hispanic whites, hispanics, blacks, share of population college educated, high school educated), and condition indicators. Column (3) restricts the estimation sample to hospitals that did not change ownership in the sample period, Col. (4) adds HRR fixed effects, Col. (5) the interaction with HHI, and Cols. (6) and (7) split the sample into hospitals that are not part of a system and those that are. *For-profit + Interaction shows the for-profit gap evaluated at the average HHI (0.149). The p-value indicates whether the sum of both coefficients is significantly different from 0.

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

Table C5: Association of different measures penalty-risk and selected covariates

	Pooled	All	Constant Ownership	HRR FE	HHI	Part of a system: No	Part of a system: Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Raw mean RR							
For-profit hospital (Yes/No)	0.002	0.002	0.002	0.003	0.002	-0.002	0.002
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)	(0.001)
× HHI-discharges					0.006	0.018	0.010
					(0.004)	(0.015)	(0.004)
For-profit + Interaction*					0.003	0.001	0.003
					p =.016	p =.220	p =.004
Panel B: Raw mean ERR, risk-adjusted							
For-profit hospital (Yes/No)	0.011	0.011	0.011	0.014	0.011	-0.010	0.009
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.010)	(0.004)
× HHI-discharges					0.025	0.058	0.049
					(0.018)	(0.062)	(0.023)
For-profit + Interaction*					0.014	-0.002	0.016
					p =.027	p =.387	p =.004
Panel C: Ebayes OLS ERR							
For-profit hospital (Yes/No)	0.008	0.009	0.009	0.015	0.008	-0.024	0.009
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.010)	(0.004)
× HHI-discharges					0.050	0.126	0.070
					(0.019)	(0.064)	(0.023)
For-profit + Interaction*					0.016	-0.005	0.019
					p =.000	p =.071	p =.000
Panel D: Raw mean penalty							
For-profit hospital (Yes/No)	0.056	0.065	0.065	0.076	0.074	0.034	0.064
	(0.017)	(0.013)	(0.014)	(0.013)	(0.019)	(0.060)	(0.022)
× HHI-discharges					0.020	-0.322	0.165
					(0.101)	(0.316)	(0.131)
For-profit + Interaction*					0.076	-0.014	0.089
					p =.290	p =.297	p =.047
Panel E: Ebayes OLS Penalty							
For-profit hospital (Yes/No)	0.050	0.054	0.057	0.079	0.059	-0.051	0.060
	(0.011)	(0.012)	(0.012)	(0.012)	(0.017)	(0.054)	(0.019)
× HHI-discharges					0.155	0.110	0.258
					(0.098)	(0.314)	(0.124)
For-profit + Interaction*					0.082	-0.035	0.098
					p =.013	p =.833	p =.004
Observations	3,197	8,713	8,072	8,072	8,072	2,153	5,919
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects				✓	✓	✓	✓

Notes: See Table notes in C4, Panel A - shows average raw readmission rate, B - raw mean of the excess readmission ratios (accounts for basic risk-adjustment), C - OLS fixed effects on ERR with post-estimation E.Bayes correction (thus adds local area risk-adjustment), D - uses the raw mean of the penalty status (policy indicator), E - again compares to the OLS with post-shrinkage.
Source: CMS 2011-2015, Dartmouth Atlas of Health Care, Census 2010, ACS, own calculations.

Table C6: Association of different measures penalty-risk and selected covariates

	Pooled	All	Constant Ownership	HRR FE	HHI	Part of a system: No	Part of a system: Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Fractional response regressions, probit: $\Phi(\alpha_i)$							
For-profit hospital (Yes/No)	0.082	0.091	0.095	0.169	0.121	-0.167	0.132
	(0.035)	(0.022)	(0.023)	(0.023)	(0.035)	(0.098)	(0.040)
× HHI-discharges					0.369	0.257	0.554
					(0.190)	(0.512)	(0.235)
For-profit + Interaction*					0.176	-0.129	0.214
					p =.003	p =.840	p =.001
Panel B: Without government hospitals							
For-profit hospital (Yes/No)	0.030	0.035	0.036	0.060	0.035	-0.059	0.041
	(0.011)	(0.010)	(0.011)	(0.010)	(0.015)	(0.042)	(0.017)
× HHI-discharges					0.205	0.334	0.181
					(0.086)	(0.194)	(0.104)
For-profit + Interaction*					0.066	-0.009	0.068
					p =.001	p =.102	p =.016
Observations$^+$	2,607	7,229	6,814	6,814	6,814	1,443	5,371
Panel C: Additional post-shrinkage empirical Bayes							
For-profit hospital (Yes/No)	0.129	0.131	0.139	0.200	0.146	-0.157	0.152
	(0.030)	(0.029)	(0.031)	(0.031)	(0.043)	(0.127)	(0.048)
× HHI-discharges					0.413	0.116	0.686
					(0.242)	(0.689)	(0.306)
For-profit + Interaction*					0.208	-0.140	0.254
					p =.009	p =.946	p =.002
Panel D: Indicator whether hospital is in top 25%							
For-profit hospital (Yes/No)	-0.043	-0.022	-0.023	-0.055	-0.024	0.174	-0.045
	(0.019)	(0.016)	(0.016)	(0.015)	(0.022)	(0.072)	(0.025)
× HHI-discharges					-0.242	-0.445	-0.184
					(0.124)	(0.364)	(0.151)
For-profit + Interaction*					-0.060	0.108	-0.073
					p =.015	p =.389	p =.085
Panel E: Too few discharges - robustness whether gaming via up/down-coding or selective admission							
For-profit hospital (Yes/No)		0.007	0.002	0.017	0.012	0.012	0.002
		(0.022)	(0.024)	(0.024)	(0.038)	(0.111)	(0.045)
× HHI-discharges					0.038	0.684	-0.056
					(0.211)	(0.607)	(0.274)
For-profit + Interaction*					0.017	0.113	-0.007
					p =.786	p =.185	p =.821
Observations$^+$		9,564	8,814	8,814	8,814	2,467	6,347
Observations	3,197	8,713	8,072	8,072	8,072	2,153	5,919
Hospitals characteristics	✓	✓	✓	✓	✓	✓	✓
County demographics	✓	✓	✓	✓	✓	✓	✓
HRR fixed effects				✓	✓	✓	✓

Notes: See notes in Table 2. Panel A fractional probit model, B - drops government run hospitals, C - adds an additional post-shrinkage to the BR probit estimates, and D - uses as dependent variable an indicator whether the hospital-condition pair is in the top 25%, and E - shows the results for too few discharges that does not vary with ownership or competition. If Observations$^+$ are not indicated in the Panel they are the same as in the main regressions and indicated at the bottom.

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`[21] to the crosswalk between 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[22] 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`,[23] 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 county.

Hospital Referral Region information

We use zip code crosswalks:[24] `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 where less. Note, that we can not distinguish, whether these are actually 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:[25] `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:[26]

- `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`

[21] downloaded from http://www.nber.org/ssa-fips-state-county-crosswalk/ (accessed 26.03.17).
[22] downloaded from https://www.ers.usda.gov/data-products/atlas-of-rural-and-small-town-america/download-the-data/ (accessed 26.03.17).
[23] downloaded from https://www.census.gov/data-tools/demo/saipe/saipe.html (accessed 26.03.17).
[24] downloaded from http://www.dartmouthatlas.org/tools/downloads.aspx?tab=39 (accessed 26.03.17).
[25] downloaded from http://www.dartmouthatlas.org/tools/downloads.aspx?tab=41 (accessed 26.03.17).
[26] 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 for 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 hospital, for the very few minor changes we set the to the first observed state, to make them time-consistent.

Next, we use Hospital Compare data archive:[27]

- 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 a robustness, see appendix Table C6, we use this as dependent variable. Finally, we summarise our definitions of the key variables in the following table.

[27] 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