Catching Them Red-Handed: Optimizing the Nursing Homes' Rating System

XU HAN, Fordham University
NIAM YARAGHI, University of Connecticut & The Brookings Institution
RAM GOPAL, Indian Institute of Management

The Centers for Medicare & Medicaid Services (CMS) launched its nursing home rating system in 2008, which has been widely used among patients, doctors, and insurance companies since then. The system rates nursing homes based on a combination of CMS's inspection results and nursing homes' self-reported measures. Prior research has shown that the rating system is subject to inflation in the self-reporting procedure, leading to biased overall ratings. Given the limited resources CMS has, it is important to optimize the *inspection* process and develop an effective *audit* process to detect and deter inflation.

We first examine if the domain that CMS currently inspects is the best choice in terms of minimizing the population of nursing homes that can inflate and minimizing the difficulty of detecting such inflators. To do this, we formulate the problem mathematically and test the model by using publicly available CMS data on nursing home ratings. We show that CMS's current choice of inspection domain is not optimal if it intends to minimize the number of nursing homes that can inflate their reports, and CMS will be better off if it inspects the staffing domain instead. We also show that CMS's current choice of inspection domain is only optimal had there been an audit system in place to complement it. We then design an audit system for CMS which will be coupled with its current inspection strategy to either minimize the initial budget required to conduct the audits or to maximize the efficiency of the audit process. To design the audit system, we consider nursing homes' reactions to different audit policies, and conduct a detailed simulation study on the optimal audit parameter settings. Our result suggests that CMS should use a moderate audit policy in order to carefully balance the tradeoff between audit net budget and audit efficiency.

7

CCS Concepts: • **Information systems → Information systems applications**;

Additional Key Words and Phrases: Audit design, nursing homes' rating system, rating inflation

ACM Reference format:
Xu Han, Niam Yaraghi, and Ram Gopal. 2019. Catching Them Red-Handed: Optimizing the Nursing Homes' Rating System. *ACM Trans. Manage. Inf. Syst.* 10, 2, Article 7 (August 2019), 26 pages.
https://doi.org/10.1145/3325522

1 INTRODUCTION

Nursing homes constitute an important segment of the U.S. healthcare system. They provide care to 1.5 million patients in America (Fowles 2012). Medicare annually spends more than $49 billion on

Authors' addresses: X. Han, Gabelli School of Business, Fordham University, 140 W 62nd St, New York, NY 10023; email: xhan44@fordham.edu; N. Yaraghi, School of Business, University of Connecticut, 1 University Pl, Stamford, CT 06901; email: niam.yaraghi@uconn.edu; R. Gopal, Department of Operations Management, Quantitative Methods and Information Systems, Indian Institute of Management, Udaipur, Rajasthan, India 313001; email: ram.gopal@iimu.ac.in.

2158-656X/2019/08-ART7 $15.00
https://doi.org/10.1145/3325522

the services provided in nearly 16,000 nursing homes in the United States (KFF 2017). The medical, social, and economic importance of nursing homes led the Centers for Medicare & Medicaid Services (CMS) to design and implement a rating system for these facilities in 2008. Given the lack of alternative information resources, this publicly available rating system has become the gold standard in the industry and is being widely used by patients, physicians, and payers (Thomas 2014).

In this rating system, CMS rates each nursing home in a 5-star scale based on its performance in three domains: *Health Inspections, Staffing*, and *Quality Measures.* The health inspections are conducted by CMS-certified inspectors, while the other two domains are self-reported by nursing homes. To rate a nursing home, CMS first conducts an on-site inspection that determines an initial rating. The nursing home will then self-report measures of its quality and staffing which can add or subtract up to two stars to or from its initial inspection rating (CMS 2015). The two self-reported measures can significantly affect a nursing home's overall ratings; for example, a nursing home that initially receives three stars from the on-site inspection can increase its overall ratings to five stars if its self-reports excellent measures on its quality and staffing. Prior research shows that at least six percent of the nursing homes inflate their self-reported measures as a strategy to gain higher overall star ratings (Han et al. 2018).

CMS has limited resources and inevitably has to partially rely on self-reported measures to evaluate nursing homes. This requires (1) an inspection strategy to determine the domain to be inspected directly by CMS and (2) an audit system to detect and deter fraud in the measures that are self-reported by the nursing homes. CMS, however, has neither an inspection strategy nor an audit system. This article, as we describe below, seeks to address these limitations by designing both an inspection strategy and an audit system for CMS to improve its current nursing homes' rating system.

Currently, CMS only inspects the *Health Inspections* domain and leaves the other two domains to be self-reported by the nursing homes. It is not yet known if inspecting this particular domain is the optimal strategy. In this research, we examine other measures which CMS can inspect directly in order to either minimize the population of nursing homes that can inflate their self-reported measures or minimize the difficulty of detecting the ones that engage in such behavior.

When it comes to self-reporting, a typical practice is to design an audit system to detect and deter fraud. A well-known example is the audit system implemented by Internal Revenue Service (IRS), which is shown to significantly reduce the extent of income tax evasion in the US (Cebula 2012). Despite its benefits, CMS currently does not have any audit system in place. To bridge this gap, we design an effective audit system for CMS to control inflation in the self-reported measures. In our design, CMS randomly audits a portion of nursing homes which have gained additional stars through self-reporting and fines the caught inflators according to a pre-announced rate. CMS adds the collected penalties to its budget which funds more audits within the same year. Under this audit mechanism, the reaction of the nursing homes that do not inflate is different from those that do. In other words, although the honest nursing homes' reporting behaviors are not affected by CMS's audit policy, the inflators will change their self-reporting behavior based on their expected profits under different CMS auditing policies.

This study presents major insights for improving the inspection system and designing an audit strategy. We develop a general inspection measure selection model and optimize the selection of inspection domains. We then focus on the case in which the settings are kept as close to CMS's current practice as possible in order to obtain insights that can directly guide CMS's current rating system reform. We tried two objectives by either minimizing the population of nursing homes that can inflate their self-reported measures or minimizing the difficulty of detecting the ones that engaged in such behavior. Our results indicate that, to minimize the population of nursing homes that can inflate, CMS will be better off inspecting the staffing domain. However, the *Health*

Inspection domain CMS is currently inspecting is optimal only in terms of minimizing the difficulty to detect inflators in the audit, and an effective audit system has to be in place to work together with the inspection procedure. We develop conditions on the audit parameter settings and justify our findings through simulation.

The findings of this research have implications for other rating systems with similar features. For example, mandated by the Medicare Access & CHIP Reauthorization Act (MACRA) of 2015, CMS has to calculate a performance score for clinicians in the US based on a similar composition of inspected and self-reported measures. Similar to the nursing homes' star rating, the clinicians' performance score, which is the basis of Medicare payments to physicians, suffers from the shortcomings that we discussed earlier. As a result, our findings about optimal inspection domain selection and audit strategy design also apply in this context.

This article proceeds as follows: In Section 2, we review the related literature on audit systems to detect and deter fraudulent behavior. In Section 3, we develop a general model to formulate the inspection strategy and convert it into a linear optimization problem. We use the publicly available data from CMS to optimize our model and select the inspection domain based on the two objectives of minimizing the population of nursing homes that can inflate, and minimizing the difficulty for detecting inflators. In Section 4, we design the audit system. We consider nursing homes' reaction to different audit policies and derive conditions on the parameter settings of the audit for two objective functions of minimizing the audit's budget and maximizing the audit efficiency. We then conduct a simulation of the audit process to examine our analytical results. Finally, Section 5 concludes the article.

2 LITERATURE REVIEW

Researchers have developed various methods to detect fraud across a wide variety of disciplines from finance and management to sports and academia (Abbasi et al. 2010, 2012, 2015; Bai et al. 2010; Cecchini et al. 2010; Duggan and Levitt 2000; Jacob and Levitt 2003; Lau et al. 2011; Li et al. 2015; Masud et al. 2011; Mayzlin et al. 2012; Wright and Marett 2010; Xiao et al. 2018; Zhang et al. 2013). The fundamental approach of all fraud detection methods is to pinpoint "abnormal patterns" embedded in the data. We group these methods into different streams based on their approach to identify such uncommon patterns.

The first stream of research constitutes of studies that apply "majority rule" to detect abnormal patterns. In this method, researchers first determine the behavior of the majority of the population as the baseline and then identify unusual behaviors by comparing individuals' behaviors with the majority's baseline. Josang and Ismail (2002) and Ma et al. (2013) develop a mechanism to detect fake ratings in which a rater is considered dishonest, if the evaluation from all other raters falls in this rater's beta distribution rejection region. Jindal and Liu (2008) build a logistic regression model using a collection of rating features to distinguish fake ratings. Lim et al. (2010) propose a scoring method to measure the degree of spam for each reviewer to identify fake ratings, and apply the method on an Amazon review dataset. Wang et al. (2011) use a graphical method to analyze the relationship among raters, ratings, and entities. Mukherjee et al. (2013) design a rating fraud detection model which identifies fake ratings by calculating their deviation from the majority. These studies rely on the majority rule for fraud detection and the assumption that the population provides a consistent evaluation of a certain subject. However, this assumption does not always hold true, especially in cases where the population by nature has diversified opinion about the same subject.

In the second stream, researchers focus on to the unusual and abrupt changes in certain indicators to identify fraud. DellaVigna and La Ferrara (2010) propose a method to detect illegal arms trade between countries under embargo using the weapon manufacturers' stock prices as a proxy and analyzing their fluctuations as turmoil and conflicts arise at certain geographical areas. Liu and Sun (2010) propose a method to detect malicious fake ratings based on overall rating as an indicator. When a large amount of fake ratings is submitted over a short period, the overall rating will show unusual sudden change.

In the third stream, researchers compare suspicious behaviors with formerly known honest peers to detect fraud. Dellarocas (2000) detects suspicious ratings using the previously identified honest ratings as a filter to explore dissimilarities. Teacy et al. (2006) evaluate the trustworthiness of a rater by comparing her ratings with the other previously identified honest raters. Liu et al. (2014) utilize difference between local and global ratings to identify fake ratings.

The methodologies discussed above cannot be directly applied by CMS to detect fraudulent self-reporting in the nursing homes' rating system. First, nursing homes are located at various locations, with different market environments and types of patients. As a result, the patients' ratings can be very much diversified, even though they may have received similar services. For example, some issues may be important for certain groups of patients but not for others. In such cases, every opinion can be truth-reflecting, and therefore the majority rule does not apply. Second, the self-reported measures have been used for years without being audited, thus the rating patterns over the year, though probably inflated, can be very consistent. Consequently, it is difficult to identify any "sudden change" in the patterns of self-reported measures. Third, the group of inflation detection methods, like the other two, require substantial data on the characteristics of those who are more likely to commit fraud. However, there is currently no audit system to identify the inflators, and therefore it is very difficult to identify the characteristics of the inflating nursing homes. Fraud detection methods are usually problem-dependent, and to the best of our knowledge, few effective fraud detection methods for the nursing home rating system have been reported in the existing literature.

Recent studies on the CMS nursing home rating system shed light on this nursing home self-reporting inflation problem. Han et al. (2018) collected CMS rating data over 2009–2013 and the corresponding financial data reported by Office of Statewide Health Planning and Development (OSHPD) and patients' complaints data reported by California Department of Public Health (CDPH) for 1219 nursing homes in California to empirically examine the key factors that affect the changes in the star rating of a nursing home. The results indicate a significant positive association between the change in nursing homes' star ratings and their financial incentives. It is also demonstrated that the improvement in ratings cannot be explained by nursing homes' legitimate efforts to improve their service qualities. A prediction model is developed to evaluate the extent of inflation among the nursing homes which identifies 6% to 8.5% nursing homes to be likely inflators in the current system. The results provide important guidelines on evaluating parameters in the nursing home rating system, such as nursing homes' financial conditions and nursing home population in each rating level. Building on the prior research by Han et al. (2018), in this article we first set up a model to identify the domain which CMS should inspect in order to reduce the possibility of inflation, and then accordingly design an audit system.

Given the limitations of auditing resources, it is important to design audit policies in which resources are optimally allocated to maximize compliance. The audit process will be more efficient when the resources are focused on auditing agents that are more likely to be noncompliant. This implies that rather than a static audit procedure in which the probability of being selected for an audit is assigned randomly, the auditing agency should adopt a dynamic strategy in which the probability of audit is adjusted by the information obtained through the previous audits (Raymond 1999; Friesen 2003; Stafford 2008).

For example, Gilpatric et al. (2011) design an auditing strategy in which the audit probability is higher for those firms that the difference between expected and reported emissions is greatest and show that firms report higher levels of emissions under this strategy. Oestreich (2015) investigates how firms change their self-reporting behavior when the auditing strategy shifts from a random process to a competitive mechanism in which firms with lower self-reported measures of emissions are more likely to be subject of an audit. He shows that a competitive audit mechanism leads to more truthful self-reporting. Macho-Stadler and Pérez-Castrillo (2006) design an optimal audit strategy to minimize the level of emissions and show that it is optimal to allocate the resources primarily to the easiest-to-monitor firms and less to those firms that value pollution.

A significant body of literature examine the role of self-reporting in compliance with and enforcement of regulations (Harford 1987; Helland 1998; Innes 1999; Kaplow and Shavell 1994; Livernois and McKenna 1999; Malik 1993). These studies conclude that self-reporting combined with an audit strategy increases compliance. Peer-reviews, as another form of self-reporting, have been shown to be associated with both perceived (Hilary and Lennox 2005) and actual (Casterella et al. 2009) audit quality of accounting firms.

In the context of nursing homes' rating system, CMS has no auditing strategy. This shortcoming has led to an increased level of inflation in the self-reported measures and unreliable ratings. Our work contributes the literature by designing an audit strategy which uses the self-reported performance measures to optimally assign audit probabilities such that the audit budget is minimized and the efficiency of the audit process in terms of the ratio of caught inflators to audited honest nursing homes is minimized.

3 INSPECTION STRATEGY

Without proper monitoring in the self-reporting process, nursing homes have significant incentives to report inflated measures to CMS in order to achieve higher overall star ratings. Such biased ratings will not only mislead those who rely on this information to make medical decisions, but will also undermine those truthful nursing homes that pursue genuine efforts to improve their ratings. The selection of inspection domain determines which domains are left self-reported, and has prolonged effects on the reliability of CMS's ratings.

Currently, CMS uses a method in which the service quality of a nursing home is first evaluated in several itemized measures, either through inspection or through self-reporting. These measures are then consolidated into three domains and each domain is then assigned a star rating. Table 1 gives a summary of the itemized measures in each of the current domain. For these three domains, CMS only inspects one of them and rely on nursing homes to self-report their performance in the other two.

We argue that this may not be the optimal mechanism to conduct inspection due to the following reasons: First, there can be information loss when consolidating itemized measures into three domains. Even there is no inflation issue, this consolidation can lead to inaccuracy in the star rating, i.e., the true service quality difference between two five-star nursing homes may be bigger than the service quality difference between a five-star and a three-star nursing home, which significantly deviate from people's common perception about these star ratings. Second, relying on self-reported data for certain domains in the star rating procedure introduces potential inflation issues. We understand that, due to limited resources, CMS is not able to inspect all the measures, but CMS could potentially divert its resources to inspect other measures other than the measures it currently does.

In this section, we examine alternative inspection strategies under different objectives to identify which measures CMS should inspect in order to ensure that the population of nursing homes that can inflate their self-reported measures is minimized or the difficulty of detecting the ones that engaged in such behavior is minimized. We assume that all nursing homes are profit maximizing. In other words, they seek for the best possible rating to maximize their expected payoff. This assumption is consistent with our data in which the majority of nursing homes are for-profit

Table 1. Itemized Measures for Each Domain (Health Inspection, Staffing, and Quality Measures)

Health Inspection (H: Health; F: Fire Safety)	*Staffing*	*Quality Measures (L: long-star; S: Short-stay)*
Count of Administration Deficiencies (H)	RN hours/day	Percent of residents whose need for help with activities of daily living has increased (L)
Count of Environmental Deficiencies (H)	LPN hours/day	Percent of high-risk residents with pressure sores (L)
Count of Mistreatment Deficiencies (H)	Nurse aide hours/day	Percent of residents who have/had a catheter inserted and left in their bladder (L)
Count of Nutrition and Dietary Deficiencies (H)	Total Licensed hours/day	Percent of residents who were physically restrained (L)
Count of Pharmacy Service Deficiencies (H)	Total Nurse hours/day	Percent of residents with a urinary tract infection (L)
Count of Quality of Care Deficiencies (H)		Percent of residents who self-report moderate to severe pain (L)
Count of Resident Assessment Deficiencies (H)		Percent of residents experiencing one or more falls with major injury (L)
Count of Resident Rights Deficiencies (H)		Percent of residents with pressure ulcers that are new or worsened (S)
Count of Building Construction Deficiencies (F)		Percent of residents who self-report moderate to severe pain (S)
Count of Corridor Walls and Doors Deficiencies (F)		
Count of Electrical Deficiencies (F)		
Count of Emergency Plans and Fire Drills Deficiencies (F)		
Count of Exits and Egress Deficiencies (F)		
Count of Exit and Exit Access Deficiencies (F)		
Count of Fire Alarm Systems Deficiencies (F)		
Count of Furnishings and Decorations Deficiencies (F)		
Count of Hazardous Area Deficiencies (F)		
Count of Illumination and Emergency Power Deficiencies (F)		
Count of Interior Finish Deficiencies (F)		
Count of Laboratories Deficiencies (F)		
Count of Medical Gases and Anesthetizing Areas Deficiencies (F)		
Count of Miscellaneous Deficiencies (F)		
Count of Building Service Equipment Deficiencies (F)		
Count of Smoke Compartmentation and Control Deficiencies (F)		
Count of Smoking Regulations Deficiencies (F)		
Count of Automatic Sprinkler Systems Deficiencies (F)		
Count of Vertical Openings Deficiencies (F)		

Table 2. Nursing Homes' True Service Quality
(0-Not good, 1-Good)

Nursing Home	Measure A	Measure B
1 (Inflator)	0	0
2	0	1
3 (Inflator)	0	0
4	0	1
5	1	1

Table 3. Rating Pattern Distribution

Rating Pattern (AB)	Probability to observe rating pattern (AB)
(00)	0.4
(01)	0.4
(10)	0
(11)	0.2

nursing homes. We formulate CMS's inspection strategy in two models then empirically test our models using CMS's historical data.

3.1 An Illustrative Example

In this section, we introduce an illustrative example to explain different inspection objectives CMS can use. We mainly consider two objectives: minimizing the population of nursing homes that can potentially inflate their ratings and minimizing the difficulty to detect inflators in the following audit procedure.

Suppose CMS is running a small system with five nursing homes. Each nursing home is evaluated based on two measures, called A and B. Each measure has two levels, and we use 1 to denote good service quality and 0 to denote bad service quality. The true service quality for all five nursing homes is shown in Table 2. Due to limited budget and human resources, CMS can only inspect one of the two measures and leave the other one to be self-reported by nursing homes. The question is to determine which measure is the optimal for CMS to inspect. Suppose that among the five nursing homes, nursing homes 1 and 3 are inflators, and they seek every opportunity to inflate their measures in order to achieve better ratings. CMS does not know which nursing home in the population is an inflator. However, CMS has the rating pattern distribution as shown in Table 3, which can be used in the audit to detect suspicious patterns.

In this example, 80% of nursing homes are not performing well on measure A, while 40% of nursing homes are not performing well on measure B. From CMS's perspective, a large population (80%) can potentially inflate on measure A, while only 40% can potentially inflate on measure B. As a result, CMS may choose to conduct inspection on A in order to minimize the potential population of inflators. However, in this case, nursing homes 1 and 3 will choose to inflate on measure B, and inspecting measure A can lead to the rating pattern as shown in Table 4. The two inflators will blend in into common rating patterns, leading to the same rating pattern for nursing homes 1–4, and can be relatively difficult to detect the inflators in the following audit. On the other hand, if CMS conducts inspection on measure B, nursing homes 1 and 3 can only inflate on measure A, leading to the rating pattern as shown in Table 5.

Table 4. Rating Patter if CMS Inspects Measure A

Nursing Home	Measure A	Measure B
1 (Inflator)	0	1
2	0	1
3 (Inflator)	0	1
4	0	1
5	1	1

Table 5. Rating Patter if CMS Inspects Measure B

Nursing Home	Measure A	Measure B
1 (Inflator)	1	0
2	0	1
3 (Inflator)	1	0
4	0	1
5	1	1

According to the rating pattern distribution, (10) is a pattern which was rarely observed before. This can trigger a red flag immediately and result in an easy detection of the two inflators in the audit. For each nursing home, it will end up with a rating pattern with certain probability after self-reporting. We define the audit difficulty to be the sum of these rating pattern probabilities for all nursing homes. In this particular example, each inflator ends up with a rating pattern with probability 0, thus the total audit difficulty is defined to be $0 + 0 = 0$. For CMS, conducting inspection on measure B minimizes the difficulty to detect inflators in the audit. In the following sections, we will discuss the mathematical formulation for minimizing potential inflator population and minimizing the difficulty to detect inflators in detail and test our models by using CMS data.

3.2 Minimizing the Population of Nursing Homes that Can Inflate

In this formulation, we minimize the population of nursing homes that can inflate. A detailed list of notations and symbols is provided in Table 6.

Consider a problem with K measures indexed by k and I nursing homes indexed by i. For each measure, there are m possible levels that a nursing home can self-report. The measure which CMS inspects is denoted by a binary decision variable y_k where $k \in \{1, \ldots, K\}$ and $y_k = 1$ *iff* the measure is selected for inspection. For each nursing home i, we introduce a binary status variable x_i, where $i \in \{1, \ldots, I\}$ and $x_i = 1$ *iff* the nursing home is inflating. x_{ik} is a binary status variable introduced to indicate whether nursing home i inflates on measure $k \cdot x_{ik} = 1$ *iff* nursing home i chooses to inflate on measure k. We use $\Delta prof_{ikm}$ to denote the profit gain if nursing home i chooses to report level m on measure k. $\Delta prof_{ikm} = 0$ if the nursing home reports truthfully on measure k. $\Delta prof_{ikm} > 0$ if the nursing home inflates its self-reported level on measure k. The problem has a two-level structure: At the lower level, nursing homes maximize their expected profits by choosing the level to report at each measure. At the higher level, CMS minimizes the population of nursing homes that can inflate by selecting measures to conduct inspection on. The detailed formulation is shown below.

$$\min_{y_k} \sum_i x_i \tag{1}$$

Table 6. Notations and Symbols used in the Inspection Measure Selection Problem

i	Nursing home index
K	Inspection measure index
J	Rating pattern index
m	Measure level index
x_i	Status variable, $x_i = 1$, if nursing home i is a nursing home that can inflate
x_{ik}	Status variable, $x_{ik} = 1$ if nursing home i will inflate measure k
y_k	Decision variable. $y_k = 1$ if measure k is selected for inspection
z_{ij}	Status variable, $z_{ij} = 1$ if nursing home i will show rating pattern z after self-reporting
c_k	The cost to inspect measure k
C	CMS's total inspection budget
$prof_{ikm}$	The profit gain nursing home i can obtain for reporting level m at measure k
p_{km}	The probability of being audited for reporting level m at measure k
$Prob(Pattern_j)$	The probability of showing pattern j
M	A big positive number used for linear conversion.

$$\text{s.t.}\quad \text{Budget constraint:}\quad \sum\nolimits_k y_k c_k \leq C, \tag{2}$$

$$\text{Inspection constraint:}\quad x_{ik} \leq 1 - y_k, \forall i \tag{3}$$

Inflation rational constraint:

$$\mathrm{M}(y_k + x_{ik}) \geq \Delta prof_{ikm}(1 - p_{km}) - p_{km} r\ \Delta prof_{ikm}, \forall i, \forall k, \tag{4}$$

where M is a big number.

$$\text{Inflator status constraint:}\quad x_i \geq x_{ik}, \forall k \tag{5}$$

Equation (1) denotes the total population of nursing homes that can inflate, which is the objective of this model. Equation (2) represents the budget constraints of CMS, where c_k is the cost for inspecting measure k, and C is the overall inspection budget. Equation (3) denotes the relationship between status x_{ik} and the inspection domain y_k. For any nursing home, if measure k is selected for inspection by CMS, i.e., $y_k = 1$, then it is impossible to inflate measure k, and x_{ik} is fixed at 0 by Equation (3).

One of the key challenges for formulating CMS's inspection measure selection problem is how to capture nursing homes' profit maximizing decisions at the lower-level problem. Equation (4) describes the condition under which nursing homes may choose to inflate. If a nursing home chooses to inflate self-reported measures, its expected payoff from inflating must be higher than the expected payoff for staying honest. Although CMS does not have an effective audit system for the data period that we are focusing on,[1] the reform direction is to gradually incorporate audits into the rating procedure, starting from the staffing domain. In our problem setting, we assume that

[1]Since launched in 2008, CMS does not incorporate an effective and systematical audit system to ensure the accuracy of nursing homes' self-reported measures. Since early 2017, CMS started to collect payroll data from nursing homes and in the latest version of "Design for Nursing Home Compare Five-Star Quality Rating System: Technical Users' Guide" published in July 2018, CMS mentions that audits are conducted on nursing homes' reported staffing hours. The audit was not mentioned in CMS's previously published rating mechanism documents. This is the first time CMS uses audit in its rating system, and it is only applied on the staffing domain. The details and the effectiveness of the audit are still to be evaluated, and no audit has been designed for the quality measures.

CMS conducts a random audit according to a probability p. If nursing homes are caught inflating, they will be fined based on a punishment rate r. This is consistent with our problem setting in Section 4 when we systematically discuss CMS's audit design.

Suppose nursing home i chooses to report level m for measure k. The profit gain is $\Delta prof_{ikm}$, and the probability for such nursing homes to be audited is p_{km}. If the nursing home is inflated and gets audited, it has to return its illegal profit gain and will be punished based on a preannounced punishment rate r. As a result, the expected payoff for nursing home i can be expressed as $\Delta prof_{ikm}(1 - p_{km}) - p_{km} r\ \Delta prof_{ikm}$.

It is important to notice that nursing homes' inflation may not necessarily result in star rating increase, because there can be nursing homes whose self-reported measures are at the one-star level which will result in losing stars according to the rating mechanism, but they inflate to avoid losing stars in the overall rating. Our formulation is able to also capture this type of inflation, since gaining additional stars in the overall rating and avoiding losing stars can both be viewed as gaining additional profits and can be reflected by $\Delta prof_{ikm}$. If an honest nursing home i is reporting truthfully, then $\Delta prof_{ikm} = 0$. In reality, the number of nursing homes whose star ratings decrease after self-reporting is far less than those who gain additional stars after self-reporting. For example, in 2013, over 60% of nursing homes are self-reporting 4 or 5 stars in the staffing domain, which may result in additional stars in overall rating according to the rules. However, less than 5% nursing homes are self-reporting 1 star, which may result in decreased overall rating. As a result, we focus more on the impact of the star increase after self-reporting.

Given nursing homes' self-reporting decision m, we can embed nursing homes' inflation rational constraint into CMS's problem formulation as a constraint. In Equation (4), the right-hand-side shows nursing homes' expected payoff, which is related to nursing homes' inflation status x_{ik} on measure k and CMS's audit decision y_k. If $y_k = 0$, i.e., measure k is not inspected, and $\Delta prof_{ikm}(1 - p_{km}) - p_{km} r\ \Delta prof_{ikm} > 0$, then x_{ik} will be fixed at 1, i.e., nursing homes are expected to have a positive profit gain by inflating this uninspected measure, and they will choose to inflate. On the other hand, if $y_k = 1$, Equation (4) itself imposes no restriction on x_{ik} and becomes redundant. In other words, if measure k is inspected, then no matter the expected payoff $\Delta prof_{ikm}(1 - p_{km}) - p_{km} r\ \Delta prof_{ikm}$ is positive or negative, nursing homes would not be able to inflate on this measure. Equation (5) defines the status of an inflator. If a nursing home inflates on any measure k, it will be identified as an inflator.

3.3 Minimizing the Difficulty of Detecting Inflators

In this section, we discuss the other objective of the inspection strategy, which facilitates the detection of fraud in the measures which are left to the nursing homes to self-report. That is, we intend to minimize the *audit difficulty* by inspecting the measures such that inflating nursing homes, after inflating the remaining uninspected measures, show some rare patterns in their rating combination, which makes them easier to be detected. This will facilitate and streamline the auditing process which we will design later in the article. For example, suppose a rating combination is very common and occurs with a relatively high probability. If an inflating nursing home achieves this rating combination after fraudulent self-reporting, it will be less likely to draw CMS's attention. On the contrary, if an inflating nursing home achieves a rare rating combination with a low probability, the case will be highly suspicious and easier for CMS to detect in its subsequent auditing process. Intuitively speaking, we try to strategically "redirect" the inflation, such that the inflators will show some rare rating pattern after self-reporting and can be detected easily.

In this formulation, we minimize the *audit difficulty* which we define to be the sum of probabilities for showing the rating patterns after self-reporting. For example, under certain inspection measure selection, if an inflator ends up with a rating pattern with probability 0.2 after self-reporting,

Table 7. Possible Ratings Patterns after Inflation and Their Probabilities

True Pattern	Possible Patterns after Inflation				
	Pattern 1	Pattern 2	Pattern 3		Pattern 2^k
m_1 m_2 m_3 ... m_k	m_1 m_2 m_3 ... m_k	m_1+ m_2 m_3 ... m_k	m_1 m_2+ m_3 ... m_k	...	m_1+ m_2+ m_3+ ... m_k+
Pattern Probabilities	Prob (Pattern$_1$)	Prob (Pattern$_2$)	Prob (Pattern$_3$)	...	Prob (Pattern$_{2^k}$)
Status Variables	$x_{i1} = 0$, $x_{i2} = 0$, $x_{i3} = 0$, ..., $x_{ik} = 0$	$x_{i1} = 1$, $x_{i2} = 0$, $x_{i3} = 0$, ..., $x_{ik} = 0$	$x_{i1} = 0$, $x_{i2} = 1$, $x_{i3} = 0$, ..., $x_{ik} = 0$	...	$x_{i1} = 1$, $x_{i2} = 1$, $x_{i3} = 1$, ..., $x_{ik} = 1$

then the audit difficulty to detect this inflator is defined to be 0.2. If there are ten such inflators in the system, then the difficulty to detect them is defined to be the sum of probabilities of the rating pattern, which is $10 * 0.2 = 2$.

Formulating the rating patterns each nursing home can end up with after inflation is the key challenge for this model. Consider a problem with k measures, and a nursing home whose true measures are $\{m_1, m_2, m_3, \ldots, m_k\}$, as shown in first column of Table 7. Each measure can have multiple levels. For any measure, if the nursing home decides to report higher than its true level, then the nursing home is inflating. For a problem with k measures, the total number of possible patterns that the nursing home can end up with after inflation is 2^k, as shown in the second column to the last column of Table 7. For each possible pattern, the probability of showing the pattern is also listed in Table 7 and is denoted as *Prob* (*Pattern*). This probability can be calculated by using CMS's data following the method presented in Han et al. (2018). For each possible pattern, the conditions on the status variables x_{ik} are shown in the last row of Table 7.

We use "+" to denote that nursing homes can inflate the measure to obtain additional stars after self-reporting. The self-reporting rules for the Staffing and Quality Measure domain, however, are a little different here. While CMS only allows nursing homes with 5 stars in the Quality Measure domain to gain additional stars, they allow nursing homes with both 4 and 5 stars in the Staffing domain to have additional stars, as long as the staffing rating is higher than the inspection rating. This complication, however, will not affect the validity of our methodology, because both 4 and 5 stars on staffing domain are qualified for "+", and can be easily captured by our model.

Table 7 thus covers all the possible rating changes after self-reporting. For example, the third column represents that the nursing home only wants to inflate on measure 1 from m_1 to m_1+, and will truthfully report other measures, ending up with a pattern $\{m_1+, m_2, m_3, \ldots, m_k\}$ with probability *Prob* (*Pattern*$_1$). In this case, we have $x_{i1} = 1$, and all other status variables x_{ik} to be 0. It is important to notice that we do not consider how much a nursing home deviates on each individual measure in this procedure. Rather, we focus on the status variable x_{ik}, showing that the nursing home can inflate on that measure. It is also important to notice that some measures in the true pattern $\{m_1, m_2, \ldots, m_k\}$ may be low and can lead to a punishment in the overall rating, e.g., one star in the current three-domain system. For such measures, self-reporting a higher level is also considered to be inflation, though the overall rating may not increase. Our model is capable of capturing these scenarios.

For nursing home i, the probability of showing certain rating pattern after inflation can be expressed by logical constraints containing status variables x_{ik} and the probability of showing that pattern. For example, to show Pattern 2 after inflation, it requires $x_{ik} = 1$ and all other status variables x_{ik} to be 0. The probability can be expressed as

$$x_{i1}(1 - x_{i2})(1 - x_{i3})\ldots(1 - x_{ik})Prob(Pattern_2) \tag{6}$$

It is important to notice that the patterns are exclusive and the probabilities of showing each pattern sum up to 1. We introduce a binary status variable z to denote the status of the ending rating pattern. $z_{ij} = 1$, iff pattern j is the pattern that nursing home i end up with after self-reporting, $j \in \{1, 2, \ldots, 2^k\}$. Equation (6) can then be written as $z_{i2}Prob(Pattern_2)$, with

$$z_{i2} = x_{i1}(1 - x_{i2})(1 - x_{i3})\ldots(1 - x_{ik}) \tag{7}$$

The probability that a nursing home will end up with after self-reporting can thus be expressed as

$$\sum\nolimits_j z_{ij}Prob(Pattern_j), \text{ where } \sum\nolimits_j z_{ij} = 1 \tag{8}$$

Equation (7) has to be converted into linear for the problem to be solved efficiently as a mixed integer programming problem. The linear conversion is described as follows: From Equation (7), it is known that $z_{i2} = 1$ iff $x_{i1}(1 - x_{i2})(1 - x_{i3})\cdots(1 - x_{ik}) = 1$, i.e., $x_{i1} = 1$ & $1 - x_{i2} = 1$ & … & $1 - x_{ik} = 1$. This is equivalent to $x_{i1} + 1 - x_{i2} + \cdots + 1 - x_{ik} = k$. As a result, we replace Equation (7) by Equation (9). z_{i2} will be fixed at 1 iff $x_{i1} = 1$ and other status variables x_{ik} are all 0. Otherwise, z_{i2} will be 0 due to minimization.

$$z_{i2} \geq x_{i1} + 1 - x_{i2} + 1 - x_{i3} + \cdots + 1 - x_{ik} - k + 1 \tag{9}$$

Following the same method, all the logical constraints can be systematically converted to linear forms.

This problem also has a two-level structure, with nursing homes maximizing their expected profits by choosing the level to report at each measure at the lower level, and CMS minimizing the difficulty to detect inflators by selecting inspection measure at the higher level. The detailed formulation is shown below.

$$\min_{y_k} \sum\nolimits_i \sum\nolimits_j z_{ij}Prob(pattern_j) \tag{10}$$

$$\text{s.t.} \quad \text{Budget constraint:} \quad \sum\nolimits_k y_k c_k \leq C, \tag{11}$$

$$\text{Inspection constraint:} \quad x_{ik} \leq 1 - y_k, \forall i \tag{12}$$

Inflation rational constraint:

$$M(y_k + x_{ik}) \geq \Delta prof_{ikm}(1 - p_{km}) - p_{km}r\, \Delta prof_{ikm}, \forall i,\ \forall k, \tag{13}$$

where M is a big number.

$$\text{Inflator status constraint: } x_i \geq x_{ik}, \forall k \tag{14}$$

and all linearized logical constraints described above.

Table 8. Rating Pattern Probabilities (1 for a 5-star Rating and 0 for other Ratings, Based on CMS Pooled Data 2009–2013)

Pattern	**000**	**100**	**010**	**001**	**101**	**110**	**011**	**111**
Probability	0.6955	0.0540	0.0349	0.1667	0.0267	0.0055	0.0119	0.0048

3.4 A Special Case for CMS

In Sections 3.2 and 3.3, we proposed two general formulations for minimizing the population of nursing homes that can inflate and minimizing the difficulty to detect inflators in the following audit. However, we are aware that CMS may have budget and human-resource constraints or regulations which may limit the implementation of our proposed model. In this section, we seek to keep the setting of the problem as close to the current CMS's practice as possible so that the limitations from the constraints mentioned above can be minimized, and the insights obtained can be easily applied to the CMS's rating system reform. Specifically, we focus on the situation in which CMS keeps the current three-domain rating structure and only select one domain to conduct inspection on. We test this situation by using CMS data to find out which domain is the optimal inspection choice for CMS.

We collect data on 1,219 California nursing homes over the 5 years since the inception of CMS's nursing home rating system from 2009 to 2013. The data consists of three parts: *ratings, finances,* and *complaints*. The *ratings* dataset is collected directly from CMS and contains nursing homes' ratings in the three domains as well as their basic information, such as location, size, certification, and ownership. The *finances* dataset is obtained from California Office of Statewide Health Planning and Development (OSHPD) and contains the detailed revenues and expenses of each nursing home. The *complaints* data is obtained from California Department of Public Health (CDPH) and contains detailed complaints, incidents, and deficiency reports of all California nursing homes. The CDPH complaint data not only covers complaints that CMS have already considered in its rating procedure, but also includes complaints and deficiency reports that are only state-wide and are not reported to CMS. The three datasets are combined and used for calculating the parameters needed in our model, such as the profit gain nursing homes can have when inflating, and the probability of showing each rating pattern. We also used the data and followed the method presented in Han et al. (2018) to calculate the distribution of nursing homes at each rating pattern. In Han et al. (2018), the authors introduced a method to systematically exclude suspicious nursing homes that are likely inflating either to gain additional stars in the overall rating or to avoid losing stars. The remaining nursing homes have no sign of any rating inflation behavior and are believed to be completely honest nursing homes. The rating distribution of these nursing homes are believed to be unbiased and can be used as historical known truth to check potential inflation. As a result, our result does not suffer from the potential distribution distortion caused by the existing inspection policy. Table 8 listed the probability of showing each rating combination at each rating combination.

Each of the two models are solved in the IBM ILOG CPLEX Optimization Studio (IBM 2015). In our dataset, there are a total of 1,219 nursing homes in the California market. If no inspection is conducted, every nursing home can inflate. If every measure is inspected, there will be no measure that a nursing home can inflate, thus the population of nursing homes that can inflate will be 0. In our analysis, we do not discuss these two trivial cases, but focus on the cases in which one or two measures are inspected. For the period of data we collected, CMS does not have an audit on any domain, thus the audit probability p_{km} is 0 and the inflation rational constraint reduces to $M(y_k + x_{ik}) \geq \Delta prof_{ikm}$, i.e., nursing homes can inflate on the domains that are not inspected.

Table 9. Inspection Measure Selection Results (H for Health Inspection, S for Staffing, Q for Quality Measures)

Inspected Measures	**H**	**S**	**Q**	**HS**	**HQ**	**SQ**
Objective Value	1209	1153	1189	1125	1178	985
Prevented Inflation (%)	0.80%	5.40%	2.50%	7.70%	3.40%	19.20%

(a) Minimizing the Population of Nursing Homes That Can Inflate

Inspected Measures	**H**	**S**	**Q**	**HS**	**HQ**	**SQ**
Objective Value	1.57	51.81	19.36	28.69	11.41	115.94
Relative Audit Difficulty	1	33.00	12.33	18.27	7.27	73.84

(b) Minimizing the Difficulty to Detect Inflators

Table 9(a) shows the result for minimizing the population of nursing homes that can inflate. The objective values are reported in the first row and the percentage of nursing homes that are prevented from inflating their self-reported measure by inspecting the corresponding domains are listed in the second row. It can be seen that inspecting two domains can prevent more nursing homes from inflating their self-reported measures than inspecting only one domain; however, there are exceptions. For example, inspecting the *staffing* domain (denoted as S in the table) alone can control more inflators than inspecting both *health inspections* (denoted as H in the table) and *quality measures* (denoted as Q in the table) domains. The reason is that according to the historical ratings data, the *staffing* is a domain that most inflators will inflate, thus the population of nursing homes that can inflate is still high even when both domains of *health inspections* and *quality measures* are inspected. The costs of inspecting each domain is difficult to estimate accurately. In our analysis, we keep the setting as close to CMS's current setting as possible, i.e., we focus on the cases in which only one domain is inspected. Our results show that if CMS intends to minimize the population of nursing homes that can inflate, it should inspect the *staffing* domain. As shown in the second row of Table 9(a), compared to no inspection at all, inspecting the *staffing* domain prevents 5.40% of the rating inflation while inspecting the *health inspections*, the domain that CMS currently inspects, only prevents 0.80% of inflators from exaggerating their self-reported measures.

The objective values of different inspection domain selections when minimizing the difficulty to detect inflators are reported in the first row in Table 9(b). It is however easier to interpret the results by comparing the audit difficulties of different domain selections to CMS's current practice, which is conducting an inspection on the *health inspection* domain. If we set the relative audit difficulty of the current CMS' inspection practice to be 1, then the relative audit difficulty for other domain selections can be calculated as reported in the second row of Table 9(b). As shown in Table 9(b), when minimizing the difficulty for detecting inflators, the current practice of conducting health infections is the optimal choice for CMS. Since CMS does not have a complete audit system for nursing homes' self-reported measures, its current choice of inspection domain can be viewed as a lost opportunity since it minimizes the difficulty of an audit which does not take place. To address this shortcoming, in the next section we design an audit strategy for CMS according to its current inspection practice. An interesting observation is that conducting inspections on more domains may not necessarily reduce the audit difficulty. For example, our results show that conducting inspection only on *quality measures* may lead to an easier inflator detection than inspecting both the *health inspection* and *staffing* domains. This is due to the fact that when more domains are inspected, there is less room left for nursing homes to inflate. In this case, the inflators will only inflate some of the measures, and may end up with some popular ratting pattern, resulting in a more difficult inflation detection.

Table 10. Notations and Symbols used in the Audit Simulation

p_1	The audit percentage of nursing homes whose star ratings increase one star after self-reporting
p_2	The audit percentage of nursing homes whose star ratings increase one star after self-reporting
r	The punishment rate
B_0	The net budget CMS has at the beginning of the year
C	The cost for auditing one nursing home
$\Delta prof_1$	The additional profit a nursing home can gain by improving its rating by one star
$\Delta prof_2$	The additional profit a nursing home can gain by improving its rating by two stars
PCI	The percentage of caught inflators
PAH	The percentage of audited honest nursing homes

4 AUDIT STRATEGY

In this section, we focus on the audit system design for CMS and conduct a one-year audit simulation based on the most recent year (2013) of the nursing homes' ratings data we collected. Table 10 lists all the notations and symbols used in the audit simulation.

To design the audit system, we assume that CMS continues to inspect the current domain (*health inspections*) to minimize the difficulty for detecting inflators. We also assume that CMS publicly announces the following audit policies:

- Nursing homes in which their overall star rating increases after self-reporting are subject to audit. An audit can distinguish honest from inflating nursing homes without any errors.
- Nursing homes in which their rating increases 1 or 2 stars after self-reporting will be randomly selected for auditing. The probabilities to be selected are p_1 and p_2, respectively.
- Each caught inflator is subject to a fine calculated based on the illegitimate profit it has gained through inflation. The punishment rate is r. For example, if a 3-star nursing home inflates its rating to 5 stars and consequently increases its per patient profit from 10 to 17, then the nursing home's illegitimate profit is $17 - 10 = 7$. If it is caught in the audit, its punishment will be $7 \times (1 + r)$. Following Han et al. (2018), the expected profit for each nursing home is calculated using the OSHPD financial data.
- CMS is a federal agency and its budget is assigned on an annual basis. To reflect this fact in the audit simulation, we assume that CMS has a fixed net audit budget (B_0) at each year. The total budget available for supporting audit includes B_0 plus the fines collected from caught inflators. We further assume that all fines collected from caught inflators and the net audit budget B_0 are all allocated to auditing nursing homes within the same year.

4.1 Nursing Homes' Reaction to Audit Strategy

Using the most recent available data and following the method presented in Han et al. (2018), we can identify the likely inflators amongst the nursing homes. Since no audit system was in place in 2013, nursing homes could inflate their self-reported measures without any consequence. We therefore identify honest nursing homes as those that did not inflate their self-reported measures in 2013, despite the fact that they would have faced no consequences had they chosen to do so. We assume that honest nursing homes will have no intention to inflate their self-reported measures, regardless of the CMS's audit policy. Unlike the honest nursing homes, the inflators will change their self-reporting behavior according to different audit policies as we discuss below.

Nursing homes that have their overall ratings increase by one or two stars are randomly selected for auditing according to pre-announced probability p_1 and p_2, respectively. If an inflator is caught, the illegitimate additional profit gained through inflating will be confiscated and a fine will be issued against the nursing home based on the pre-announced rate r. The expected profit for nursing homes in each of the rating level is determined by using OSHPD financial data. We denote the additional profit that a nursing home can achieve as a result of inflating one or two stars as $\Delta prof_1$ and $\Delta prof_2$, respectively. According to the rating mechanism, nursing homes receiving one star or four stars in health inspections can only increase their overall ratings by one star, and thus for them, $\Delta prof_1$ and $\Delta prof_2$ are equal.

As a result, for a given combination of p_1, p_2, and r, a nursing home that considers inflating will calculate its expected payoff for the following three choices:

- Being honest: $\text{Payoff}_0 = 0$
- Inflating one domain: $\text{Payoff}_1 = \Delta prof_1\ (1 - p_1) - p_1 r\ \Delta prof_1 = \Delta prof_1 (1 - p_1 - p_1 r)$
- Inflating both domains: $\text{Payoff}_2 = \Delta prof_2 (1 - p_2) - p_2 r \Delta prof_2 = \Delta prof_2 (1 - p_2 - p_2 r)$

The nursing home will inflate its rating if $1 - p_1 - p_1 r > 0$ or $1 - p_2 - p_2 r > 0$. That is

$$r < \frac{1 - p_1}{p_1}, \text{ or} \tag{15}$$

$$r < \frac{1 - p_2}{p_2}. \tag{16}$$

If either Equation (15) or Equation (16) is satisfied, then the nursing home will choose to inflate. If Equations (15) and (16) are both satisfied, then the nursing home compares the expected payoffs and will inflate the measures of two domains instead of only one, if

$$\Delta prof_2 (1 - p_2 - p_2 r) > \Delta prof_1 (1 - p_1 - p_1 r). \tag{17}$$

4.2 Objective Functions of Audit Strategy

An increase in the overall star rating is not necessarily due to inflating self-reported measures, but can also be the result of legitimate efforts by honest nursing homes to improve their performance. From CMS's point of view, auditing an honest nursing home wastes audit resources, while auditing an inflating nursing home will deter others and result in collection of additional penalties, which will fund more audits.

Suppose the population of nursing homes in which the ratings increase by 1 and 2 stars are π_1 and π_2, respectively. Note that π_1 and π_2 include both honest and inflating nursing homes. Both π_1 and π_2 are functions of p_1, p_2 , and r. The unit cost for auditing a nursing home is denoted by c. The fine collected from the caught inflators is also a function of p_1, p_2, and r and is denoted by $F(p_1, p_2, r)$. We consider the following two objectives for designing the audit system, both of which are important indicators of the performance of an audit system.

- Minimizing the audit budget.
- Maximizing the audit efficiency as denoted by the ratio of caught inflators to audited honest nursing homes.

For both of these objectives, the following budget constraint should be satisfied.

$$\pi_1(p_1, p_2, r) p_1 c + \pi_2(p_1, p_2, r) p_2 c \leq B_0 + F(p_1, p_2, r). \tag{18}$$

The left-hand side of Equation (18) denotes the total costs for auditing selected nursing homes which includes both honest and inflating nursing homes. The inflators audited will be fined, and the collected fines will be added to CMS's initial budget, B_0, as shown on the right-hand side of Equation (18). For simplicity, we assume the fines are immediately collected and can be used toward auditing more nursing homes within the same year.

4.3 Simulation Results

Neither the initial budget, nor the audit efficiency can be easily formulated in a linear form. In view of this, we conduct a simulation to summarize useful insights on parameter settings rather than solving the problem analytically.

Consider a one-year setting in which CMS's audit policy is announced at the beginning of the year. Given CMS's audit policy, inflating nursing homes react differently according to their expected payoffs; some choose to inflate measures in two domains, some choose to inflate measures only in one domain, and some choose not to inflate. By the end of the year, CMS conducts the audit by randomly selecting p_1 percent of the nursing homes that their overall rating increases by one star, and p_2 percent of the nursing homes that their overall rating increases by two stars. The caught inflators are fined according to the punishment rate r. CMS keeps the total cost of the audits under the initial budget B_0 plus the collected fines. We set the simulation parameters according to the 2013 data, which is the latest available year in our dataset. Following Han et al. (2018), we first identify the likely inflators in the nursing home population. We then exhaust all possible combinations of $p_1,\ p_2$, and r with step size 0.01, and calculate the net budget B_0 and the audit efficiency.

The analysis of the relationship between the initial audit budget B_0, the probability of audits, $p_1,\ p_2$, and the rate of penalties, r, is important, albeit complicated. On the one hand, increasing the audit probabilities and the penalties results in more inflators being caught, and more funds being collected and therefore lowers the level of required initial budget. On the other hand, it deters some of the nursing homes from inflating, which in turn reduces the number of inflators being caught and the penalties being collected, thus the level of required initial budget may increase.

4.3.1 Minimizing the Audit Budget. It is of significant importance for CMS to determine the optimal parameter settings under which it can completely deter inflation while minimizing the initial audit budget. According to Equation (17), nursing homes will stop inflating when the expected payoff of inflation equals 0. The conditions are expressed in Equations (19) and (20).

$$\Delta prof_1(1-p_1) - \Delta prof_1 p_1 r = 0, \tag{19}$$

$$\Delta prof_2(1-p_2) - \Delta prof_2 p_2 r = 0. \tag{20}$$

Solving Equations (19) and (20), we can obtain the marginal probability for deterring inflation:

$$p_1 = p_2 = \frac{1}{1+r}. \tag{21}$$

Equation (21) is consistent with Equations (15) and (16), and defines the minimum audit probabilities, $p_1,\ p_2$ for a given penalty rate, r, to absolutely deter inflation and ensure that no nursing home has any incentive to inflate. Equation (21) also indicates that when p_1, p_2 are set higher than $\frac{1}{1+r}$, CMS's resources will be *wasted* on auditing honest nursing homes that report true improvements, as no nursing home will have any incentive to inflate under such audit policies. Table 11 lists the simulation results of minimum audit probabilities p_1 and p_2 to completely deter inflation for a given penalty rate r, and the corresponding initial budget B_0. The minimum levels of p_1 and p_2 obtained from the simulation are equal, which is consistent with Equation (21). When punishment rate r increases, the corresponding minimum p_1 and p_2 required to absolutely deter inflation

Table 11. Optimal Policy Parameters for Deterring Inflation

R	0.1	0.3	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9
p_1	0.91	0.77	0.67	0.59	0.53	0.48	0.44	0.4	0.38	0.35
p_2	0.91	0.77	0.67	0.59	0.53	0.48	0.44	0.4	0.38	0.35
B_0	340.34	287.98	250.58	220.66	198.22	179.52	164.56	149.6	142.12	130.9

decrease. This result shows a tradeoff between audit probabilities and punishment rate to achieve the same outcome. The results in Table 11 also show that when the penalty rate r, increases, the corresponding net budget B_0 decreases. The reason is that when inflation is completely deterred, CMS collects no fine from auditing, and has to rely solely on net budget B_0 to fund the audit process. In this situation, the smaller p_1 and p_2 lead to fewer audits and smaller net budget. The results also indicate that though increasing audit probabilities and penalty rate can both deter inflation, the way they function is different. While increasing audit probabilities increases the audit workload and its required budget, increasing the punishment rate does not affect the audit budget and can therefore be a more economical way for CMS to deter inflation.

4.3.2 Maximizing the Audit Efficiency. Completely deterring rating inflation, though desirable, may not be feasible, since the audit probability is limited by the initial budget, and the maximum punishment rate can be restricted by law. While a certain level of inflation may be inevitable, CMS can set audit parameters under the given net budget in order to maximize its audit efficiency. Since audits impose unnecessary burdens on honest nursing homes, CMS should focus its resources to audit *more* inflators and *fewer* honest nursing homes. Following this idea, we formulate the *audit efficiency*. We define the Percentage of Caught Inflators (*PCI*) to be the ratio between caught inflators and the total inflators in the system, and the Percentage of Audited Honest nursing homes (*PAH*) to be the ratio between audited honest nursing homes and the total honest nursing homes in the system. By definition, both *PCI* and *PAH* are between 0 and 1. The audit efficiency curve, defined as the corresponding *PCI* given a certain *PAH*, denoted as *PCI*(*PAH*), can then be plotted in a 1×1 square area, where the y-axis represents the *PCI*, and x-axis represents the *PAH*. Note that a given *PAH* can be achieved by multiple combinations of p_1, p_2, and r, resulting in different *PCI*s.

To study the properties of the audit efficiency, we focus on the upper and lower limits of *PCI*, as the maximum and the minimum of audit efficiency. In the year 2013, the overall star ratings of 496 nursing homes increased as a result of their self-reported measures, of which 122 are identified as likely inflators (Han et al. 2018). We set up our simulation based on these statistics. The following propositions are derived.

Proposition 1:

The maximum audit efficiency is achieved at $r = 0$, and does not change with respect to B_0, i.e., $\forall B_{0x}, B_{0y} \in \{0, R^+\}, B_{0x} \neq B_{0y}, \max_{p_1,p_2,r} PCI(PAH,\ B_{0x}) = \max_{p_1,p_2,r} PCI(PAH, B_{0y}) = \max_{p_1,p_2,0} PCI(PAH)$.

Proposition 2:

(2.1) Given p_2 and r, the maximum audit efficiency converges to p_2 monotonically when $p_1 \rightarrow 1$.

(2.2) Given p_1 and r, the maximum audit efficiency converges to p_1 when $p_2 \rightarrow 1$.

(2.3) If $\Delta prof2 \cdot p_2 - \Delta prof1 \cdot p_1 < 0$, or $\frac{p_1}{p_2} > \frac{\Delta prof2}{\Delta prof1} > 1$, then the maximum audit efficiency does not change with respect to r. If $\frac{p_1}{p_2} < \frac{\Delta prof2}{\Delta prof1}$, then the maximum audit efficiency depends on the values of p_1, p_2, $\Delta prof1$ and $\Delta prof2$, and is piecewise linear.

The proofs of the propositions are provided in the Appendix.

We justify the propositions using simulation results. Figure 1(a) shows the maximum audit efficiency that CMS can achieve given different levels of net budget B_0. On the contrast, Figure 1(b) shows the lowest audit efficiency CMS can achieve given the same levels of the net budget B_0.

As shown in Figure 1, the maximum audit efficiency is achieved at $r = 0$ and increasing B_0 will not further increase the audit efficiency. When $r = 0$, inflating nursing homes will face no consequence if they are caught, thus all the non-honest nursing homes that may inflate, will choose to do so. As a result, the inflator proportion in the suspect group reaches the maximum, and the audit efficiency also reaches the maximum at the given audit probabilities. The result is consistent with what is proved in Proposition 1. In this case, increasing B_0 will not result in detecting more inflators, and can only put more honest nursing homes under audit. On the other hand, the lowest audit efficiency will decrease even more when B_0 increases. This is due to the fact that for the same percentage of caught inflators (*PCI*), more honest nursing homes will be audited and thus *PAH* will also increase. From Figure 1, we also know that the maximum audit efficiency curve is piecewise linear and monotonically non-decreasing. These properties of the maximum audit efficiency curve are also consistent with the results derived in Proposition 2.

4.3.3 Profit Gain Estimate. The implementation of our model relies on the estimates of inflators' profit gain after self-reporting, i.e., $\Delta prof1$ and $\Delta prof2$. It can be challenging to evaluate these statistics for each individual nursing home. However, from CMS's perspective, it is possible to estimate the profit gain for each rating group, and it is not uncommon to see government agencies using average statistics when the actual number is difficult to be calculated. For example, IRS uses average statistics to determine tax liabilities (IRS 2016).

In our analysis, we collect nursing homes' financial data from Office of Statewide Health Planning and Development (OSHPD) for the year 2009–2013. The data contains detailed information about the financial situation of each nursing home, including revenues, expenses and profits from each payer category, e.g., Medicare, Medicaid, Self-paying, etc. By combining the OSHPD financial data with CMS rating data, and following the method presented in Han et al. (2018), we are able to calculate the average daily per patient profit for each rating group, as shown in Table 12. These statistics are used in our simulation.

According to the current rules for calculating overall rating, the average profit gain for nursing homes in each rating group can be calculated, as shown in Table 12. For example, for a three-star nursing home with an average profit 10.79, if it inflates one domain to obtain four stars in overall rating, its expected profit gain is $\Delta prof1 = 2.812 = 13.602 - 10.790$. If the three-star nursing home inflates two domains, then its expected profit gain is $\Delta prof2 = 9.011 = 19.801 - 10.790$. For nursing homes in other rating groups, the profit gain can be calculated in a similar way. For four-star nursing homes, the highest rating they can achieve after inflation is five stars, thus $\Delta prof1 = \Delta prof2 = 6.199$. For five-star nursing homes, their star ratings cannot be improved through inflation, thus $\Delta prof1 = \Delta prof2 = 0$. For one-star nursing homes, according to the rule, their overall ratings cannot exceed two stars after self-reporting, thus $\Delta prof1 = \Delta prof2 = 0.822$.

5 CONCLUSION

CMS's Nursing Home Compare is the most widely used and respected evaluation system for nursing homes. Despite its importance, the system has been shown to suffer from exaggerated and inflated self-reported measures which significantly undermine the reliability of its ratings. This happens for two reasons: first, CMS does not have the budget to directly inspect all three domains and has to rely on the self-reported measures in two domains. Second, CMS does not have any audit system to detect and deter inflation of the self-reported measures. In this research, we address both of these issues.

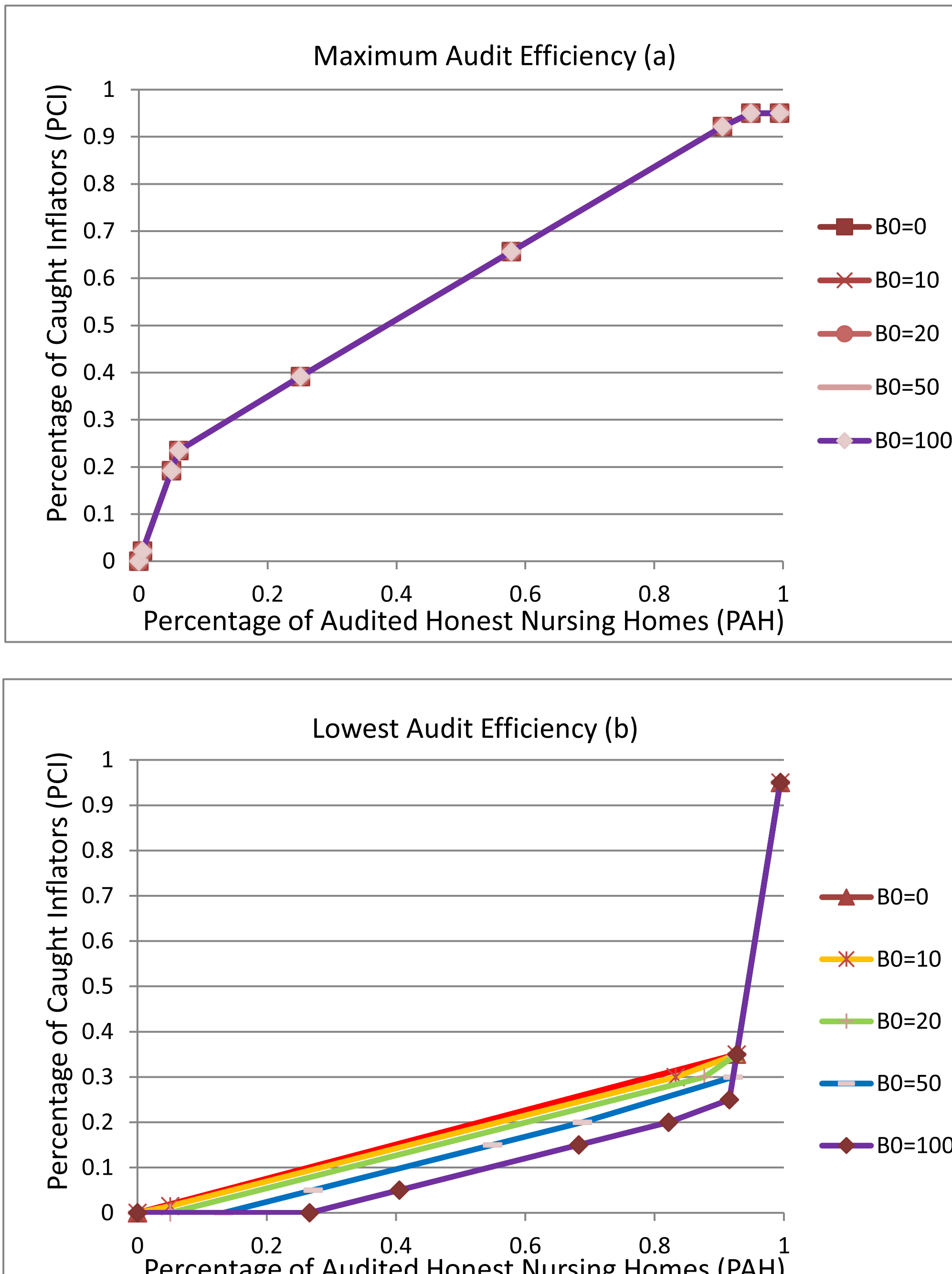

Fig. 1. Efficiency curves given different net budget B_0.

Table 12. Estimated Profit Gain for Each Rating Group

Star Rating	Expected profit	$\Delta prof1$	$\Delta prof2$
5	19.801	0	0
4	13.602	6.199	6.199
3	10.790	2.812	9.011
2	10.108	0.682	3.494
1	9.286	0.822	0.822

We first examine if the domain that CMS currently inspects is the best choice in terms of reducing the population of nursing homes that can inflate and reducing the difficulty of detecting inflators. To do this, we develop a general formulation with different objective functions and solve a particular case that is close to CMS's current practice based on publicly available data on nursing home ratings. We show that CMS's current choice of inspection domain is not optimal if it intends to minimize the population of nursing homes that can inflate their reports, and CMS will be better off if it inspects the staffing domain instead. We also show that CMS's current choice of inspection domain is only optimal had there been an audit system in place to complement it. This is due to the fact that the current inspection strategy of CMS minimizes the difficulty of detecting the inflating nursing homes, yet it does not have any audit system to utilize this potential.

We then design an auditing system for CMS which will be coupled with its current inspection strategy with two objective functions: to minimize the initial budget required to conduct the audits and to maximize the efficiency of the audit in terms of the ratio of audited inflators to audited honest nursing homes. To design the audit system, we first derive analytical conditions on the parameter settings, and then conduct a simulation to justify the results.

Given the limitations of datasets and the assumptions, this research may have the following limitations: When optimizing the inspection strategy, we found it very difficult to accurately measure the cost of inspecting each domain. Our result is obtained based on the assumption that the inspection cost for all domains is equal while in practice the inspection cost for each of the three domains may be different as it is driven by many factors such as local regulations and differences in difficulty of inspecting measures within each domain. However, we keep our problem setting close to the current practice as much as possible, thus our results are reasonable estimates of the current system and can provide a theoretical framework for optimizing the inspection structure.

When designing the audit system, we assume that the fines can be collected immediately to fund more audits within the same year. In view of the fact that the process of collecting fines is long and time-consuming, all fines may not be available within the same year. As a result, the quantitative results we obtained will be different in this case, but most qualitative results will still hold. In view that no complete audit system is currently adopted by CMS, our design provides a reasonable plan and can serve as a benchmark for other audit system designs.

Our results not only provide insights and theoretical support for CMS's current rating system but also provide guidelines for the future audit mechanism design. Moreover, this research has important managerial application for other rating systems that share similar features. A good example is the Merit-Based Incentive Payment System (MIPS), which is used by CMS to rate the physicians based on a combination of inspected and self-reported measures. With the fast development of healthcare IT solutions, self-reported measures can have a wide application in different rating systems. Our research provides a good framework to systematically control the quality of self-reported measures, which guarantees the accuracy and reliability of subsequent ratings.

APPENDIX A

Proof for Proposition 1:

We first divide the nursing home population into four different types, as shown in Table A.1.

The PAH is calculated as

$$PAH = \frac{\pi_{HI2}\, p_2 + \pi_{HI1}\, p_1}{\pi_{HI2} + \pi_{HI1}}$$

Note that PAH is a linear combination of p_1 and p_2, since π_{HI1} and π_{HI2} are constant.

The PCI is calculated as

$$PCI = \frac{\pi_{LI2}(p_1,\ p_2,\ r) \cdot p_2 + \pi_{LI1}\ (p_1,\ p_2,\ r) \cdot p_1}{\pi_{LI2}(p_1,\ p_2,\ r) + \pi_{LI1}(p_1,\ p_2,\ r)} = \frac{1}{\pi_{LI}}(\pi_{LI2}(p_1,\ p_2,\ r) \cdot p_2 + \pi_{LI1}\ (p_1,\ p_2,\ r) \cdot p_1)$$

where $\pi_{LI} = \pi_{LI2}(p_1,\ p_2,\ r) + \pi_{LI1}\ (p_1,\ p_2,\ r)$ denotes the total number of inflators in the system. The first term in the parenthesis denotes the caught inflators from the 2-star suspect group, and the second term denotes the caught inflators from the 1-star group.

Since $\pi_{LI2}(p_1,\ p_2,\ r)$ and $\pi_{LI1}\ (p_1,\ p_2,\ r)$ are both non-increasing in r, the linear combination $\pi_{LI2}(p_1,\ p_2,\ r) \cdot p_2 + \pi_{LI1}(p_1,\ p_2,\ r) \cdot p_1$ is maximized at $r = 0$. For higher B_0, higher punishment rate r can be used in the audit, however, the audit efficiency is maximized at $r = 0$, and higher net budget B_0 does not result in higher the audit efficiency, i.e., $\exists B_{0x}, B_{0y} \in \{0, R^+\}, B_{0x} \neq B_{0y}, \max_{p_1,p_2,r} TPR(FPR,\ B_{0x}) = \max_{p_1,p_2,r} TPR(FPR, B_{0y}) = \max_{p_1,p_2,0} TPR(FPR)$

Proof for Proposition 2:

(2.1) The nursing homes having intention to inflate make their decisions based on the expected payoffs of the following three cases:

0. Not Inflate. Π_0=0
1. Inflate 1 star. Π_1=Δprof1 (1-p_1-$p_1 r$)
2. Inflate 2 stars. Π_2=Δprof2 (1-p_2-$p_2 r$)

In our problem, we have $\Delta prof1 \leq \Delta prof2$ for all rating levels.

Given r and p_2, and set $p_1 = 0$, $\Pi_2 = \Delta prof2(1 - p_2 - p_2 r)$, denote the profit when $p_1 = 0$ as $\Pi_{10} = \Delta prof1$

a) If $\Pi_{10} \leq \Pi_2$, $PCI_0 = p_2$. When p_1 increases, Π_1 decreases, thus $\Pi_1 \leq \Pi_2$ holds, and PCI = p_2

b) If $\Pi_{10} > \Pi_2$, $PCI_0 = p_1$. When p_1 increases, Π_1 decreases.

Table A.1. Nursing Home (NH) Population Partition

		True Service Quality	
		H	**L**
Star Increase	**I**	**π_{HI}: NHs improve star rating legitimately, including π_{HI1} and π_{HI2}** π_{HI2} – Honest NHs reporting 2-star improvement π_{HI1} – Honest NHs reporting 1-star improvement	**π_{LI}: Inflators, including π_{LI1} and π_{LI2}** π_{LI2} – Inflators inflating 1 star[a] π_{LI1} – Inflators inflating 2 stars[b]
	NI	**π_{HNI}: Established NHs**	**(π_{LI0}): Potential inflators who choose not to inflate[c]** **π_{LNI}: "Abandoned" NHs**

a. Nursing homes choosing to inflate 2 stars, if any, are confounded with the population π_{HI2}.

b. Nursing homes choosing to inflate 1 star, if any, are confounded with the population π_{HI1}.

c. Nursing homes with the intention to inflate but decide not to inflate due to unfavorable expected payoff. They are confounded with other honest nursing homes and are also not the focus of audit.

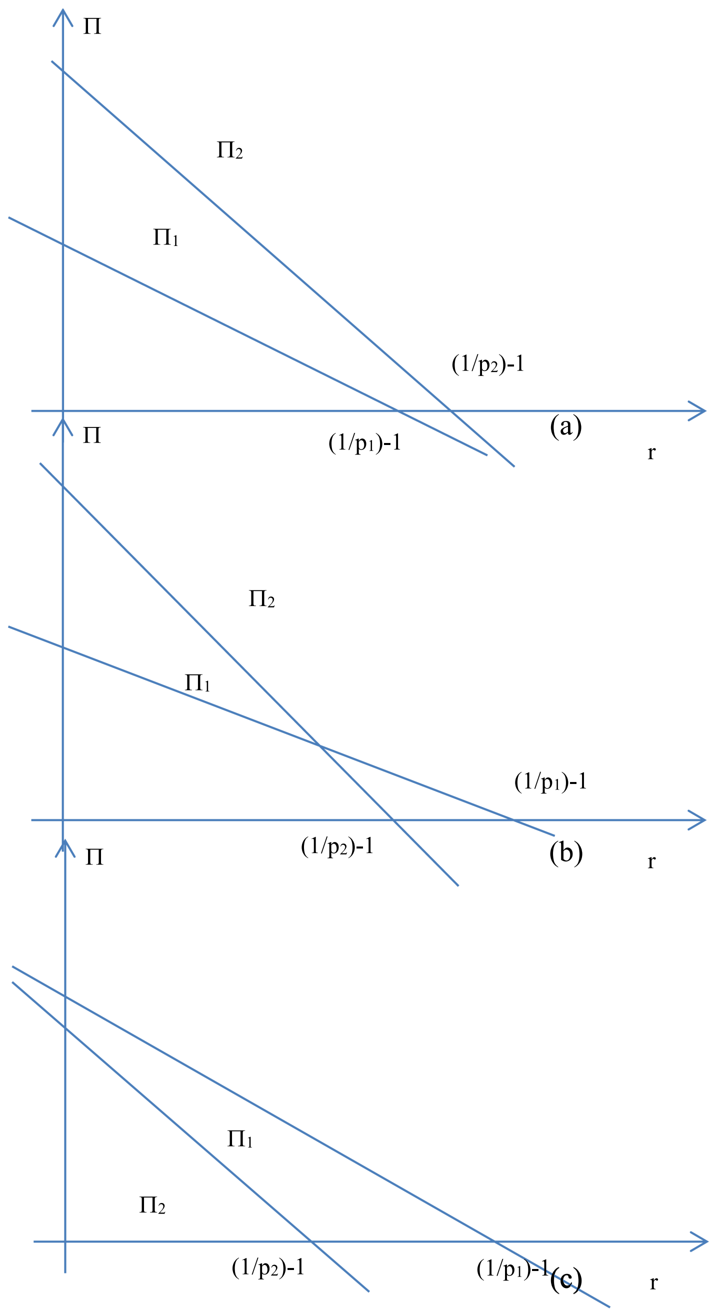

Fig. A.1. Three cases of the profit function.

The breakpoints is $\Pi_2 = \Delta prof2(1 - p_2 - p_2 r) = \Delta prof1(1 - p_1 - p_1 r) = \Pi_1$, i.e.,

$$p_1* = \frac{\Delta prof1 - \Delta prof2 + \Delta prof2(1+r)p_2}{\Delta prof1(1+r)}$$

When $p_1 > p_1*$, we have $\Pi_1 < \Pi_2$, and PCI= p_2. In other words,

$$\text{PCI} = p_1, \quad \text{when } p_1 \leq p_1*$$

$$p_2, \quad \text{when } p_1 > p_1*$$

Thus, if $p_1* < p_2$ holds, then PCI is monotonically non-decreasing in p_1

$$p_1 * - p_2 = \frac{\Delta prof1 - \Delta prof2 + \Delta prof2(1+r)p_2 - \Delta prof1(1+r)p_2}{\Delta prof1(1+r)}$$

$$= \frac{\Delta prof1 - \Delta prof2}{\Delta prof1(1+r)}[1-(1+r)p_2]$$

Since $\frac{\Delta prof1 - \Delta prof2}{\Delta prof1(1+r)} \leq 0$, and $1-(1+r)p_2 > 0$ when Π_2>0, thus $p_1 * - p_2 \leq 0$, i.e., $p_1* < p_2$ always holds when Π_2 >0. In other words, PCI is monotonically non-decreasing in p_1.

(2.2) Given r and p_1, and set $p_2 = 0$, $\Pi_1 = \Delta prof1(1 - p_1 - p_1 r)$, $\Pi_{20} = \Delta prof2 > \Pi_1$, and $\text{PCI}_0 = p_2$. When p_2 increases, Π_2 decreases.

The breakpoints are $\Pi_2 = \Delta prof2(1 - p_2 - p_2 r) = \Delta prof1(1 - p_1 - p_1 r) = \Pi_1$, i.e.,

$$p_2* = \frac{\Delta prof2 - \Delta prof1 + \Delta prof1(1+r)p_1}{\Delta prof2(1+r)}$$

When $p_2 > p_2*$, we have $\Pi_1 > \Pi_2$, and PCI= p_1. In other words,

$$\text{PCI} = p_2, \quad \text{when } p_2 \leq p_2*$$

$$p_1, \quad \text{when } p_2 > p_2*$$

Thus, if $p_2* < p_1$, then PCI is monotonically non-decreasing in p_2

$$p_2 * - p_1 = \frac{\Delta prof2 - \Delta prof1 + \Delta prof1(1+r)p_1 - \Delta prof2(1+r)p_1}{\Delta prof2(1+r)}$$

$$= \frac{\Delta prof2 - \Delta prof1}{\Delta prof2(1+r)} = [1-(1+r)p_1]$$

Since $\frac{\Delta prof2 - \Delta prof1}{\Delta prof2(1+r)} \geq 0$, and $1-(1+r)p_1 > 0$ when Π_1>0, thus $p_2 * - p_1 \geq 0$, i.e., $p_2* > p_1$, as a result, when p_2 increases, PCI increases to p_2^*>p_1 first, then drops back to p_1 and stay at p_1. PCI is NOT monotonic in p_2.

(2.3) Given p_1 and p_2, set $r_0 = 0$, thus $\Pi_{20} = \Delta prof2(1 - p_2)$, $\Pi_{10} = \Delta prof1(1 - p_1)$,

If $\Pi_{20} > \Pi_{10}$, we have

$$\Delta prof2(1 - p_2) > \Delta prof1(1 - p_1)$$

$$\Delta prof2 - \Delta prof1 > \Delta prof2p_2 - \Delta prof1p_1$$

Since $\Delta prof2 - \Delta prof1 \geq 0$, if $\Delta prof2 \cdot p_2 - \Delta prof1 \cdot p_1 < 0$, then $\Pi_{20} > \Pi_{10}$ holds.

When r increases, if $\Pi_2 > \Pi_1$ still holds, then PCI does not change. In other words,

$$\Delta prof2(1 - p_2 - p_2 r) > \Delta prof1(1 - p_1 - p_1 r)$$

Or

$$\Delta prof2 - \Delta prof1 > (\Delta prof2p_2 - \Delta prof1p_1)(1+r)$$

Case 1. If $\Delta prof2p_2 - \Delta prof1p_1 < 0$, when r increases, the right-hand side decreases, thus the above equation always holds. In this case, $\Pi_2 > \Pi_1$ holds when r increases. PCI = $p2$, and will not change.

Case 2. If $\Delta prof2p_2 - \Delta prof1p_1 > 0$, three subcases can be discussed.

a) If $p_1 > p_2$, then $\Pi_{20} > \Pi_{10}$, the payoff functions are shown in Figure A.1(a).
In this case, $\Pi_2 > \Pi_1$ holds for $\Pi > 0$; thus PCI $= p_2$, and will not change.
b) If $p_1 < p_2$, then if $\Pi_{20} > \Pi_{10}$, PCI $= p_2$, the payoff functions are shown in Figure A.1(b).
In this case, if $r < \frac{\Delta prof2 - \Delta prof1}{\Delta prof2 p_2 - \Delta prof1 p_1} - 1$, then $\Pi_1 < \Pi_2$, and PCI $= p_2$.
If $r > \frac{\Delta prof2 - \Delta prof1}{\Delta prof2 p_2 - \Delta prof1 p_1} - 1$, then $\Pi_1 > \Pi_2$, and PCI $= p_1$.
c) If $p_1 < p_2$, then when $\Pi_{20} < \Pi_{10}$, the payoff functions are shown in Figure A.1(c).
In this case, $\Pi_1 > \Pi_2$ always holds, PCI $= p_1$, and will not change.

REFERENCES

Ahmed Abbasi, Zhu Zhang, David Zimbra, Hsinchun Chen, and Jay F. Nunamaker, Jr. 2010. Detecting fake websites: The contribution of statistical learning theory. *MIS Quarterly* (2010), 435–461.

A. Abbasi, C. C. Albrecht, A. Vance, and J. V. Hansen. 2012. MetaFraud: A meta-learning framework for detecting financial fraud. *MIS Quarterly* 36, 4 (2012), 1293–1327.

A. Abbasi, F. M. Zahedi, D. Zeng, Y. Chen, H. Chen, and J. F. Nunamaker, Jr. 2015. Enhancing predictive analytics for anti-phishing by exploiting website genre information. *Journal of Management Information Systems* 31, 4 (2015), 109–157.

Xue Bai, Ram Gopal, Manuel Nunez, and Dmitry Zhdanov. 2012. On the prevention of fraud and privacy exposure in process information flow. *INFORMS Journal on Computing* 24, 3 (2012), 416–432.

Jeffrey R. Casterella, Kevan L. Jensen, and W. Robert Knechel. 2009. Is self-regulated peer review effective at signaling audit quality? *The Accounting Review* 84, 3 (2009), 713–735.

Richard J. Cebula. 2012. Impact of income-detection technology and other factors on aggregate income tax evasion: The case of the United States. *PSL Quarterly Review* 54, 219 (2012).

M. Cecchini, H. Aytug, G. J. Koehler, and P. Pathak. 2010. Detecting management fraud in public companies. *Management Science* 56, 7 (2010), 1146–1160.

Centers for Medicare and Medicaid Services (CMS). 2015. Design for Nursing Home Compare Five-Star Quality Rating System: Technical Users' Guide. https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/Downloads/usersguide.pdf.

Centers for Medicare and Medicaid Services (CMS). Changes to Nursing Home Compare and the Five Star Quality Rating System. http://www.medicareadvocacy.org/changes-to-nursing-home-compare-and-the-five-star-quality-rating-system/.

Chrysanthos Dellarocas. 2000. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In *Proceedings of the 2nd ACM Conference on Electronic Commerce*, 150–157.

Stefano DellaVigna and Eliana La Ferrara. 2010. Detecting illegal arms trade. *American Economic Journal: Economic Policy* 2, 4 (2010), 26–57.

Mark Duggan and Steven D. Levitt. 2000. Winning isn't everything: Corruption in sumo wrestling. *National Bureau of Economic Research* w7798 (2000).

Donald G. Fowles. 2012. *A Profile of Older Americans*. Administration on Aging. US Department of Health and Human Services.

Lana Friesen. 2003. Targeting enforcement to improve compliance with environmental regulations. *Journal of Environmental Economics and Management* 46, 1 (2003), 72–85.

Scott M. Gilpatric, Christian A. Vossler, and Michael McKee. 2011. Regulatory enforcement with competitive endogenous audit mechanisms. *The RAND Journal of Economics* 42, 2 (2011), 292–312.

Xu Han, Yaraghi Niam, and Ram Gopal. 2018. Winning at all costs: Analysis of inflation in nursing homes' rating system. *Production and Operations Management* 27, 2 (2018), 215–233.

Jon D. Harford. 1987. Self-reporting of pollution and the firm's behavior under imperfectly enforceable regulations. *Journal of Environmental Economics and Management* 14, 3 (1987), 293–303.

Eric Helland. 1998. The enforcement of pollution control laws: Inspections, violations, and self-reporting. *Review of Economics and Statistics* 80, 1 (1998), 141–153.

Gilles Hilary and Clive Lennox. 2005. The credibility of self-regulation: Evidence from the accounting profession's peer review program. *Journal of Accounting and Economics* 40, 1–3 (2005), 211–229.

IBM. 2015. IBM ILOG CPLEX Optimization Studio Getting Started with the IDE. https://www.ibm.com/support/knowledgecenter/SSSA5P_12.6.3/ilog.odms.studio.help/pdf/gsoplide.pdf, (2015).

IRS. 2016. Information About Your Notice, Penalty and Interest. https://www.irs.gov/pub/irs-pdf/n746.pdf, (2016).

Robert Innes. 1999. Remediation and self-reporting in optimal law enforcement. *Journal of Public Economics* 72, 3 (1999), 379–393.

Brian Jacob and Steven Levitt. 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *National Bureau of Economic Research* (2003) w9413.

N. Jindal and B. Liu. 2008. Opinion spam and analysis. In *Proceedings of 1st ACM International Conference on Web Search and Data Mining*. Stanford, California, USA.

A. Josang and R. Ismail. 2002. The beta reputation system. In *Proceedings of the 15th Bled Electronic Commerce Conference.*

Kaiser Family Foundation (KFF). 2017. Distribution of Medicaid Spending on Long Term Care. http://kff.org/medicaid/state-indicator/spending-on-long-term-care/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D.

Louis Kaplow and Steven Shavell. 1994. Optimal law enforcement with self-reporting of behavior. *Journal of Political Economy* 102, 3 (1994), 583–606.

Raymond Y. K. Lau, S. Y. Liao, Ron Chi-Wai Kwok, Kaiquan Xu, Yunqing Xia, and Yuefeng Li. 2011. Text mining and probabilistic language modeling for online review spam detection. *ACM Transactions on Management Information Systems (TMIS)* 2, 4 (2011), 25.

Shing-Han Li, Yu-Cheng Kao, Zong-Cyuan Zhang, Ying-Ping Chuang, and David C. Yen. 2015. A network behavior-based botnet detection mechanism using PSO and K-means. *ACM Transactions on Management Information Systems (TMIS)* 6, 1 (2015), 3.

E. Lim, V. Nguyen, N. Jindal, B. Liu, and H. Lauw. 2010. Detecting product review spammers using rating behaviors. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*. Toronto, Canada.

Y. Liu and Y. L. Sun. 2010. Anomaly detection in feedback-based reputation systems through temporal and correlation analysis. In *Proceedings of the 2nd IEEE International Conference on Social Computing*.

S. Liu, J. Zhang, C. Miao, Y. Theng, and A. Kot. 2014. An integrated clustering-based approach to filtering unfair multinominal testimonies. *Computational Intelligence* 30 (2014).

John Livernois and Chris J. McKenna. 1999. Truth or consequences: Enforcing pollution standards with self-reporting. *Journal of Public Economics* 71, 3 (1999), 415–440.

X. Ma, L. Khansa, Y. Deng, and S. Kim. 2013. Impact of prior reviews on the subsequent review process in reputation systems. *Journal of Management Information Systems* 30, 3 (2013), 279–310.

Ines Macho-Stadler and David Perez-Castrillo. 2006. Optimal enforcement policy and firms' emissions and compliance with environmental taxes. *Journal of Environmental Economics and Management* 51, 1 (2006), 110–131.

Arun S. Malik. 1993. Self-reporting and the design of policies for regulating stochastic pollution. *Journal of Environmental Economics and Management* 24, 3 (1993), 241–257.

Mohammad M. Masud, Tahseen M. Al-Khateeb, Kevin W. Hamlen, Jing Gao, Latifur Khan, Jiawei Han, and Bhavani Thuraisingham. 2011. Cloud-based malware detection for evolving data streams. *ACM Transactions on Management Information Systems (TMIS)* 2, 3 (2011), 16.

Dina Mayzlin, Yaniv Dover, and Judith Chevalier. 2012. Promotional reviews: An empirical investigation of online review manipulation. *National Bureau of Economic Research* w18340.

A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh. 2013. Spotting opinion spammers using behavioral footprints. In *Proceedings of SIGKDD International Conference on Knowledge Discovery and Data Mining*. Chicago, USA.

Andreas Marcel Oestreich. 2015. Firms' emissions and self-reporting under competitive audit mechanisms. *Environmental and Resource Economics* 62, 4 (2015), 949–978.

Mark Raymond. 1999. Enforcement leverage when penalties are restricted: A reconsideration under asymmetric information. *Journal of Public Economics* 73, 2 (1999), 289–295.

Sarah L. Stafford. 2008. Self-policing in a targeted enforcement regime. *Southern Economic Journal* (2008), 934–951.

W. Teacy, J. Patel, N. Jennings, and M. Luck. 2006. Travos: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems* 12, 2 (2006).

K. Thomas. 2014. Medicare star ratings allow nursing homes to game the system. *The New York Times* (2014).

G. Wang, S. Xie, B. Liu, and P. S. Yu. 2011. Identify online store review spammers via social review graph. *ACM Transactions on Intelligent Systems and Technology* (2011).

R. T. Wright and K. Marett. 2010. The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. *Journal of Management Information Systems* 27, 1 (2010), 273–303.

Keli Xiao, Qi Liu, Chuanren Liu, and Hui Xiong. 2018. Price shock detection with an influence-based model of social attention. *ACM Transactions on Management Information Systems (TMIS)* 9, 1 (2018), 2.

He Zhang, Sanjay Mehotra, David Liebovitz, Carl A. Gunter, and Bradley Malin. 2013. Mining deviations from patient care pathways via electronic medical record system audits. *ACM Transactions on Management Information Systems (TMIS)* 4, 4 (2013), 17.

Received May 2018; revised November 2018; accepted April 2019