Gaming of Quality Inspection: Evidence from High Frequency Staffing Data

Sean S. Huang and Richard A. Hirth[‡]

Very Preliminary: Please Do not Circulate or Cite

October, 2015

ABSTRACT

We use a unique high frequency daily staffing dataset to study whether and to what extent nursing homes manipulate the results of annual health inspections. We find nursing homes can precisely anticipate the inspection and inflate staffing during inspection days by 6.3% on average and as much as 23.6% for specific groups. We also find that gaming behavior is amplified after the 5-star quality rating system, with inspection day staffing changes rising from 5.0% to 7.9%. Interestingly, the comparison of staffing on non-inspection days before and after the implementation of quality rating suggests that the rating system also has meaningful positive effects on staffing level. Overall, nursing homes can and do inflate staffing during inspection days. Our results suggest amplified gaming behaviors and meaningful increase in staffing level after the implementation of quality rating.

[‡] Sean Huang, Georgetown University (sean.huang@georgetown.edu). Richard Hirth is at University of Michigan (rhirth@umich.edu).

We thank Ian McCarthy, Ethan Lieber, and seminar participants at 2015 iHEA World Congress, AHRQ's CFACT, and University of Washington for valuable comments and suggestions. All errors remain our own.

1. Introduction

"The inspection period is so crucial that in 2010, an administrator at a home on Long Island described it as "our Super Bowl" and explained that staff levels would drop once the inspection was completed." *-- The New York Times, August 2014*

Quality inspection has been used extensively to ensure quality standards are met. The results of these inspections are often disseminated publicly to encourage consumers to purchase on the basis of quality. For inspection to reveal true quality, one of two assumptions, the randomization of timing and static quality proxies, has to hold. However, despite wide adoption of quality inspection across different economic sectors, little is known about how accurately inspection reflects the underlying quality. In this paper, using a novel high frequency (daily) nursing home staffing dataset (day, evening, and night shifts), we examine whether providers can anticipate and manipulate annual inspection by temporarily altering staffing. More importantly, the dataset spans the period before and after the implementation of a more effective quality-rating and reporting system. We are therefore able to examine if the public dissemination of quality information may unintentionally exacerbate the manipulation of quality inspections.

These high frequency staffing data also enable us to address an important policy debate. Staffing levels have substantially improved after the implementation of quality rating system since 2009 (Williams et al., 2013). In fact, the magnitude of improvement is so substantial that concerns have been expressed about its plausibility and potential artificial inflation (New York Times, 2014). Comparing daily staffing levels distant from inspection days helps to mitigate the concerns of the inflation of staffing. A finding of significant and substantial increase of staffing on the non-inspection days can address such concerns and provide strong evidence that quality rating indeed achieve meaningful improvement of nursing staffing.

The empirical strategy relies on a high frequency staffing dataset that includes day-shift-level staffing for all Oklahoma nursing homes from 2007 to 2012. The primary result is based around 240,000 observations. Matching daily staffing data with annual inspection days, we can precisely observe abnormal staffing patterns before, during and after health inspections. On average, we find that staffing increases by 6.3% during an inspection. We also find that this effect is exacerbated under the recently implemented five-star quality rating among those with worse patient outcomes by at least 58% (from 5.0% to 7.9%). We also identify 20% of facilities that substantially inflate staffing by an average of 23.6%. Contrasted to the public perception, we find that nonprofit nursing homes have more intense gaming behaviors compared to the for-profit nursing homes. Also, independent nursing homes actually inflate their staffing more aggressively than chain-affiliated homes. These heterogeneous effects by organizational characteristics provide an alternative perspective to regulate and monitor health care firms.

Overall, our results call for more extensive reviews of the validity of inspection results, in both health care and other sectors. Last but not least, the substantial staffing increase (9.2%) during non-inspection days after the implementation of the five-star system suggests that quality rating may indeed achieve meaningful quality enhancement, at least on the staffing domain.

2. Nursing Home Staffing and Concern of Gaming

Nurses and nurse aides are the major workforce in the nursing homes. Because of the asymmetric quality information to consumers and low salaries of nursing home workers, policy makers have concerned that inadequate staffing level in nursing homes may lead to suboptimal quality (GAO, 2001). The decentralized reimbursement and regulation of nursing homes cause substantial variations of nursing home staffing level (Harrington 2010). Both state and federal

governments have adopted policies to address this challenge, including various mandates of minimum staffing requirements, Medicaid's wage pass-through payments, and public disclosure of staffing information. Utilizing these policy and other exogenous changes, previous studies provide evidence of the casual relationship between higher (lower) staffing level and better (worse) quality. For example, Cawley et al. (2006) uses changes in minimum wage laws to instrumental for nursing homes' staffing level and find that lower staffing leads to worse patient outcomes; Lin (2014) uses the variation and change of state minimum staffing laws to instrument nursing homes' staffing level, and find that higher Registered Nurse (RN) staffing ratio has positive impacts on quality. Feng et al. (2010) and Foster and Lee (2014) study Medicaid wage pass-through payments and find consequentially increase on Certified Nurse Aides (CAN) staffing and improvement on quality. Despite these positive findings, because the staffing ratios are self-reported by nursing homes, the reliability of the reported staffing level remains a major concern to consumers. This concern continues grow substantially after the implementation of a more user-friendly rating system.

Recent studies show nursing home quality responds to quality rating based on inspection results. However, particularly on the staffing domain, the magnitude of improvement is sometimes so large that concerns have been expressed about its plausibility and potential artificial inflation (New York Times, 2014). For example, 37.8% of nursing homes receive 4- or 5-star rating on staffing domain when the rating system was launched in January 2009. By October 2013, about 53% of nursing homes receive at least 4 star ratings. On the other hand, the percentage of nursing homes receive 1 star reduces from 22.9% to 12.1% during the same period (CMS 2014). The substantial improvement in staffing and the increasingly dissemination of quality rating, motive us to closely study nursing homes' staffing patterns.

3. Hypotheses

We examine three main hypotheses in this paper. First, we examine the dynamics of staffing patterns on days when facilities are inspected relative to non-inspection days. If nursing facilities are willing and able to game quality inspection, we will see significantly higher staffing on inspection days than on non-inspection days. Because inspections are done every 9 to 15 months, nursing home administrators theoretically can only approximately anticipate the next inspection day, or even the quarter in which it occurs. In fact, it is written specifically in the state health statutes that inspection should be unannounced and disclosure to any unauthorized person may face administrative and legal consequences[1]. However, if the law is not strictly enforced, with potential information sharing between inspectors and facilities, administrators may still be able to anticipate inspections. If nursing home administrators can anticipate inspections and inflate their staffing temporarily, we should see staffing level spike abnormally during the inspection.

Second, we examine whether the intensity of gaming is amplified under the recently implemented five-star quality rating. Recent studies on prices (Hirth and Huang, 2015) and quality (Konetzka et al., 2015) both show that consumers and providers are responsive to the quality rating. When inspection results are more widely disseminated, higher marginal benefits (costs) of better (worse) reported quality may further encourage facilities to inflate staffing.

Third, to address the growing public concerns that the substantial increase in staffing may be mainly due to gaming of rating system rather than meaningful improvement, we compare the staffing ratio on the non-inspection days before and after the implementation of quality rating. Unless nursing homes consistently inflate staffing data for everyday through the year, and

[1] Oklahoma Statutes – Title 63. Public Health and Safety (http://law.justia.com/codes/oklahoma/2014/title-63/section-63-1-829)

substantially intensify the inflation after the implementation of the rating system, a finding of significant increase of staffing level on non-inspection days during the post-rating period will favor meaningful quality enhancement.

4. High Frequency Data and Empirical Strategy

Previous studies about nurse staffing mostly rely on the annual administrative data. For example, in the nursing home literature, most studies extracted staffing data from the annual Online Survey, Certification, and Reporting (OSCAR) database. In OSCAR dataset, staffing information is self-reported during the annual health inspection. It is unclear whether a one-time snapshot represents the underlying staffing levels across the entire year (Harrington et al., 2000). The availability of daily staffing information enables us to take a finer look at staffing patterns, particularly any abnormalities on or near health inspections. High frequency data has been used in the finance and economics literature to study market microstructure (Engle and Russell, 2004), information flows, and consumer demand (Levin et al., 2015). To our knowledge, our study is the first to utilize high frequency staffing data in studying nursing homes[2]. We obtained a daily nurse staffing dataset from the Oklahoma State Department of Health[3]. All Oklahoma certified nursing homes are required to submit staffing reports monthly to the health agency to ensure the compliance of the state minimum staffing requirement. Figure 1. shows a sample of the staffing report. These staffing reports contain daily staffing records of all nursing facilities from 2002 to 2014, before and after the implementation of five-star quality rating in December 2008. Our

[2] Evans and Kim (2006) and Dobkin (2003) use the fluctuations of hospital admissions between weekdays and weekends to examine the causal relationship between staffing and hospital quality. In both studies, the actual daily staffing levels are not observable and rely on the assumption that staffing levels are predetermined. .

primary analysis excludes transition years (2008 and 2009) and includes the staffing ratios of day, evening, and night shifts in 2007, 2010, 2011, and 2012 data[4]. We merge the daily staffing data with inspection dates extracted from the Center of Medicare and Medicaid Services (CMS). Our sample has more than 240,000 observations. We use the recorded health survey dates as the inspection days (633 inspection dates). We create three staffing ratios by dividing direct care staffing hours by number of patients during day, evening, and night shifts. Our main analysis is based on a facility-fixed effect model. The baseline econometric model can be described as follows:

$$S_{i,t} = \alpha + \beta_{ins} * Inspection_{i,t} + \beta_p * post + \beta_r * Inspection_{i,t} * post + T_t + \varepsilon_{i,t}$$

where the main dependent variables $S_{i,t}$ are daily staffing ratios of day, evening, and night shifts. The staffing ratio is defined as the direct care staff hours divided by peak in-house resident counts during each shift. We also create a shift-weighted average staffing ratio to measure the average daily staffing level. The key independent variable $Inspection_{i,t}$ is a binary variable indicating inspection days. *Post* is a dummy variable indicating the period after the launch of five-star rating system in December 2008. The coefficient of the interaction term, β_γ, represents the difference in the effects of inspection days on staffing ratios, before and after the high dimensional quality information is summarized into quality star ratings. If nursing homes can anticipate the inspection dates and inflate their staffing during inspections, $\beta_{\text{ins},}$, the coefficient of *Inspection,* will be positive and significantly different than zero. Second, if ,the summarization of the higher dimensional quality data into publicly reported star rankings

[4] Data before January 2007 and after October 2012 are kept in a different format and difficult to use. We are processing these data and will include them in the final analysis to provide more robust analysis. For the years included in the study, 30%-50% of facilities submitted paper-based reports, and we are in the process of transcribing these data. Due to this issue, our preliminary results are subject to selection bias and should be taken with caution.

enhanced facilities' incentives to game health inspection, β_γ, the coefficient of the interaction term would be positive and significantly different than zero (that is, nursing homes would raise their staffing more aggressively during inspections after the star ranking became available). To account for any seasonality, for all regressions, we also control for month, calendar day, and weekday fixed effects. T_t represent these time-fixed effects. The facility-fixed effect accounts for all time-invariant firm-level and market-level variables. We exclude hospital-based nursing homes because they have very different patient population and operating environments. In our preliminary analysis, we do not include patient-level characteristics but will include in our following work.

5. Results and Discussion:

As shown in Figure 2, nursing homes are able to adjust their staffing levels at the beginning of inspection, maintain higher staffing for 3 days (6-8% higher than the 30 days median), and rapidly decrease to their normal level thereafter. We find very strong and significant evidence that facilities are able to increase staffing level around the inspection dates by 6.3% (coefficient=0.11/median=1.58, Table 2). We also find that the effect is exacerbated by 2.9% (coefficient=0.045/median=1.58, Table 2) after the five-star quality rating. Interestingly, such exacerbation mostly occurs on the day shift (column (2)) but not on the evening and night shifts (column (3) and (4)). This heterogeneous effect among shifts is likely because staffing during the daily shifts is more observable to inspectors and hiring of temporary staffers or adding on-call staffs is easier and perhaps cheaper.

To measure the heterogeneous effects among facilities, we identify a subgroup of approximately 20% of the facilities that have an abnormal spike of staffing on inspection days (> 1 std) and their staffing on inspection days, rising by 23.6% (coefficient=0.373/median=1.58, Table 3).

Under five-star ratings, facilities with worse patient outcomes may be more likely to manipulate staffing to improve their overall rating. We split facilities with above or below median patient outcomes. We select four representative quality measures, including the percentage of patients with pressure sores, contractures, unexpected weight change, and worsen mental status. These four quality measures are all included in the calculation of CMS' star quality rating. Because patient outcomes are measured based on quarterly reporting that is independent of the annual health inspection date (when staffing level is reported), we argue that the reported patient outcomes are less subject to the timing of health inspection as well as the temporary inflation of staffing during the inspection. We find that quality rating only exacerbates gaming behaviors for the facilities with worse patient outcomes (Table 4, column (2) (4) (6) and (8)).

We further explore the heterogeneous effects according to the ownership status (for-profit vs. nonprofit) and chain-affiliation. Studying these heterogeneous effects provides an important policy implication to identify the types of nursing facilities that are more likely to game, and the regulator may allocate its monitoring resources more efficiently. In our sample, about 89% of nursing homes are for-profit and only 9% are nonprofits. 35% of the nursing homes are chain-affiliated. We stratify the sample by the for-profit and chain-affiliation status and the results are reported in Table 5 and Table 6. Interestingly, we find that both for-profit and nonprofit nursing homes have similar gaming patterns. The magnitudes of gaming are even large for the nonprofit

homes (0.155, Panel B in Table 5) than the for-profits (0.102, Panel A in Table 5). In addition to the profit-maximizing and altruistic objectives argument, an alternative explanation is that nursing facilities use for-profit and nonprofit status signal customers and stack up the low-end and high-end markets respectively. If nonprofits target the high quality segment and their consumers are more responsive to revealed quality, nonprofits may actually have similar if not stronger incentives to game.

Contrasted to the public perception (Kitchener et al., 2008; Harrington, 2011), the results from Table 6 suggest that the magnitudes of gaming are actually larger for the independent facilities than the chain-affiliated facilities (0.114 vs. 0.097). To explain the seeming counter-intuitive finding, we link our results to the literature on legal risks and asset shielding. Brickley et al. (2014) finds that when facing higher malpractice litigation risks, large nursing home chains are more likely to divest their ownership and exit from markets with higher litigation risks. In the same spirit, chain-affiliated facilities may bear higher legal risks of staff inflation and participate less aggressively. In addition to legal risks, large health chains also often face stricter scrutiny from the media, for example, The Washington Post on Manor Care (2007), AseraCare (2013) and The Wall Street Journal on Kindred Healthcare (2015).

5 A. Comparison of Inspection and Non-inspection Days:

One of the major concerns of current rating system is that facilities *only* increase their staffing levels during the inspection, to receive better quality rating without meaningful improvement during most of the year. To address such concerns, we also compare the staffing levels on the non-inspection days, before and after the implementation of quality rating. Non-inspection days span most of the days of a year, and more importantly, nursing facilities have

weaker incentives to inflate staffing during these days. As shown in Tables 2, 3, and 4, we find the coefficients of the *post* variable are positive and statistically significant. The results suggest that there was significantly and substantially higher staffing after the implementation of the five-star system (not necessarily casual).

5 B. Mechanisms:

We further explore the plausible mechanisms enabling nursing homes to inflate their staffing in such a sharp and rapid fashion. For such a short-term staffing adjustment, one plausible explanation is that the nursing homes reduce working hours prior and after the inspection, and reallocate these work hours to inspection days. The reallocation of resources to game quality inspection is not new. For example, Lu (2012) finds that nursing homes reallocate inputs to improve quality only on the domains that are included in public reporting and quality on other domains deteriorate. If nursing homes do temporarily reallocate staffing around inspection days, not only would reported staffing be inflated, but patients' health and safety could also be jeopardized just before and after the inspection periods. To test this hypothesis, we create the indicator variables, *before* and *after*, which take values of 1 if the days fall into 7 days before and after the inspection periods. If the coefficients are negative and significantly different than zero, it would suggest that the staffing levels are lower than the annual average and raise a flag that nursing homes reduce the staffing immediately before and after inspection, in order to boost their staffing on inspection days. The regressions are shown in Table 7 and Table 8. We do not find significant reduction of staffing levels on the days immediately prior and after inspection.

It is also possible that nursing homes discharge or rehospitalize patients around inspection days to temporarily reduce the numbers of patients, and increase staffing ratios while not increase the actually numbers of nurses and aides work in the facility. We test this alternative hypothesis by replacing staffing ratios with patient hours as the new dependent variables and replicate the same regressions. The results are shown in Table 9. The coefficients are negative, consistent with the hypothesis, but are not statistically significant and the magnitude is quiet small. For example, using the estimated coefficient from Table 9 (-0.215, Column 1), it only translates to 0.38% fewer patients (median=57 patients). Therefore it seems nursing homes do not aggressively discharge patients on inspection days in order to improve staffing ratios. Perhaps the potentially loss of revenues from discharge outweighs the gains of inflating staffing ratios and intentionally discharging patients is not a likely practice. Based on these results, alternatives such as contracting temporary staffing or adding additional hours/shifts during inspections may be more plausible mechanisms and worthy of further investigation.

5 C. Potential Impacts on Health Deficiencies

Because the spike of staffing level concentrates on the inspection days, other quality measures that are accessed during the inspection may also be inflated due to the temporarily staffing increase. Particularly, the number and severity of health deficiencies that are observed during inspection may be lower than the average level through the year. If so, the inspection results can overstate the quality of nursing homes. Because health deficiency is the only not self-reported quality matrix, it is considered the most important and objective quality domain in the calculation of the quality rating. The potential spillover to deficiency results is particularly of concern. While we can't directly observe the deficiencies level on the non-inspection days, we

use the estimated elasticity of staffing and deficiency from the literature to naively approximate the potential impacts on deficiency results. According to Lin (2014), a 1% increase of RN staffing level reduces the numbers of health deficiency by 2.48%. If the inflated staffing hours on the inspection days are all RN hours, on average, it may temporarily reduce the number of deficiencies by 17.26%. While the impact on deficiencies is substantial, because Oklahoma data does not provide the information of professional mix of the staffs (e.g. RN, LPN, NA), the approximation likely overestimates the true effect and should be taken with caution.

5 D. Generalizability

To enhance the generalizability of our findings, we also obtained quarterly staffing data from the New Jersey's Department of Health. The New Jersey data consists of 363 unique facilities and 3,306 facility-quarter observations from the 1st quarter of 2012 to the 2nd quarter of 2014. While not as detailed as Oklahoma's daily data, New Jersey's quarterly data separates staffing ratios by the level of professional training (i.e. RN, LPN, and CNA). Consistent with the Oklahoma analysis, we find statistically significant evidence that nursing homes are timing inspection and gaming staffing, on both the extensive and intensive margins. About 36.4% of inspection quarters are also the highest staffing quarters of the year (baseline =25%). At the intensive margin, during inspection quarters, CNA and RN ratios significantly improve by 1.0-3.3% and by 2.1-3.9%. The moderate magnitudes of the results are consistent with the Oklahoma data in Figures 2 and 3, showing that facilities can precisely time the inspection and the effect is short-lived. While our study uses nursing home data, the results may also generalize to other economic activities where quality information for an entire time period is absent and rely on

point-in-time administrative inspections to ensure quality floors (e.g. restaurants, fire safety, airline safety).

6. Alternative Hypothesis and Selection Bias:

An alternative explanation of our finding is that the temporarily surge of staffing hours on inspection days is to fulfill the additional workloads related to the inspection, for example, additional administrative and logistic tasks during the inspectors' visits. However, because the reported staffing level only includes direct care hours, the staffing hours related to additional administrative and logistic tasks by definition should not be included in our staffing measures. In addition, after the introduction of quality rating system, most of the amplification of gaming behavior concentrates on nursing facilities with lower quality (Table 4, column (2), (4), (6), and (8)). This asymmetric amplification between good and poor quality facilities suggests that the inflated quality rating should not be entirely resulted from additional administrative and logistic tasks during the inspection.

Since our current sample is based on only 50% of the Oklahoma nursing homes that submitted staffing reports electronically, other facilities that submit paper-based staffing reports may behave very differently. Thus, our results are subject to selection problem and may be biased. We are in the process of transcribing the paper-based staffing reports and include these data in the later analysis should mitigate the concern of selection bias.

7. Conclusions and Implications:

Our preliminary results suggest that nursing facilities are able to boost staffing during inspection days. It is particularly striking that facilities can adjust their staffing very quickly. We also find that the effect is amplified after the implementation five-star quality rating, though we cannot establish whether this relationship is causal. Manipulating staffing level around inspection days could be a widely used strategy to game quality inspection and reporting, in addition to the rehospitalization of sicker patients (Kontzka et al., 2013) for better reported quality. These results are worrisome because not only the staffing levels but also other inspection-based quality domains may be intentionally or unintentionally inflated (deficiencies, 75% weight of 5 star rating). On the other hand, the results on the non-inspection days also suggest staffing levels have increased significantly after the implementation of quality rating, especially among the facilities with poor patient outcomes. Overall, despite her the gaming on inspection days and increase in staffing on non-inspection days in the post-rating period, the quality rating should achieve meaningful quality enhancements, though possibly not as great as those documented in literature. To mitigate such gaming behaviors, policy makers should enforce greater randomization of the timing of health inspections. With the advance of information technology, continuous reporting of staffing between inspections may also be another feasible option to deter the manipulation of quality inspections.

References

Brickley, James, Susan F. Lu, Gerard J. Wedig, 2014, "Malpractice Laws and Incentives to Shield Assets: Evidence from Nursing Homes", Working Paper

Centers for Medicare & Medicaid Services, 2014, "Nursing Home Compare Five-Star Quality Rating System: Year Five Report", prepared by Abt Associates

Dobkin, Carlos, 2003, "Hospital Staffing and Inpatient Mortality", Working Paper

Engle, Robert F. and Jeffery R. Russell, 2004, "Analysis of High Frequency Data", Handbook of Financial Econometrics, Vol 1: Tools and Techniques, Ch 7, 382-424

Evans, William N. and Beomsso Kim, 2006, "Patient Outcomes when Hospitals Experience a Surge in Admissions", Journal of Health Economics, 25, 365-388

Feng, Zhanlian, Yong Suk Lee, Sylvia Kuo, Orna Intrator, Andrew Foster, and Vincent Mor, 2010, "Do Medicaid Wage Pass-through Payments Increase Nursing Home Staffing", Health Service Research, 728-747

Foster, Andrew D. and Yong Suk Lee, 2015, "Staffing Subsidies and the Quality of Care in Nursing Homes", Journal of Health Economics, 41, 133-147

General Accounting Office, 2001, "Nursing Workforce: Recruitment and Retention of Nurses and Nurse Aides Is a Growing Concern", Testimony before the Committee on Health, Education, Labor, and Pension, U.S. Senate

Harrington, Charlene, 2010, "Nursing Home Staffing Standards in State Statues and Regulations"

Harrington, Charlene, Brian Olney, Helen Carrillo, and Taewoon Kang, 2011, "Nurse Staffing and Deficiencies in the Largest For-Profit Nursing Home Chains and Chains Owned by Private Equity Companies", Health Services Research, 106-128

Harrington, Charlene, David Zimmerman, Sarita L. Karon, James Robinson, and Patricia Beutel, 2000, "Nursing Home Staffing and Its Relationship to Deficiencies", Journal of Gerontology: Social Science, 55(5), 278-287

Hirth, Richard A. and Sean S. Huang, 2015, "Quality and Private Prices: Evidence from the Nursing Home Industry", Working Paper

Konetzka, R. Tamara, Daniel Polsky, and Rachel M. Werner 2013, "Shipping Out Instead of Shaping Up: Rehospitalization from Nursing Homes as an Unintended Effect of Public Reporting", Journal of Health Economics, 32, 341-352

Koneetzka, R. Tamara, David C. Grabowski, Marcelo Coca Perraillon, and Rachel M. Warren,

"Nursing Home 5-Star Rating System Exacerbates Disparities in Quality, by Payer Source", Health Affairs, 34:5, 819-827

Kitchener, Martin, Janis O'Meara, Ab Brody, Hyang Yuol Lee, and Charlene Harrington, 2008, "Shareholder Value and Performance of a Large Nursing Home Chain", Health Services Research, 1062-1084

Levin, Laurence, Matthew S. Lewis, and Frank A. Wolak, 2015, "High Frequency Evidence on the Demand for Gasoline", Working Paper

Lin, Haizhen, 2014, "Revisiting the Relationship between Nurse Staffing and Quality of Care in Nursing Homes: An Instrumental Variables Approach", Journal of Health Economics, 37, 13-24

Lu, Susan, 2012, "Information Disclosure, Multitasking and Product Quality: Evidence from Nursing Homes", Journal of Economics and Management Strategy, 21(3), 673-705

New York Times, 2014, "http://www.nytimes.com/2014/08/25/business/medicare-star-ratings-allow-nursing-homes-to-game-the-system.html

The Wall Street Journal, 2015, "Hospital Discharges Rise at Lucrative Times: Facilities Release Medicare Patients After Rules Trigger Higher Payments"

The Washington Post, 2007, "After Delays, Carlyle Wins Manor Care"

The Washington Post, 2013, "Hospice Firms Draining Billions from Medicare"

Williams, Christianna, Alan White, Allison Muma, and Louis Hadden, 2013, "Nursing Home Compare: The First Four Years of the Five-Star Quality Rating System", Abt Associates

Figure 1. Sample of Daily Staffing Report

The following is a sample of nursing home daily staffing report collected by Oklahoma state agency. Nursing homes submitted the reports monthly. The report documents the number of residents and the direct care staff hour three shifts per day.

A) Direct Care Staffing*

Day of The Month	Day Shift 6:30 to 2:30		Evening Shift 2:30 to 10:30		Night Shift 10:30 to 6:30	
	Peak In-House Resident Count	Direct Care Staff Hours	Peak In-House Resident Count	Direct Care Staff Hours	Peak In-House Resident Count	Direct Care Staff Hours
1	27	40	27	24	27	24
2	27	48	27	24	27	24
3	27	48	27	24	27	24
4	27	48	27	24	27	24
5	27	56	27	32	27	16
6	27	32	27	24	27	24
7	27	32	27	24	27	32
8	27	48	27	24	27	24
9	27	48	27	32	27	24
10	27	48	27	24	27	24
11	27	48	27	24	27	24
12	27	56	27	24	27	24
13	27	32	27	24	27	16
14	27	40.	27	24	27	16
15	27	48	27	24	27	32
16	28	64	28	24	28	24
17	28	40	28	24	28	24
18	28	48	28	24	28	24
19	28	40	26	32	27	32
20	27	40	27	24	27	16
21	27	32	27	24	27	32
22	27	40	27	24	27	24
23	27	40	27	24	27	24
24	27	40	27	32	27	16
25	29	48	29	24	29	16
26	29	40	29	24	29	24
27	29	32	29	32	29	40
28	29	40	29	24	28	24
29	28	40	28	32	28	24
30	28	32	28	32	28	16
31	28	32	28	24	28	16

Figure 2. Daily Staffing Dynamics Around Inspection Days (Day Shift)
We show the relative daily staffing ratios 15 days before (t-15) and after (t+15) the first inspection date (x=0). The baseline (y=0) is the average of 30 days around inspection dates (weekends are excluded).

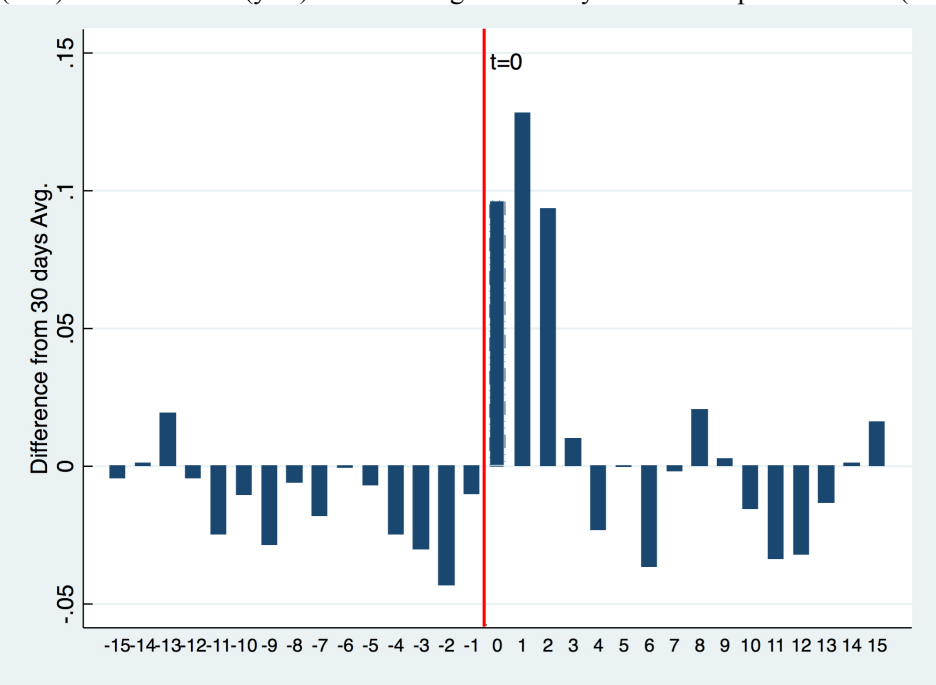

Figure 3. Monthly Staffing Dynamics Around Inspection Month (Day Shift)
We show the relative daily staffing ratios 6 month before (t-6) and after (t+6) the inspection month (x=0). The baseline (y=0) is the 13 months average of daily staffing around inspection dates. We find the monthly effect is only about 10% of the daily effect as shown in Figure 2. The comparison suggests most homes can precisely target the exactly inspection dates.

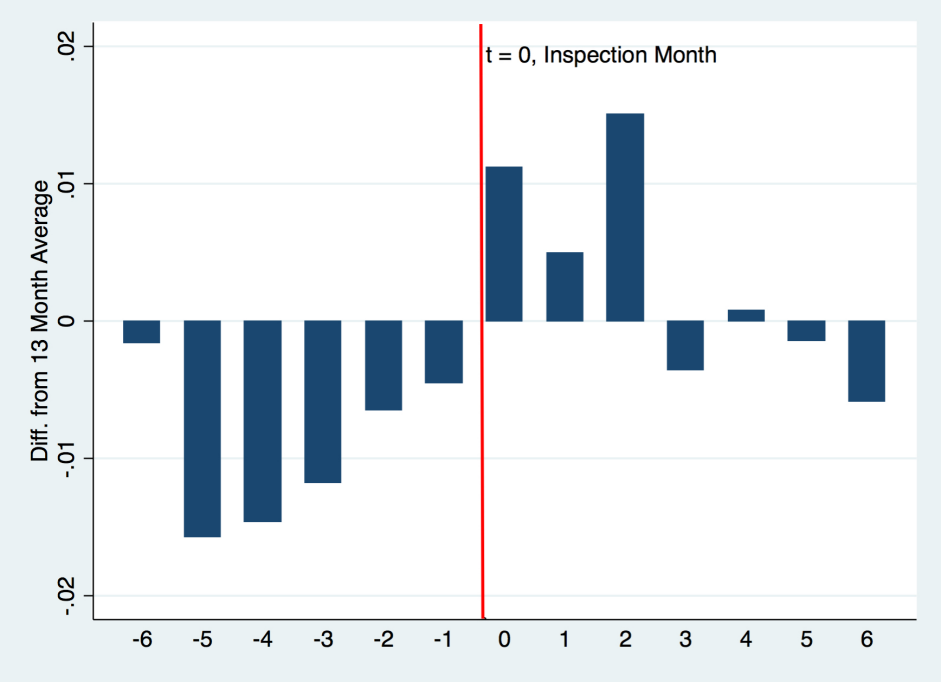

Figure 4. The Distribution of Staffing Ratios (Hrs/Patient, Day Shift) by Inspection Days
Inspection periods are normally last 3 to 4 days. We use the reported health survey date as the beginning of the inspection period, and include 2 following days as the inspection period.

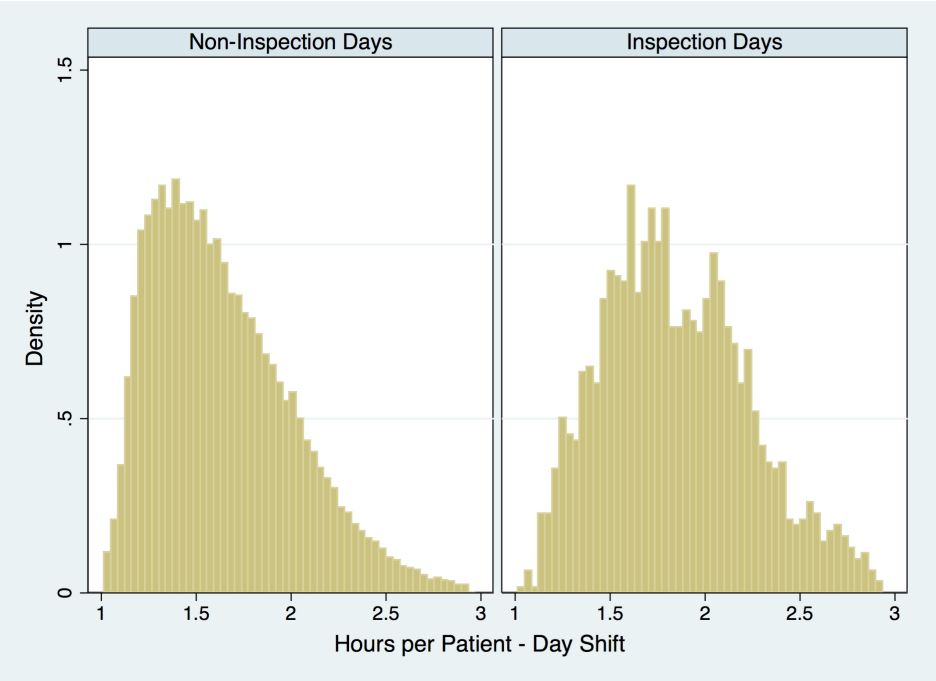

Figure 5. Before and After the Implementation of Five-Star Rating
The five-start quality rating was implemented in 2008/2009, so we selected 2007 and 2012 to represent the distributions in the pre-and post-periods.

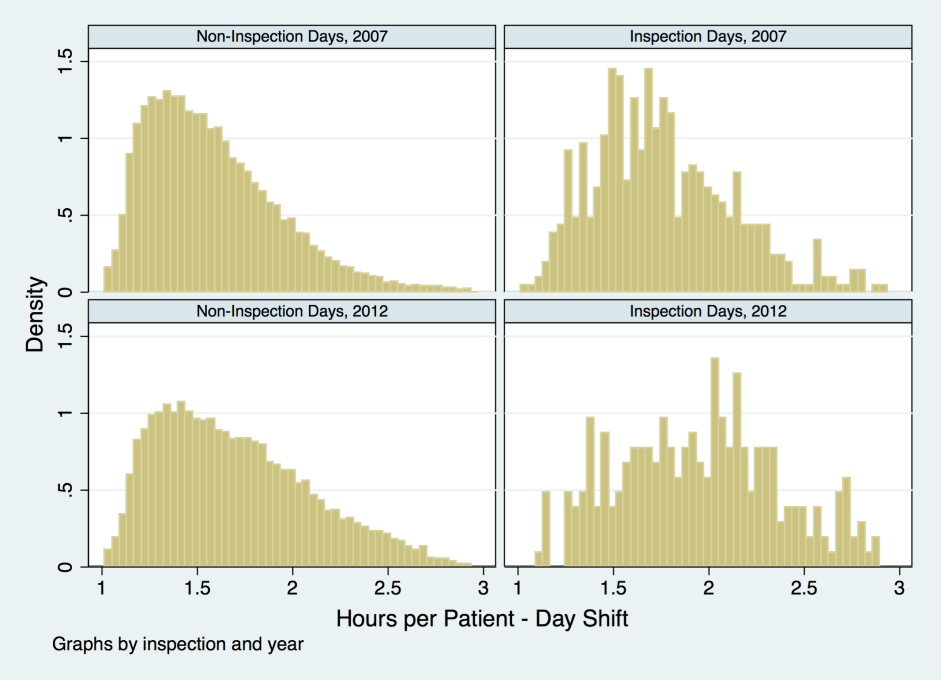

Table 1. Summary Statistics

	Median	Mean	Std	Obs
Staffing Level				
Day Shift	1.58	1.67	2.78	240,039
Evening Shift	1.16	1.20	2.91	240,038
Night Shift	0.66	0.70	0.43	240,003
Weighted Average	3.44	3.56	4.09	240,040
Facility-level Variables				
For-profit	1.00	0.89	0.32	240,040
Nonprofit	0.00	0.09	0.29	240,040
Total Residents	58.00	62.94	26.70	240,040
Chain-Affiliated	0.00	0.35	0.48	240,040
Hospital-based	0.00	0.01	0.09	240,040

Table 2. Regression Results of Staffing Ratios On Inspection Days

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The ***Post X Inspection*** represents the marginal effect of implementation of five-star rating. Hospital-based nursing facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing ratios to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical significance at the 10%, 5%, 1%, and 0.5% level.

	Day Shift		**Evening Shift**	**Night Shift**	**Weighted Total**
	(1)	(2)	(3)	(4)	(5)
Inspection Days	0.110****	0.079****	0.073****	0.022****	0.181****
	[0.009]	[0.014]	[0.010]	[0.006]	[0.024]
Inspection X Post		0.045***	0.015	-0.004	0.051*
		[0.017]	[0.012]	[0.007]	[0.030]
Post	0.093****	0.092****	0.048****	0.015**	0.154****
	[0.015]	[0.015]	[0.009]	[0.006]	[0.025]
Cons	1.148****	1.148****	1.065****	0.653****	2.864****
	[0.023]	[0.023]	[0.010]	[0.006]	[0.032]
R-squared	0.328	0.328	0.059	0.011	0.246
Observations	232,848	232,848	233,120	232,991	233,048

Table 3. Subsample Analysis: "Gaming" Facilities

We identify facilities as gaming facilities if their staffing-levels on inspection days are one standard deviation higher than yearly average. About 20% of the facilities fall into this category. We consider this group of facilities is more likely to intentionally manipulate the inspection results.

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The ***Post X Inspection*** represents the marginal effect of implementation of five-star rating. Hospital-based nursing facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing ratios to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical significance at the 10%, 5%, 1%, and 0.5% level.

	Day Shift		**Evening Shift**	**Night Shift**	**Weighted Total**
	(1)	(2)	(3)	(4)	(5)
Inspection Days	0.373****	0.353****	0.110****	0.042**	0.521****
	[0.016]	[0.028]	[0.028]	[0.019]	[0.055]
Inspection X Post		0.028	0.007	-0.032	-0.02
		[0.031]	[0.036]	[0.021]	[0.067]
Post	0.080*	0.080*	0.071*	-0.023	0.134
	[0.046]	[0.046]	[0.036]	[0.023]	[0.097]
cons	1.122****	1.122****	1.043****	0.689****	2.831****
	[0.053]	[0.053]	[0.035]	[0.017]	[0.091]
R-squared	0.468	0.468	0.098	0.022	0.369
Observations	48,246	48,246	48,763	48,413	48,466

Table 4. Comparison of Staffing Patterns by Aggregated Patient Outcomes

Because patient outcomes are based on quarterly reports and are independent from the health inspection, the aggregated patient outcome is less affected by temporarily staffing inflation. We separate the nursing facilities into better and worse quality groups, and compare their staffing patterns. Facilities with lower quality outcome may face higher pressures to inflate staffing in order to improve inspection results and higher staffing scores. Therefore, we expect to see more substantial gaming behaviors of facilities with worse patient outcomes especially after the implementation of five star quality rating.

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The ***Post X Inspection*** represents the marginal effect of implementation of five-star rating. Hospital-based nursing facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing ratios to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical significance at the 10%, 5%, 1%, and 0.5% level.

*****Dependent Variable is Day Shift Staffing Ratio and sample is split by better or worse than median of quality measures.***

	Pressure Sores		Contractures		Weight Change		Mental Status	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Good	Poor	Good	Poor	Good	Poor	Good	Poor
Inspection Days	0.107****	0.114****	0.085****	0.082****	0.102****	0.067****	0.089****	0.070****
	[0.017]	[0.037]	[0.022]	[0.017]	[0.022]	[0.017]	[0.021]	[0.018]
Inspection X Post	0.025	0.098**	0.032	0.039*	0.017	0.049**	0.032	0.050**
	[0.021]	[0.044]	[0.025]	[0.023]	[0.025]	[0.022]	[0.025]	[0.021]
Post	0.078****	0.206****	0.080****	0.112****	0.064***	0.088****	0.056***	0.137****
	[0.020]	[0.046]	[0.023]	[0.023]	[0.024]	[0.021]	[0.021]	[0.023]
cons	1.143****	2.869****	1.171****	1.133****	1.165****	1.158****	1.153****	1.136****
	[0.027]	[0.047]	[0.030]	[0.029]	[0.033]	[0.026]	[0.029]	[0.028]
R-squared	0.302	0.296	0.313	0.366	0.326	0.343	0.331	0.348
Observations	125974	106987	115208	117640	98478	134370	106973	125875

Table 5. Regression Results of Staffing Ratios By For-Profit Status

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The ***Post X Inspection*** represents the marginal effect of implementation of five-star rating. Hospital-based nurs facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing r to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical signifi at the 10%, 5%, 1%, and 0.5% level.

Panel A. For-profit Nursing Homes

	Day Shift		Evening Shift	Night Shift	Weighted Total
	(1)	(2)	(3)	(4)	(5)
Inspection	0.102****	0.070****	0.033***	0.012	0.119****
	[0.011]	[0.017]	[0.012]	[0.008]	[0.028]
Post X Inspection		0.046**	0.02	-0.005	0.055
		[0.019]	[0.015]	[0.009]	[0.033]
Post	0.092****	0.092****	0.045****	0.012*	0.148****
	[0.016]	[0.016]	[0.010]	[0.007]	[0.027]
R-squared	0.321	0.321	0.054	0.01	0.238
Observations	207,830	207,830	208,545	208,134	208,168

Panel B. Non-profit Nursing Homes

	Day Shift		Evening Shift	Night Shift	Weighted Total
	(1)	(2)	(3)	(4)	(5)
Inspection	0.155****	0.108**	0.057	-0.002	0.188
	[0.038]	[0.051]	[0.050]	[0.028]	[0.114]
Post X Inspection		0.066	0.063	0.013	0.113
		[0.073]	[0.067]	[0.037]	[0.151]
Post	0.101*	0.101*	0.078**	0.048*	0.226**
	[0.058]	[0.058]	[0.030]	[0.025]	[0.092]
R-squared	0.398	0.398	0.086	0.053	0.311
Observations	21,311	21,311	20,814	21,176	21,145

Table 6. Regression Results of Staffing Ratios By Chain Affiliations

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The ***Post X Inspection*** represents the marginal effect of implementation of five-star rating. Hospital-based nursing facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing ratios to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical significance at the 10%, 5%, 1%, and 0.5% level.

Panel A. Chain- affiliated Nursing Homes

	Day Shift		Evening Shift	Night Shift	Weighted Total
	(1)	(2)	(3)	(4)	(5)
Inspection	0.097****	0.068**	0.03	0.019	0.118**
	[0.019]	[0.029]	[0.019]	[0.012]	[0.049]
Post X Inspection		0.044	0.031	-0.003	0.059
		[0.036]	[0.024]	[0.015]	[0.059]
Post	0.069**	0.069**	0.017	-0.002	0.085
	[0.032]	[0.032]	[0.018]	[0.013]	[0.055]
R-squared	0.289	0.289	0.039	0.01	0.209
Observations	81,139	81,139	80,728	81,142	81,006

Panel B. Independent Nursing Homes

	Day Shift		Evening Shift	Night Shift	Weighted Total
	(1)	(2)	(3)	(4)	(5)
Inspection	0.114****	0.084****	0.030**	0.006	0.126****
	[0.011]	[0.019]	[0.015]	[0.009]	[0.031]
Post X Inspection		0.043*	0.029	-0.004	0.062
		[0.023]	[0.019]	[0.010]	[0.039]
Post	0.113****	0.113****	0.063****	0.024****	0.198****
	[0.019]	[0.019]	[0.012]	[0.008]	[0.032]
R-squared	0.353	0.353	0.068	0.018	0.27
Observations	151,709	151,709	152,392	151,849	152,042

Table 7. Regression Results of Staffing Ratios Before Inspection Days

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The ***Post X Before*** represents the marginal effect of implementation of five-star rating. Hospital-based nursing facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing ratios to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical significance at the 10%, 5%, 1%, and 0.5% level.

	Day Shift		**Evening Shift**	**Night Shift**	**Weighted Total**
	(1)	(2)	(3)	(4)	(5)
Before	0.030****	0.011	0.014*	0.012**	0.038*
	[0.006]	[0.011]	[0.007]	[0.005]	[0.019]
Post X Before		0.027**	0.022**	-0.004	0.045*
		[0.012]	[0.009]	[0.006]	[0.024]
Post	0.093****	0.092****	0.048****	0.015**	0.154****
	[0.015]	[0.015]	[0.009]	[0.006]	[0.025]
R-squared	0.325	0.325	0.055	0.011	0.242
Observations	232,848	232,848	233,120	232,991	233,048

Table 8. Regression Results of Staffing Ratios After Inspection Days

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The Post X After represents the marginal effect of implementation of five-star rating. Hospital-based nursing facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing ratios to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical significance at the 10%, 5%, 1%, and 0.5% level.

	Day Shift		**Evening Shift**	**Night Shift**	**Weighted Total**
	(1)	(2)	(3)	(4)	(5)
After	0.008	-0.001	-0.002	0.004	0
	[0.006]	[0.010]	[0.007]	[0.004]	[0.018]
Post X After		0.013	0.011	0	0.026
		[0.013]	[0.009]	[0.006]	[0.023]
Post	0.093****	0.092****	0.048****	0.015**	0.154****
	[0.015]	[0.015]	[0.009]	[0.006]	[0.025]
R-squared	0.325	0.325	0.055	0.011	0.242
Observations	232,848	232,848	233,120	232,991	233,048

Table 9. Regression Results of Patient Hours On Inspection Days

All regressions control for facility-, month-, calendar day-, and week day-fixed effects. The ***Post X Inspection*** represents the marginal effect of implementation of five-star rating. Hospital-based nursing facilities are excluded. We also exclude the observations with the top 99% and bottom 1% staffing ratios to account for reporting errors and unlikely values. *, **, ***, and ****, represent statistical significance at the 10%, 5%, 1%, and 0.5% level.

	Day Shift		**Evening Shift**	**Night Shift**
	(1)	(2)	(3)	(4)
Inspection	-0.215	-0.403	-0.373	-0.455
	[0.136]	[0.278]	[0.277]	[0.288]
Post X Inspection		0.275	0.233	0.331
		[0.335]	[0.334]	[0.342]
Post	0.224	0.224	0.231	0.195
	[0.646]	[0.646]	[0.647]	[0.648]
R-squared	0.001	0.001	0.001	0.002
Observations	233,306	233,306	233,178	233,181