8-4-2017

Rating Game – Inflation Detection, Audit Design, and Technology Adoption on the Nursing Home Compare System

Xu Han
University of Connecticut - Storrs, xu.han@business.uconn.edu

Rating Game – Inflation Detection, Audit Design, and Technology Adoption on the Nursing Home Compare System

Xu Han, PhD

University of Connecticut, 2017

Abstract

The Nursing Home Compare System supported by the Centers for Medicare & Medicaid Services (CMS) is being widely used by patients, medical providers and payers. However, evidence suggests that the rating system is prone to self-reporting inflation, leading to biased and misleading ratings. This dissertation consists of three essays analyzing a series of issues that arise in this rating system, including inflation detection, performance evaluation, audit design, and technology adoption.

In the first essay, we use data over 2009-2013 for 1219 California nursing homes to empirically examine the key factors affecting a nursing home's rating. We find a significant association between a nursing home's rating change and its profits, and then demonstrate this association does not always lead to legitimate efforts to improve service quality, but can induce self-reporting inflation. A prediction model is then developed to evaluate the extensiveness of inflation based on which 6 to 8.5% of the nursing homes are identified as likely inflators.

Given limited CMS resources, it is important to optimize the inspection process and develop an effective audit process to control inflation. In the second essay, we first formulate the inspection problem by using an innovative graph-based method, and solve the problem based on CMS data. The results support CMS's current practice in term of minimizing inflation detection difficulty, and suggest an audit system. We then conduct a detailed simulation study on the optimal audit parameter settings. Our result suggests a moderate audit policy to balance the tradeoff between audit net budget and efficiency.

IoT technologies enable automatic data collection, which can release nursing homes from self-reporting burden and reduce the possibility of misreporting. However, IoT technologies can be costly,

and CMS may consider subsidizing IoT adoption to control inflation. In the third essay, we develop a two-level game theoretical framework to study how IoT adoption can affect nursing homes' operational decisions, and how CMS should subsidize IoT adoption. We analyze reactions of honest and inflating nursing homes to IoT adoption, and analyze how CMS can control IoT adoption by auditing and subsidization. We also obtain insights on budget allocation between subsidization and audit.

Rating Game – Inflation Detection, Audit Design, and Technology Adoption on the Nursing Home Compare System

Xu Han

B.S., University of Science and Technology Beijing, 2005

M.S., University of Science and Technology Beijing (Ctrl. Theory & Ctrl. Engineering), 2008

M.S., University of Connecticut (Electrical & Computer Engineering), 2012

M.S., University of Connecticut (Statistics), 2017

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2017

APPROVAL PAGE

Doctor of Philosophy Dissertation

Rating Game – Inflation Detection, Audit Design, and Technology Adoption on the Nursing Home Compare System

Presented by

Xu Han

Major Advisor ____

Ram Gopal

Associate Advisor ____

James Marsden

Associate Advisor ____

Robert Day

University of Connecticut

2017

ACKNOWLEDGEMENTS

I would like to first express my sincere grateful attitude to all the members on my committee for their invaluable contributions: Dr. James Marsden, Dr. Robert Day, and especially my major advisor, Dr. Ram Gopal. I started working with Ram closely since the beginning of this project. His brilliant research ideas opened my mind and greatly motivated me. I appreciate his patience, understanding during those tough times in my research, and I cannot achieve all of this without his guidance. Dr. James Marsden has offered me a great amount of help in my research, my qualifying paper, my job market preparation, etc. With his help, I was able to shape my mind to be an IS researcher. Dr. Robert Day also gives me a lot of helps and advices, not only in research, but also in teaching. I was his TA in the Operations Management class. He gave me a lot of suggestions on class preparation, presentation, and class management. Following his advices, I recast my teaching, and improved it significantly.

I would like to express my special appreciation to my external advisor Dr. Niam Yaraghi for his tremendous contribution to this project. He is very knowledgeable in Healthcare IS research, and his input sparks a lot of great ideas in this project. We won the 2015-2016 NIHCM Research Grant together, and had our research results introduced to the public in the C-SPAN live broadcasting. I cannot imagine achieving this without his contribution.

I would like to thank all the OPIM Ph.D. students for the great days we shared together. Dr. Lei Wang, Dr. Gang Wang and Dr. Mark Schneider helped me a lot since I started in the program, and have set great role models for me to follow. Brian Lee, Yuan Jin, Aindrilla Chakraborty, Yucheng Chen, Hongfei Li, Hoda Atef Yekta, Mohsen Emadikhiav, Arda Zuber,

Teng Huang and Guneet Kaur worked with me together through out these years. I'm very glad to be one member in the OPIM big family, with all of you as my brothers and sisters.

I also want to thank Gail Zaicek and Melissa Burk for all the help throughout the years. Our department thrive on their continuous support and effort, and we all benefit from it.

Finally, my deep and sincere gratitude to my family for their continuous help and unparalleled love. I am forever indebted to my parents for supporting me and encouraging me unconditionally. My beloved wife Jinzi is always there supporting me through the good times and bad. Their love made me who I am. My journey would not have been successful if not for them, and I dedicate this milestone to them!

Table of Contents

Chapter 1. Introduction

Ratings are commonly used to compare and evaluate alternative choices. The rating mechanisms vary across different application domains: some ratings, such as the vehicle safety ratings, are generated through a rigorous inspection procedure or expert evaluation. Some other ratings are obtained based on customer reviews, such as Amazon product ratings. Another type of rating is created based on self-reported measures, such as MBA rankings. There are also rating systems in which ratings are calculated based on a combination of the above methods. The Nursing Home Compare supported by Centers for Medicare and Medicaid Services (CMS) which is based on a combination of inspection results and nursing homes' self-reported measures is an example of these rating systems,

There are over 16000 nursing homes in the United States currently. They provide care to over 1.5 million residents (Jones et al. 2009) (Donald and Fowles 2012). In year 2012 only, CMS spent $140 billion on long-term services and supports (Eiken et al. 2014). Given the importance of nursing homes in the quality of life of patients and the billions of dollars spent on these facilities, in 2008, CMS launched its nursing home rating system, which has been widely used by patients, doctors, and insurance companies since its inception (Duhigg 2007).

The system gives a star rating in a 1-5 scale for each nursing home based on three domains of *Health Inspection*, *Staffing* and *Quality Measures(QM)*. The Health Inspection is conducted annually by licensed inspectors, while the other two measures are self-reported by nursing homes. Once the inspection and self-reported data are collected, CMS will assign a star rating to each of the three domains. The overall rating is then calculated by using the Health Inspection rating as a baseline, adding one star if any self-reported domain is 5 stars and subtracting one star if any self-reported domain is 1 star.

The presence of self-reported measures in the overall rating calculating procedure renders the possibility of inflation. Cases have been reported where patients' personal experience differs significantly from what the star ratings suggest. Some highly rated nursing homes are providing substandard cares,

even leading to the death of patients. Despite the importance of these issues, there have been very limited research on these topics. Does rating inflation exist? If so, how extensive is the inflation? What are the characteristics of those inflators? How to effectively conduct inspection? How to design an audit system in order to detect and deter inflation? How can information technologies can be used to improve the rating system? In this dissertation, we will address these questions in detail.

In Chapter 2, we first conduct an empirical analysis to demonstrate the existence of rating inflation. We collect the CMS rating data over 2009-2013 and the corresponding financial data reported by Office of Statewide Health Planning and Development and patients' complaints data reported by California Department of Public Health for 1219 nursing homes in California. To understand the importance of high ratings to nursing homes, we empirically examine the key factors affecting the star rating change of a nursing home. We find a significant association between the changes in a nursing home's star rating and its profits, which points to a financial incentive for nursing homes to inflate the ratings. It is however still possible that the star rating increase comes from nursing homes' legitimate efforts to improve their service qualities. To demonstrate rating inflation does exist, we use the independent patients' complaint data as a proxy of the nursing homes' true service quality, and demonstrate that the association between rating change and financial incentive is beyond what can be explained by legitimate efforts. A prediction model is then developed to evaluate the extensiveness of inflation. The results suggest that among the suspect population, 6 to 8.5% of the nursing homes are likely inflators. We then conduct a variable importance analysis and summarize the key characteristics of likely inflators, which provide useful insights for CMS to conduct future audit. In this chapter, we are able to empirically demonstrate the existence of rating inflation in CMS's nursing home rating system, and provide a quantifiable estimate of the system performance. To the best of our knowledge, this is the first study to do so. We also reveal the underlying driving force of rating inflation, and show the key features of likely inflators, which provides useful information for CMS's future audit design and regulation.

In chapter 3, we discuss the methods to optimize and improve the rating system. The self-reporting in the rating procedure brings in possibility of inflation. If all the three domains are inspected, there will be no room for nursing homes to inflate. However, due to limited resources, CMS can only inspect one domain. Therefore, it is important to know which domain is the optimal choice for inspection. To address this question, we develop an innovative graph-based method to formulate the problem, and solve it with different objective functions based on CMS's historical data. The results of our analysis show that the staffing measure is optimal if CMS wants to minimize the population of nursing homes who can inflate, but the measure that CMS is currently inspecting is optimal in term of minimizing the difficulty to detect inflators, and will work well together with an effective audit system. Unfortunately, CMS currently does not have an audit system for the rating system. When designing the audit system, we consider nursing homes' reactions to different audit policies, and conduct a simulation to study the optimal audit parameter settings. Our results show a tradeoff between the audit net budget and audit efficiency. Increasing punishment rate is an economical way for CMS to save net budget, but will lead to more resources wasted on auditing honest nursing homes. CMS should use a moderate audit policy in order to carefully balance the tradeoff between audit net budget and audit efficiency.

In recent years, Internet of Things (IoT) technologies are becoming popular among nursing homes. IoT devices are being used to track patients' movements and health conditions, and the data is automatically collected and communicated through network, which improves the productivity and efficiency of nursing staff, and the overall care quality. IoT can change the way a nursing home operates, and thus affect the nursing home's optimal operational decisions. From CMS's perspective, the self-reporting procedure imposes heavy burden on nursing homes' operations. IoT can release nursing homes from this burden, and can reduce the possibility of misreporting, and thus can be a good method to control self-reporting inflation. Despite these advantages, IoT technologies can be costly for nursing homes to adopt, and thus CMS may consider subsidizing nursing homes for IoT adoption. In chapter 4, we develop a game theoretical framework to study how IoT adoption can affect nursing homes optimal staffing

decisions, and how CMS should subsidize IoT adoption as an alternative way to control self-reporting inflation. We develop a two-level problem structure, with CMS optimizing audit policies and IoT subsidization amount to minimize self-reporting inflation damage at the higher level, and nursing homes optimizing IoT adoption and staffing level to maximize profits at the lower level. We analyze the different reactions of honest and inflating nursing homes to IoT adoption, and find that inflators are more reluctant to adopt IoT since it limits the amount they can inflate. For both inflators and honest nursing homes, we find diversified staffing level changes after IoT adoption due to the complementary and substitution effects of IoT adoption, i.e., for highly ranked nursing homes, IoT adoption decreases its optimal staffing level and for poorly ranked nursing homes, IoT adoption increases its optimal staffing level. We also study how CMS can affect nursing homes' IoT adoption and in turn control self-reporting inflation by auditing and subsidization. At last, we provide insights for CMS to allocate limited budget between subsidization and auditing. The model presented in this chapter suggests a new direction that CMS may want to consider in its rating system reform. By subsidizing IoT adoption, CMS's incentive to control self-reporting inflation can be aligned with nursing homes' incentives to maximize profit, and the overall service quality of nursing homes can also be improved at the same time. Comparing with wasting resource auditing honest nursing homes, allocating limited budgets on IoT subsidization under certain conditions can be CMS's optimal choice.

We conclude the thesis in chapter 5. CMS's nursing home rating system comes a long way to evolve to today's form, and the reforming never stops. It is extremely important for CMS to fully understand the issues in the current system, such that the reforms can move towards the correct and most effective direction. This dissertation provides insights on various aspects of the current rating system. The results can be very helpful and serves as a guideline for CMS to improve the current rating system, and allocate limited budget more efficiently.

Chapter 2. Winning at All Costs: Analysis of Inflation in Nursing Homes' Rating System

2.1 Introduction

Nearly two million Americans spend an average of 835 days of their life in one of the 15,700 nursing home facilities in the United States (National Center for Health Statistics 2009). The Department of Health and Human Services estimates that in 2009, 4.1% of Americans over 65 years old lived in these facilities. This percentage increases with age, ranging from 1.1 percent in the population of 65 to 74 years old to 13.2% in the population older than 85 (Fowles 2012). In 2012 only, Medicaid spent $140 billion on long-term services and supports (Eiken et al. 2014). Despite the importance of nursing homes in the quality of life of millions of Americans and the billions of dollars spent on them, very little information has been available about their service quality. The Centers for Medicare & Medicaid Services (CMS) designed and implemented its nursing home rating system after a congressional hearing in 2007 where Senator Ron Wyden asked "*why it was easier to shop for washing machines than it is to select a nursing home*" (Duhigg, 2007). Given the lack of alternative information resources on nursing homes, the publicly available CMS rating has become the gold standard in the industry since its inception, and has been widely popular among patients, physicians and payers (Thomas 2014). The recent study of Werner et al. (2016) sheds light on the importance of CMS ratings for nursing homes; according to their analysis after the release of the ratings the market share of 1-star facilities decreased by eight percent while the market share of 5-star facilities increased by more than six percent.

Given the important role of CMS's nursing home rating system, nursing homes would have a significant incentive to improve their ratings. However, these ratings may not always reflect true quality. Cases have been reported in which highly rated nursing homes only provide sub-standard care, even lead to the death of patients. It is possible that the rating system is prone to inflation by nursing homes, and

the objective of this research is to investigate the existence and the extent of inflation in the CMS's nursing home rating system.

This chapter is based on the publicly available data provided by multiple government agencies including CMS, California Statewide Health Planning and Development (OSHPD), and California Department of Public Health (CDPH). Our empirical strategy consists of four steps as discussed below. First, we explore the financial incentives for nursing homes to improve their star ratings using a combination of CMS rating data and OSHPD financial data. We find a significant positive association between the change in star ratings and the financial incentives. That is, nursing homes with higher financial incentives are more likely to improve star ratings after self-reporting. Second, to prove the existence of rating inflation, we initially analyze the correlation between the CMS inspection and nursing homes' self-reported results. If the self-reported improvement is legitimate, we expect it to be reflected in the inspection results of the subsequent period. We also expect CMS inspection rating and self-reported ratings within the same year to be closely associated. Our correlation analysis results, however, shows almost no correlation between the inspection and self-reported results, and sheds doubt on the legitimacy of self-reported measures. We then further corroborate the results of our correlation analysis by examining additional data on patient complaints provided by CDPH: if we assume that the ratings are not inflated, then we should observe similar service qualities among the nursing homes with similar *overall* ratings. Moreover, we should observe increased service quality among the nursing homes that initially had the same *inspection* rating but ended up with a higher *overall* rating as a result of their high self-reported measures. Our results, however, show significant differences between the service qualities of the nursing homes with the same *overall* rating. Moreover, no significant difference exists in the service quality of nursing homes with the same *inspection* rating. The result serves as strong evidence on the existence of inflation in the current rating system as it points to the fact that the service quality is predicted by the health inspection ratings which cannot be inflated, rather than the overall ratings which can be inflated. Third, to estimate the extent of rating inflation, we develop a prediction model and apply

it to estimates the proportion of nursing homes that have inflated their self-reported ratings. By using a 95% confidence interval, we identify around 6% of nursing homes in the suspect population to be likely inflators in the current system. Fourth, we conduct a variable importance analysis to classify the factors that their change contributes the most to the probability of being an inflator. Our results demonstrate the shortcomings of the current rating systems and call for significant reforms in how CMS and other payers evaluate the quality of nursing homes.

The chapter proceeds as follows. In Section 2.2, we discuss the background and conceptual framework of our research, including the history and evolvement of the rating system, the rating generation mechanism and potential issues, and propose the theoretical framework to deal with these issues. In Section 2.3, we review related literature on nursing home quality measures, misbehavior detection and quantifying methods. In Section 2.4, we describe our data collection procedure and explore the underlying financial incentives for nursing homes to improve their ratings. In Section 2.5, we first perform correlation analysis between the CMS-conducted inspection and self-reported measures, which cast doubt on the existence of rating inflation. We then demonstrate our conclusion by performing a more rigorous complaint-based analysis. A prediction model is developed in Section 2.6 to identify likely rating inflators and evaluate the performance of the system. A variable importance analysis is then conducted to show key characteristics of the inflators. We conclude the whole chapter in Section 2.7, and discuss the limitations and future work.

2.2. Background and Theoretical Framework

2.2.1 The History of Nursing Home Rating System

The standardization of nursing home service quality begins before CMS was founded. In 1961, the Public Health Service (PHS) began studying nursing homes' state licensures, after a series of problems being reported by the Commission on Chronic Illness from several states. The Nursing Home Standards Guide, issued by the Public Health Service (PHS), specified 77 service standards in health and safety, which established the foundation of nursing home service standards. From then on, this Nursing Home

Standards Guide gradually developed and more standards were included. By 1974, a total of 90 standards were included, covering various aspects in health and safety. In 1977, the Health Care Financing Administration (HCFA) was created as a new federal organization, and continued the standardization and certification of nursing home service qualities. The HCFA commissioned the Institute of Medicine (IOM) to examine the standards in nursing home services.

A major reform on nursing homes' regulation took place in 1987, when the Nursing Home Reform Act (OBRA-87) was passed. The OBRA-87 established more stringent inspection, and further specified and revised the regulations on nursing home services, including nurse training, care standards, sanctions and remedies. It also established the use of Resident Assessment Instrument, of which the Minimum Data Set is a major component, and is widely used today in nursing home research.

The Health Care Financing Administration (HCFA) changed its name to the Centers for Medicare and Medicaid Services (CMS) in 2001. CMS released its Nursing Home Compare (NHC) system in October 1998, in the form of report card, which provides information on Medicare/Medicaid certified nursing homes via internet. The initial system only includes nursing homes' basic information and the deficiencies on health and safety found in inspection, which are also covered in today's Health Inspection. The Staffing measure was included in the system in June 2000, and the Quality Measures were included in November 2002 (General Accounting Office [GAO], 2002). This is the early form of today's 3-measure nursing home rating system. The NHC report card system is influential since it is one of the earliest systems presenting publicly available standardized quality information on nursing homes. However, the report card method suffers challenges such as the lack of consumer awareness and access (Stevenson 2006), and the difficulties for consumers to understand the information on the report card (Shugarman & Brown 2006).

In order to address these issues, CMS launched its Nursing Home Compare system in December 2008, which is the current system being used. This reformed rating system followed the 3-measure setting in the previous report card system, but uses a 5-star scale on each of the three measures, which

greatly improved the usability of the rating system. The 5-star nursing home rating system gradually becomes the gold standard of nursing home selection. As reported by CMS (CMS 2015), the system gets more than 1.4 million visitors per year, with 85% users reporting that they found the information they are looking for on nursing homes. The dataset used in our research covers nursing homes' rating, complaint, and financial data from 2009 to 2013, which are the first 5 years since the inception of the 5-star rating system.

Starting from February 2015, CMS has gradually announced new policies to improve its nursing homes rating system (Medicare 2016). These policies include expanding the targeted surveys, adding two additional measures in the quality measure domain, revising the staffing algorithm, etc. However, the framework of the 5-star rating system is not changed. By the end of 2016, CMS requires all nursing homes to report payroll-based staffing information (CMS 2015).

2.2.2 The Current Rating Mechanism

The CMS rating system is based on three domains: *Health Inspection*, *Staffing*, and *Quality measures*. While independent, CMS-certified inspectors conduct and report the health inspections, the other two domains are self-reported by nursing homes. CMS first assigns an initial star rating to all nursing homes based on their annual *health inspection* results. The *health inspection* looks into areas such as medication management, nursing home administration, environment, food service, and residents' rights and quality of life. Ratings are given based on the number, scope, and severity of deficiencies identified during the three most recent annual inspections (1/2 for current year, 1/3 for the previous year, and 1/6 for the second prior year). According to CMS's rating mechanism design, the top 10% nursing homes in health inspection receive 5 stars, while the bottom 20% nursing homes receive 1 star. Nursing homes which rank in between receive 2-4 stars according to a fixed proportion. There is no such restriction for self-reported measures. Nursing homes are then assigned star ratings for the Staffing and Quality Measures domains. The Staffing domain is evaluated based on the self-reported CMS Certification and Survey Provider Enhanced Reports (CASPER) staffing data. Staffing domain shows the average staffing level per resident

day of a nursing home throughout a year. The two measures covered under staffing domain are the Registered Nursing (RN) hours per resident day, and the total nursing hours, which is the sum of Registered Nurse(RN) hours, Licensed Practical Nurses(LPN) hours and nurse aide hours per resident day. The results are adjusted for case-mix based on the Resource Utility Group (RUG-III) case-mix system derived from the Minimum Data Set (MDS). The staffing star rating is then updated by the end of the quarter when raw data is collected. The Quality Measure domain rating uses 9 out of 18 quality measurement criteria developed from the MDS, which covers 7 aspects from long-stay terms and 2 aspects from short-stay terms. The quality measure data is collected by the end of each quarter and the quality measure star rating is updated by using the results from three most recent quarters.

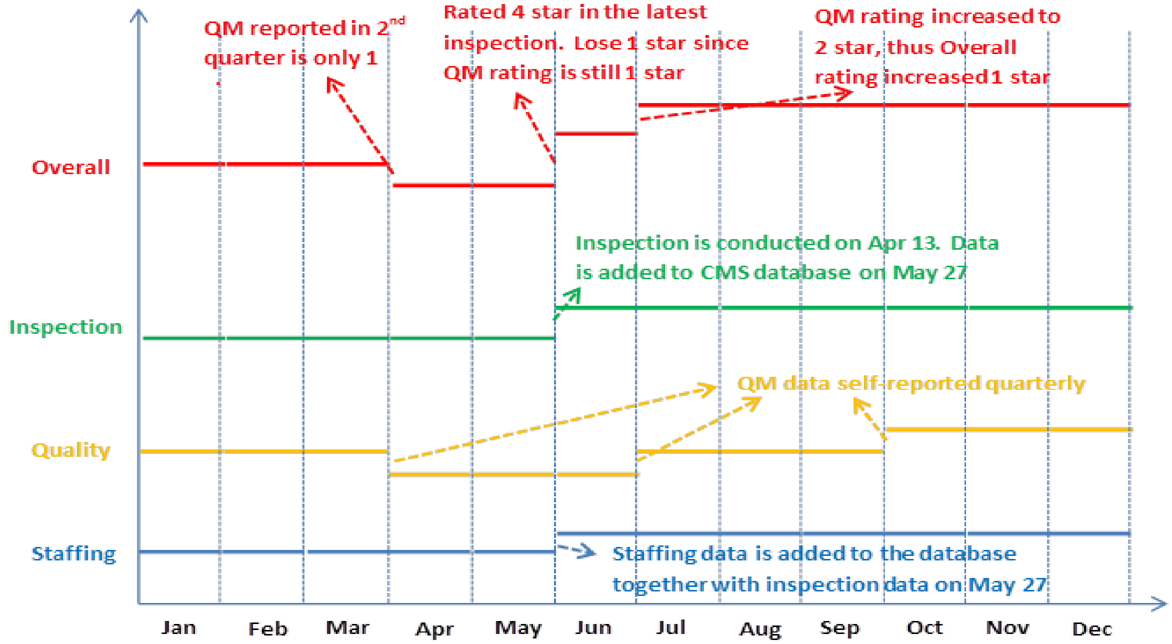
The overall star rating is then calculated by considering the health inspection rating as the baseline, adding 1 star if any self-reported domain is 5 stars and subtracting 1 star if any self-reported domain is 1 star. Nursing homes who only got 1 star in the health inspection can only have one additional star after self-reporting[1]. The overall star rating cannot be more than 5 stars or less than 1 star. An example is provided in Table 2-1 and Figure 2-1 to demonstrate the rating dynamics and the corresponding events for a randomly selected nursing home in 2009.

[1] Additional conditions apply to nursing homes which are in the CMS's Special Focus Facility (SFF) program.

Table 2-1. An Example of a Nursing Home's Rating Dynamics

Month	*Overall*	*Inspection*	*Quality Measurement*	*Staffing*
January	2	3	2	1
February	2	3	2	1
March	2	3	2	1
April	1	3	1	1
May	1	3	1	1
June	3	4	1	2
July	4	4	2	2
August	4	4	2	2
September	4	4	2	2
October	4	4	3	2
November	4	4	3	2
December	4	4	3	2

Figure 2-1. The Graphical Representation of a Nursing Home's Rating Dynamics



Note: [a] In the first quarter of 2009, the nursing home received 3 stars in inspection. It reports 2 stars in quality measures and 1 star in staffing. The resulting overall rating is 2 stars.

[b] In April, the reported quality measure reduces to 1 star, with the other two domains unchanged. As a result, the overall rating reduces to 1 star.

[c] In June, a new inspection is conducted, in which the nursing home receives 4 stars. The staffing data is also reported together with the inspection in June to be 2 stars. The resulting overall rating is 3 stars.

[d] In July, the quality measures are newly reported to be 2 stars. With the other domains unchanged, the overall rating increases to 4 stars, since none of the self-reported domains are 1 star.

[e] In October, the quality measures are newly reported to be 3 stars. This change, however, does not affect the overall rating.

The detailed items covered in each measure are listed in Table 2-2. They measure nursing homes' service qualities from three different angles. Generally speaking, the measures covered under health inspection reflect how organized the nursing facility is operating; The staffing measures cover the number of working professionals in the facility; The quality measures reflect how healthy the patients are living in the facility. Though measuring from different perspectives, there exist close connections among these measures (Munroe et al. 1990; Harrington et al. 2000&2012; Konetzka et al. 2004; Zhang et al. 2004; Castle et al. 2008; Kim et al. 2009). For example, urinary tract infection (UTI) is a common health problem found among nursing home patients, and the percentage of UTI is an important measure under the QM domain. Research has shown that UTI is closely related to catheter insertion (Gokula et al. 2004), which requires frequent and timely care, and as a result, an adequate level of staffing coverage. It has also been shown that the improper use of anti-biotic agent is one of the major reasons causing UTI, and the anti-biotic agent misusage is covered in the pharmacy service deficiencies, which are under the health inspection domain. As a result, UTI associated problems are reflected in all the three domains. A similar example can also be found for pressure ulcer associated problems. As a result, we argue that the three measures, though measuring from different angles, should be correlated at certain level. An unexpected low correlation is suspicious, and can be a preliminary evidence of misreporting.

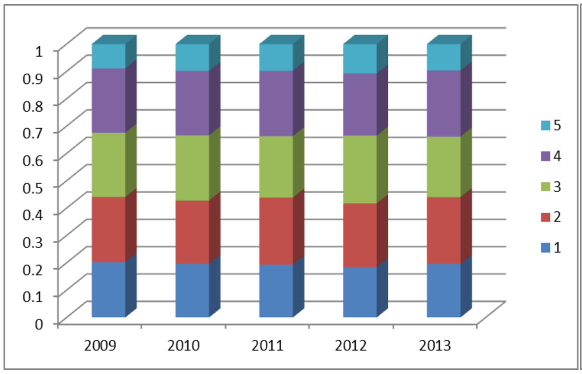
Table 2-2. Coverage of Each Measure (Health Inspection, Staffing, and Quality Measures)

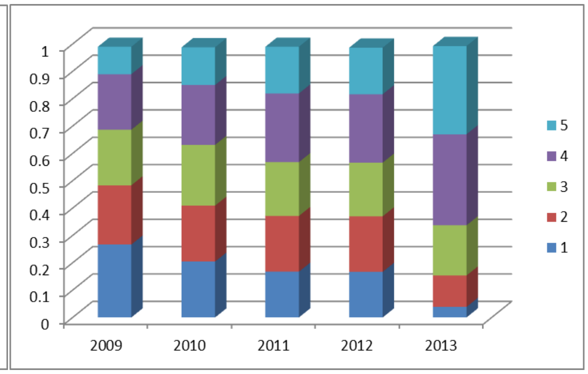
Health Inspection (H: Health; F: Fire Safety)	*Staffing*	*Quality Measures (L: long-star; S: Short-stay)*
Count of Administration Deficiencies (H)	RN hours/day	Percent of residents whose need for help with activities of daily living has increased (L)
Count of Environmental Deficiencies (**H**)	LPN hours/day	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**)		

2.2.3 Potential Issues of the Current Rating Mechanism

The two self-reported domains can fundamentally change a nursing home's overall rating. For example, it is possible for an average nursing home that has received 3 stars in the health inspection to gain two additional stars based on self-reported measures and become an excellent 5-star nursing home. As a result, the overall rating can be quite different from the health inspection rating. Figure 2-2 shows how the ratings in each of these measures have shifted to higher stars during a period of five years from 2009 to 2013. By design, the proportions of health inspection star rating remain unchanged in the 5 years, as shown in Figure 2-2(a). However, the number of nursing homes that claim high performance in the self-reported domains has continuously increased over the past five years. As shown in Figure 2-2(b), in 2009, about 40% of nursing homes self-reported to be 4 or 5 stars in the quality measures domain. This percentage has increased to 60% in 2013. On the other hand, about 20% of nursing homes self-reported to be 1 star in 2009, but less than 10% of nursing homes self-reported to be 1 star in 2013. In the staffing domain, the number of highly rated nursing homes also significantly increased over this period, as shown in Figure 2-2(c). Consequently, the overall rating is consistently skewed to the higher end over time. As shown in Figure 2-2(d), the portion of 4 or 5 nursing homes increased from 35% to 55% over the 5 years.

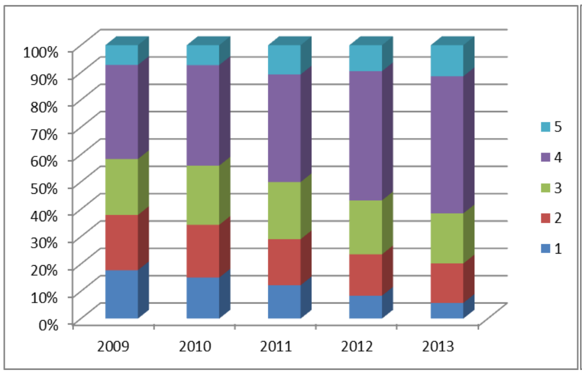
Figure 2-2. Distribution of Nursing Home Ratings from 2009 to 2013*

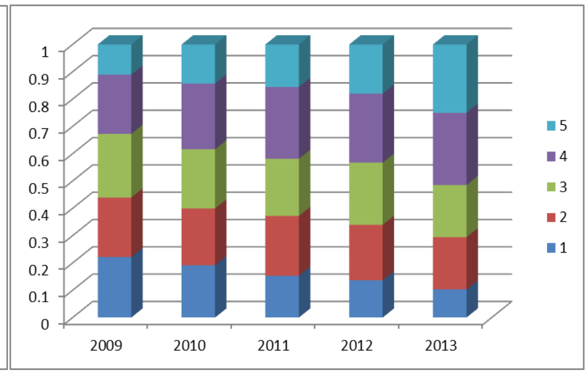




(a) On-site inspection

(b) Quality Measure





(c) Staffing

(d) Overall rating

Note: Colors represent different star rating groups

The trend we observe in Figure 2-2 can be interpreted in two ways: On the one hand, supporters can argue that increased levels of self-reported measures are genuine and represent an honest effort by nursing homes to constantly improve their services. On the other hand, however, skeptics may argue that the improved ratings are not legitimate but are rather a result of nursing homes' success in developing strategies to manipulate the system and inflate their ratings. Cases have been reported in which patients' experiences differ significantly from the star ratings. Some highly-rated nursing homes are sued for substandard care, even causing death of patients due to improper medical treatments (Thomas 2014).

Since late 2014, CMS has gradually announced new policies to improve nursing homes' rating system (Medicare 2016). These policies include the expansion of targeted surveys, including additional measures in the quality measure domain, and adding the payroll information to the staffing reports. Despite these amendments, the structure of the rating system has not been changed and it still heavily relies on the self-reported domains and thus the newly revised system continues to be prone to manipulation by false self-reported measures. It is not clear whether the rating increase is a result of nursing homes' legitimate efforts to improve their services or a signal of rating inflation, and the objective of this research is to answer this question by investigating the existence and the extent of inflation in this rating system.

2.2.4 Proposed Methodology and Theoretical Framework

Rating system may use different mechanisms to generate ratings but the rating data generally comes from three sources: authority inspection (e.g., vehicle safety ratings), customer reporting (e.g., Amazon ratings) or self-reporting (e.g., business school rankings). In the CMS's nursing home, ratings are generated by combining authority inspection and self-reported data in a unique way. When justifying such ratings, it is desirable to bring in information from the customer reporting. In our rating inflation detection method, we use the number of patient complaints as a proxy of the true service quality (Carman, et al. 1990; Dabholkar et al. 1995; Tsaur et al. 2002). Out logic is as follows:

1. If the ratings are not inflated, then for nursing homes with similar overall ratings, we should expect similar service qualities, reflected in a similar number of complaints.

2. If there is no inflation, we expect increased service quality for the nursing homes whose star rating increased after self-reporting, comparing with nursing homes who initially had the same inspection rating but did not increase after self-reporting.

Our results, however, do not support any of the above inference. We observe a clear difference in the number of complaints for nursing homes with the same overall rating, indicating that their service qualities are quite different. We also observe no significant difference for nursing homes whose star rating increased after self-reporting, indicating that their reported improvements are highly questionable. The combined results indicate that the self-reported star rating increase cannot be simply explained by legitimate efforts, and rating inflation does exist.

To give a quantifiable estimate of the inflation, we incorporate ideas from the decomposition model developed by Oaxaca (Oaxaca 1973; Fairlie 2005; Bauer and Mathias 2008), which has been commonly used for quantifying group differences. Specifically, it decomposes the total difference between the groups into two parts: the differences caused by the differences in individual characteristics, and the differences caused by inconsistency in the measures. The model we developed is in line with the Oaxaca's idea. We divided the nursing homes into honest ones and potential inflators. We obtain the unbiased coefficient of the honest nursing homes and use the coefficients to predict the star ratings of the potential inflators. By doing this, we systematically control for differences caused by individual characteristics. A maximum predicted rating is then calculated for each of the potential inflators by using selected confidence intervals, and is compared with the observed rating. If the observed rating is higher than the maximum predicted rating, then significant inconsistency exists in the measures, which points to the inflation of self-reported measures. By running this prediction model, we can identify likely inflators in the system, and give a quantifiable system performance evaluation.

2.3. Literature Review

The rating inflation problem is an important topic related to many inter-connected fields, including healthcare facility operations management, healthcare policy research, and misbehavior detection. In this section, we first review literature in each related field, and then discuss the contribution of our work to the existing literature.

2.3.1 Healthcare Facility Operations Management

The research on the operations management of healthcare facilities includes an abundance of scholarly work, and research topics can be categorized based on the size of the facility. The first stream of research analyzes the efficiency of operations and quality of care at hospitals. This line research includes improving the patient scheduling systems (Cayirli and Veral 2003; Helm et al. 2011) and developing strategies to address the demand fluctuations (Jack and Powers 2004), analysis of the effects of patients' arrival time (D. Anderson et al. 2014) or the hospitals objectives (Andritsos and Aflaki 2015) on the quality of care and creating alternative operations planning and control systems for curbing the increasing costs of hospital services (Roth and Van Dierdonck 1995). The other line of operations research focuses on individual physicians and small clinics. This stream includes design of public policies and novel scheduling strategies to reduce waiting time (Chen et al. 2015) and increase clinic performance (LaGanga and Lawrence 2012; Salzarulo et al. 2015), optimization of capacity and resources allocations and the effect of such improvements on quality of medical services in both primary (Dobson et al. 2011; McCoy and Johnson 2014; Zepeda and Sinha 2016) and specialty care settings (Chow et al. 2011; Güneş et al. 2015).

Though nursing homes are an important part of the U.S. health care system, operations and production management literature often neglects them. To the best of our knowledge is limited to a few studies on minimizing waiting times (Zhang et al. 2012) and analyzing the effects of non-profit status of nursing homes on their service quality (Chesteen et al. 2005).

2.3.2 Healthcare Policy Research

The health policy literature related to nursing homes is rich, and we summarize them into three major categories. The first category of studies tries to answer the question how service quality can be quantifiably measured. Berg et al. (2002) evaluated existing quality indicators for long-term cares. Mor et al. (2003) used the MDS to point out that the incident-based nursing home quality measures can be unstable. Shwartz et al. (2015) discussed the importance of using composite measures to measure healthcare provider performance. The second category of research mainly focuses on how to improve nursing home service qualities. Kane et al. (2001) compared the senior patients and patients with disabilities to show the key needs for senior people in long-term cares. Kieran (2001) discussed how improper regulation can potentially detract from its effectiveness and lead to disappointing results. Grabowski et al. (2014) discussed how telemedicine can reduce hospitalizations for nursing home residents. Stavropoulou et al. (2015) examined the function of incident-reporting system in improving patient safety. Mor et al. (2010) pointed out that the CMS payment incentives do not encourage the incentive alignment of care providers and care beneficiaries. Many related studies on incentive alignment problems are also conducted (Rosalie 2003; Mor et al. 2004 R. Werner and Konetzka 2010; R. Werner, Stuart, and Polsky 2010). The third category of research discusses major problems existing in the current nursing home market, some of which are major barriers for achieving good service quality. These problems include racial segregation in nursing homes (Smith et al. 2007), public images distortion (Robert J. et al. 2006; Miller et al. 2012), payment policies and litigation issues (Stevenson and David 2008; Stevenson and David 2003; Smith et al. 2007; Fennell et al. 2010; Charlene et al. 2001; William et al. 2014; David, Angelelli, and Mor 2004; David et al. 2004).

In the above healthcare policy research, the ultimate goal is to understand how good services can be delivered to patients. The nursing home star rating system is CMS's attempt to implement the quality measures developed in the literature and convey the service quality information to the public in a transparent manner. The number of studies on this rating system is growing since its inception in 2009.

Li et al. (2013) studied the nursing home satisfaction rate in Massachusetts and found that incorporating consumer's perspective would improve the CMS nursing home reporting efforts. Konetzka et al. (2015) found that the rating system exacerbates disparities in quality by payer source. To the best of our knowledge, there has been no healthcare policy literature looking into rating inflation issues.

2.3.3 Misbehavior Detection

Rating inflation is a typical misbehavior that frequently occurs in system operations. The detailed method used in detecting each type of misbehavior can be different, but the common strategy is to first identify the abnormal phenomenon which cannot be rationalized should the misbehavior not exist, then explore the underlying incentive, usually financial-oriented, driving the phenomenon. Mayzlin, et al. (2012) found significant differences in reviews from a given hotel between Expedia and TripAdvisor. Since Expedia only allows its customers to post a review, its posting cost is significantly higher than TripAdvisor, where everyone can post. Consequently, competitors have the incentive to post fake reviews on the "free" TripAdvisor, but not on the "costly" Expedia, and the results gave a good explanation to the observed difference in the two websites' reviews. Duggan, et al. (2000) conducted a study on Japan's elite sumo wrestlers to detect statistical evidence of match rigging, and found that the winning ratio for players on the margin is significantly higher than players who are not. They showed that the incentive structure of promotion leads to gains from trade between wrestlers on the margin for achieving a winning record, and the observed higher winning ratio cannot be simply explained by legitimate effort. Jocob et al. (2003) studied teachers' cheating behavior using data from Chicago public schools. He found evidence indicating that high-powered incentive systems, especially those with bright line rules, may induce unexpected behavioral distortions such as cheating.

In the above studies, a measure of the abnormal phenomenon, such as review scores, winning ratio or consistent wrong answer patterns, can be easily accessible. However, due to the illicit nature, the people committing misbehaviors usually attempt not to leave evidence. As a result, sometimes a good measure of the abnormal phenomenon cannot be easily identified, and a good proxy variable is needed to

perform the analysis. DellaVigna, et al. (2010) proposed a method to detect illegal arms trade between countries under arms embargo by using the weapon manufacturers' stock prices as a proxy and analyzing their fluctuations as turmoil and conflicts arise at certain geographical areas. Engelberg, et al. (2014) used the geographic distance between a doctor's office and drug company headquarters to instrument for the likelihood of pecuniary transfers. They found evidence that doctors tilt prescriptions in favor of the paying firm's drugs, shifting away from both branded and generic substitutes.

2.3.4 Contribution of This Research

Our research makes contribution to the existing literature in several dimensions. For operation management literature, nursing homes have not been the research focus though they are very important in the U.S. healthcare system. The reasons for this neglect of research in nursing home operations can be complicated, but can be partially attribute to the lack of effective and objective performance measures. Different from hospitals and clinics, nursing homes' patients are residents at the same time, and a lot of efficiency-related performance measures for hospitals and clinics, such as waiting time, readmission rates, do not apply for nursing homes. Many nursing home problems, however, are the results of chronical misbehaviors in the daily care, which may not be objectively measured. Our research results provide a better understanding of the nursing home performance measures, and the potential inflation in the measures, thus fill up the gap in the nursing home operations management field.

For healthcare policy literature, the existing studies are based on the assumption that the reported data is truth-reflecting and the unbiased results can be delivered to the public. If inflation exists in the rating procedure, then no matter how complete the quality measures are developed or how effective the policies are set, they do not have a truth-reflecting and solid ground, and the results will be biased and misleading. Our research targets the authenticity of the reported data and the ratings directly, and provides a solid ground for other research which rely on these data.

Nursing home rating system inflation belongs to the type of problems in which the phenomenon, or the difference between honest nursing homes and inflators, can be difficult to identify, both cross-

sectional and longitudinal. The difficulties lie in the following aspects. First, the inflators are confounded with the honest nursing homes whose star ratings also increase after self-reporting, thus ratings cannot be directly used as a measure of inflation. Furthermore, there has been no audit system implemented for the self-reported measures, and there is no data for caught inflators available, which can be used to summarize unique characteristics of inflators. As a result, there is no training data for machine-learning based techniques, making it challenging to design detection methods. From the time dimension perspective, the self-reported measures have been used for years without being audited, thus the rating patterns, though probably inflated, can be very consistent over the years. The lack of external shock also makes it challenging to identify abnormal patterns in the rating data. To overcome these difficulties, we bring in the information from the patients' side, and use the number of complaints as a proxy variable of the true service quality. We then derive contradiction to show that the self-reported rating increase is beyond what can be explained by legitimate efforts. Theoretically, our research provides a framework for detecting rating system inflations: For any product or service to be rated, the ratings are generated from authority inspection, self-reporting, consumer reporting or their combinations. The three are correlated and can be good proxy variables for justifying others and detecting rating inflation.

Most of the existing research only focuses on proving the existence of misbehaviors. However, system reform often takes time, and it is always necessary to give a quantifiable evaluation on the current system's performance. To the best of our knowledge, few papers have addressed this issue before. Our research makes contribution to the existing literature by not only demonstrating the existence of rating inflation, but providing a systematical method to quantifiable estimate the extensiveness of rating inflation.

2.4. Data Collection and Financial Incentive Analysis

2.4.1 Data Collection

Our analysis is based on publicly available datasets from three sources: CMS, OSHPD and CDPH. The CMS dataset includes performance details on each of the criteria used within the three domains of inspection, staffing and quality measures. For each nursing home, these detailed metrics are accompanied with the corresponding star rating in the three domains as well as the overall star rating. This dataset also includes other descriptive details for nursing homes such as location, size, certification, ownership information and council type. The pooled dataset consists of records from 1219 nursing homes in the state of California over the first five years since the inception of the 5-star rating system, i.e., from year 2009 to 2013.

The OSHPD data includes detailed financial information on California nursing homes over the same period of time. In this dataset, nursing homes' source of revenue is categorized into healthcare and non-healthcare sections. The healthcare section is further classified by revenue source into Medicare, Medicaid[2], Self-paying, Managed Care and others. The corresponding revenue and expense details for each section are provided, and the profits can be easily calculated.

The CDPH data is provided through the Health Facilities Consumer Information System (HFCIS) website. A consumer portal is also available on the HFCIS website through which a complaint against a facility can be filed directly. CDPH inspects nursing homes at least once every 6 to 15.9 months in response to these complaints as well as other accidents or incidents that are required to report by nursing homes themselves, such as fires, disasters, suspected abuse, etc. Depending on the deficiencies found during the investigation, various types of citations will be issued. A deficiency violating state laws will be issued a state citation, and if it also violates federal law, it will also be reported to CMS and included in the federal inspection for determining star rating. The CDPH data we collected contains detailed

[2] In California, Medicaid is referred to as Medi-Cal. However, we use Medicaid as the category name in this chapter, in order to avoid confusion for readers from other states.

patient complaints, which will be used as a proxy of the nursing home's service quality. Note that the state level agency CDPH and the federal level agency CMS, though may overlap sometimes, have independent jurisdictions on nursing home inspections. The CDPH complaints may not be included in CMS's star rating procedure, and the deficiencies covered in CMS's inspection may not result from a CDPH complaint.

2.4.2 Nursing Homes' Financial Incentive

The observed rating improvement consists of both legitimate efforts and self-reported inflation. In order to demonstrate the existence of rating inflation we should show that the rating increase is beyond a range which can be explained by legitimate efforts. In our model, we perform the financial incentive analysis to establish the connection between a nursing home's financial incentive and the increase in its star rating. We then show that this increase is far beyond the limit which can be explained by legitimate efforts.

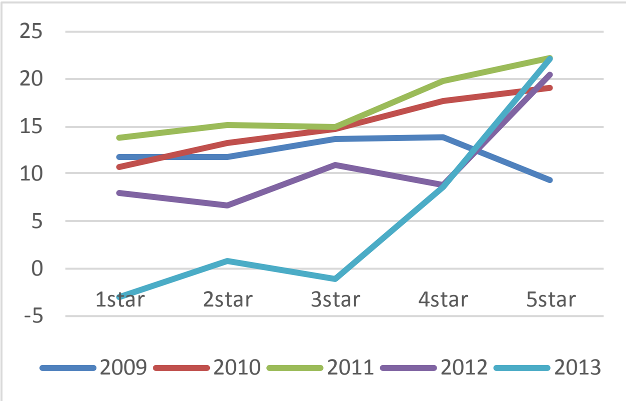
We combine the CMS rating data and OSHPD financial data to demonstrate the financial implications of star ratings for nursing homes. The combined data has 4433 records for California nursing homes over the 5 years. The average profit per day per patient is calculated for nursing homes in each overall rating group, as shown in Table 2-3. These averages serve as an estimate of the daily profit that a nursing home can expect per resident for the corresponding overall rating. The difference is significant. For example, a nursing home that receives 3 stars in health inspection may only expect a $10.79 profit from treating one patient for one day. However, if it gains two additional stars after self-reporting and achieves an overall rating of 5 stars, its expected profit increases to $19.8. Figure 2-3 shows the profit trend for each of the star rating group over the 5 years.

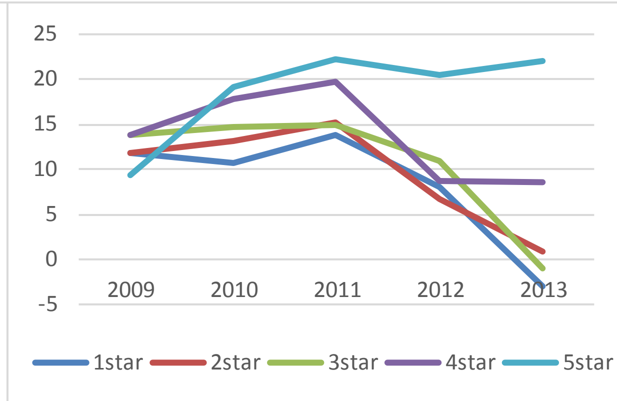
Table 2-3. Definition of Financial Incentive

Health inspection rating	*Expected profit* [a]	*Maximum possible overall rating*	*Maximum expected profit* [b]	*Financial Incentive* [c]
5	19.801	5	19.801	0
4	13.602	5	19.801	6.199 (Level 5– Level 4)
3	10.790	5	19.801	9.011 (Level 5– Level 3)
2	10.108	4	13.602	3.494 (Level 4– Level 2)
1	9.286	2	10.108	0.822 (Level 2– Level 1)

Notes:
[a] If inspection rating unchanged. The expected profit is the average per patient per day profit for the corresponding star rating group.
[b] If maximum possible overall rating realized.
[c] Difference between expected profit and expected loss.

Figure 2-3. Profit Trend over the Period of 2009-2013





(a) **Using Star Ratings as the Horizontal Axis**

(b) **Using Years as the Horizontal Axis**

Nursing homes' total net profits consist of healthcare part and non-healthcare part. The price of healthcare related services is regulated by CMS's Nursing Home Prospective Payment System (PPS) (CMS, Prospective Payment Systems - General Information, 2015) (CMS, Skilled Nursing Facility Prospective Payment System, 1997), which does not consider nursing homes' star ratings. As a result, highly rated nursing homes do not necessarily gain higher healthcare profits than the low-rating nursing homes. The non-healthcare related services, however, are not regulated by CMS. Such services include residential care services, unrestricted contributions, and interest income and gains from investments. Historical data shows that highly rated nursing homes can attract more patients who are in good financial conditions (typically self-paying and other resources). These patients are willing to pay more for good quality non-healthcare services. As a result, the non-healthcare profits for high-rating nursing homes can be significantly higher comparing with their low-rating counterparts. Moreover, the increased demand for services that happens as a result of high star ratings (R. M. Werner, Konetzka, and Polsky 2016) can reduce their overhead costs and thus lead to an increase in the net per-patient profit. The results demonstrate nursing homes' incentives to achieve the highest possible ratings from the financial perspective, and provide a quantifiable metric to measure such incentives. In our model, we define the financial incentive of a nursing home to be the profit difference between its inspection rating and the highest overall rating it could potentially obtain after self-reporting, as shown in Table 2-3. Note that the financial incentive arises from the expectations in both profits and losses. It is possible for a nursing home that has received 5 stars from the health inspection to lose two stars if it receives one star in the self-reported domains. However, it is very unlikely that a nursing home with perfect health inspection can be significantly under staffed or provides very poor quality of care. In our dataset, while 125 nursing homes initially rated 3 stars in inspection gained two additional stars after self-reporting, only 4 nursing homes initially rated five stars in inspection lost two stars after self-reporting.

2.4.3 Empirical Model Specification

We focus on the change in the star rating that happens as a result of the self-reported measures. Our dependent variable, *StarChange,* is equal to the difference between the overall rating and the health inspection rating. For example, if the nursing home receives 3 stars from health inspection but receives a 5-star overall rating after including its self-reported measures on staffing and quality measure domains, then the *StarChange* would be equal to two.

By definition, *StarChange* can only take discrete values of 2, 1, 0, -1 and -2, and thus we use an ordinal logistic specification in which *StarChange* is modeled as a function of a vector of independent variables. *StarChange* is determined by a set of parameters, α_{-2}, α_{-1}, α_0, α_1, which define the cutoff points of the five levels. *StarChange* for nursing home *i* at year *t* can be modeled as follows

$$P(StarChange_{it} \leq j) = \frac{\exp(\alpha_j + \boldsymbol{x}'_{it}\boldsymbol{\beta})}{1+\exp(\alpha_j + \boldsymbol{x}'_{it}\boldsymbol{\beta})} \tag{2.1}$$

where $j \in \{-2,-1,0,1\}$ and $\boldsymbol{x}$ is a vector of the following independent variables: *Incentive, BedCert, OccuRate, MarketShare, HHI, ForProfit, Medicare, Medicaid, CouRes, CouFam, PctgMedicare, PctgMedicaid, PctgSelfPay, PctgMGD, Chain.*

Among the independent variables, the main effect we consider in our model is the nursing homes' financial incentive, denoted by *Incentive,* and as shown in Table 2-3 varies depending on the inspection rating of a nursing home. The capacity of each nursing home is measured by the number of certified bed, and is denoted by variable *BedCert*. The occupancy of a nursing home is denoted by variable *OccuRate*, *OccuRate*$\in$[0, 1]. Variables *BedCert* and *OccuRate* together, define the average number of residence of a nursing home. Nursing homes are located in different areas, and may face different market conditions. To capture local market features, we use variable *MarketShare* to denote the market share of each nursing home in its local market, defined by Health Service Area (HSA). Based on market share, we also calculate the *Herfindahl-Hirschman Index (HHI),* which is widely used for capturing local market competition, and included it in our empirical model. Variable *ForProfit* defines a nursing home's

ownership type and is equal to one if the nursing home is for-profit and zero otherwise. Variables *Medicare* and *Medicaid* define a nursing home's certification. *Medicare* is equal to one if the nursing home is Medicare certified, likewise, *Medicaid* is equal to one if the nursing home is Medicaid certified. By law, nursing homes are required to allow councils set up by residents or their family members. These councils facilitate the communication with staff and get problems resolved more efficiently. Since nursing home residents may be more vulnerable than normal people due to their health conditions, the residential council and family council can function very differently in resolving issues and handling complaints. In our model, binary variables *ResCouncil* and *FamCouncil* are included to respectively, denote the council types as residential and family. A nursing home can have both types of councils. The OSHPD data categorize nursing home payers into five categories: Medicare, Medicaid, Self-Pay, Managed Care, and Others. To capture the impact of different payer percentage on nursing homes' star rating changes, we incorporate the percentage of each type of payers. Four variables, *PctgMedicare*, *PctgMedicaid*, *PctgSelfPay*, *PctgMGD* are added to denote the percentage of Medicare payers, Mediaid payers, Self-paying payers, and Managed care payers. The percentage of other type payers are excluded due to multicollinearity. In the nursing home industry, a certain amount of nursing homes is running under some chains. Comparing with nursing homes working as separate facilities, nursing homes in chains may have different operational rules and self-reporting behaviors. In our pooled California data, we have over 1500 records of nursing homes in a chain, and there are totally 101 distinct chains. As a result, we do not have sufficient observations for each of the chains to conduct a fixed effect analysis. Rather than adding a chain-level fixed effect, we regroup the nursing homes and add binary variable *chain*, which equals 1 if the nursing home is operating in a chain and 0 the nursing home is operating separately. Table 2-4 provides the summary statistics of all variables in our model.

Table 2-4. Variable Summary Statistics

Variable	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*
Incentive	4.497	3.201	0	9.011
BedCert	101.964	49.579	19	391
OccuRate	0.874	0.172	0.0497	1
ForProfit	0.891	0.292	0	1
Chain	0.767	0.423	0	1
Medicare	0.963	0.188	0	1
Medicaid	0.965	0.183	0	1
CouRes	0.979	0.143	0	1
CouFam	0.230	0.421	0	1
MarketShare	0.0165	0.0159	0.000277	0.125
HHI	5.249	11.271	0.000765	156.314
PctgMedicare	0.154	0.127	0	0.921
PctgMedicaid	0.647	0.235	0	1
PctgSelfPay	0.0838	0.129	0	1
PctgMGD	0.663	0.103	0	0.999

2.4.4 Estimation Results

We estimate equation (2.1) by different methods, as shown in Table 2-5. The first column shows the estimation results for the pooled data. To deal with potential endogeneity, we take nursing homes' fixed effects into account and run a panel data regression which its estimates are shown in the second column. Some of the variables in our model are time-invariant. For example, if a nursing home is Medicare certified in year 1, it will most likely remain Medicare certified throughout the following years. As a result, we cannot estimate their coefficients directly through the fixed-effect method. To obtain the coefficients of these time-invariant variables, we implement Hausman-Taylor method, as shown in the third column. In the estimates from all methods, the main effect *Incentive* is positive and statistically significant, which indicates that nursing homes with higher financial incentives are more likely to improve their star ratings after self-reporting. In all the three models, we observe negative significant coefficients for variable *chain*, indicating that for nursing homes operating in chains, their star rating increases are less likely to be driven by their financial incentives.

Table 2-5. Estimates of Equation (2.1)

Variables	*Pooled data*	*Fixed effect*	*Hausman-Taylor*
Incentive	0.0325***	0.074***	0.074***
	(0.0906)	(0.0144)	(0.0143)
BedCert	0.000284	-	0.00162
	(0.000655)		(0.00207)
OccuRate	-0.584***	0.627	0.445
	(0.166)	(0.560)	(0.468)
ForProfit	-0.0128	-	-1.578***
	(0.112)		(0.332)
Chain	-0.269***	0.0265	-0.444**
	(0.0698)	(0.247)	(0.148)
Medicare	-0.805***	-	-2.684***
	(0.170)		(0.472)
Medicaid	-0.622***	-	-1.323*
	(0.184)		(0.595)
CouRes	-0.231	-	-0.293
	(0.202)		(0.536)
CouFam	-0.279***	-	-0.315
	(0.0693)		(0.174)
MarketShare	-6.79	-71.073*	-60.348*
	(4.49)	(34.363)	(28.568)
HHI	0.020***	0.0777*	0.0781*
	(0.00614)	(0.0343)	(0.0339)
PctgMedicare	-2.283***	4.280***	4.242***
	(0.339)	(1.290)	(1.269)
PctgMedicaid	-0.771***	1.610	1.819
	(0.254)	(1.090)	(1.080)
PctgSelfPay	-0.861***	-4.851***	-4.447***
	(0.321)	(1.063)	(1.042)
PctgMGD	-0.446	6.291***	6.386***
	(0.369)	(1.207)	(1.198)

2.4.5 Alternative Incentive Definition

In the above section, we define the financial incentive of nursing homes based on the average per patient daily profit over the 5-year period. The financial incentive, however, may vary over the years. For example, the difference in the average per patient daily profit between 3-star nursing homes and 5-star nursing homes in year t+1 may be bigger than that in year t. To capture this change over the years and to test the robustness of our result, we propose an alternative incentive definition in this section. Instead of looking at a 5-year average level, we instead use the per patient daily profit difference of the year t to define nursing homes' financial incentive of year t+1. Table 2-6 lists the new financial incentive under the new definition. Table 2-7 then gives the regression results under the alternative financial incentive definition. Similar to the discussion in the previous section, we also run three models: the pooled data model, fixed effect model and Hausman Tylor model. In all the three models, the main effect financial incentive is positive significant, which demonstrates the robustness of our results.

Table 2-6. Alternative Definition of Financial Incentive

Health inspection rating	*2010 Financial Incentive*	*2011 Financial Incentive*	*2012 Financial Incentive*	*2013 Financial Incentive*
5	0	0	0	0
4	0	0	0.38	3.655
3	0	0.687	3.399	5.183
2	2.129	3.525	3.97	2.858
1	0	1.126	1.447	0.937

Notes:
The financial incentive of year t is defined by using the year t-1 data. Since the panel we collected is from 2009-2013, we have no data to define incentives for year 2009, and the year 2013 data (which should be used for 2014 according to the definition) is not used in this definition.
In early years (2009 and 2010), there is no significant difference in per patient daily profit for some of the rating levels, thus the financial incentive for improving star rating is defined as 0.

Table 2-7. Estimates of Equation (2.1) based on the Alternative Financial Incentive Definition

Variables	*Pooled data*	*Fixed effect*	*Hausman-Taylor*
Incentive	0.323***	0.412***	0.412***
	(0.0202)	(0.0253)	(0.0252)
BedCert	-0.000111	-	-0.000531
	(0.000733)		(0.00210)
OccuRate	-0.515**	0.050	0.257
	(0.190)	(0.625)	(0.517)
ForProfit	-0.0384	-	-0.907**
	(0.126)		(0.352)
Chain	-0.288***	-0.336	-0.522***
	(0.790)	(0.265)	(0.148)
Medicare	-0.900***	-	-2.119***
	(0.189)		(0.486)
Medicaid	-0.461*	-	-1.309*
	(0.203)		(0.623)
CouRes	-0.273	-	-0.244
	(0.227)		(0.495)
CouFam	-0.214**	-	-0.142
	(0.0777)		(0.166)
MarketShare	-8.128	-5.752	-19.306
	(5.101)	(39.159)	(32.179)
HHI	0.0218**	0.0308	0.0344
	(0.00706)	(0.0392)	(0.0388)
PctgMedicare	-2.189***	0.644	0.828
	(0.375)	(1.574)	(1.551)
PctgMedicaid	-0.862**	1.241	1.312
	(0.282)	(1.264)	(1.254)
PctgSelfPay	-0.377	-2.206	-1.990
	(0.364)	(1.281)	(1.248)
PctgMGD	-0.563	3.397*	3.481**
	(0.406)	(1.357)	(1.350)

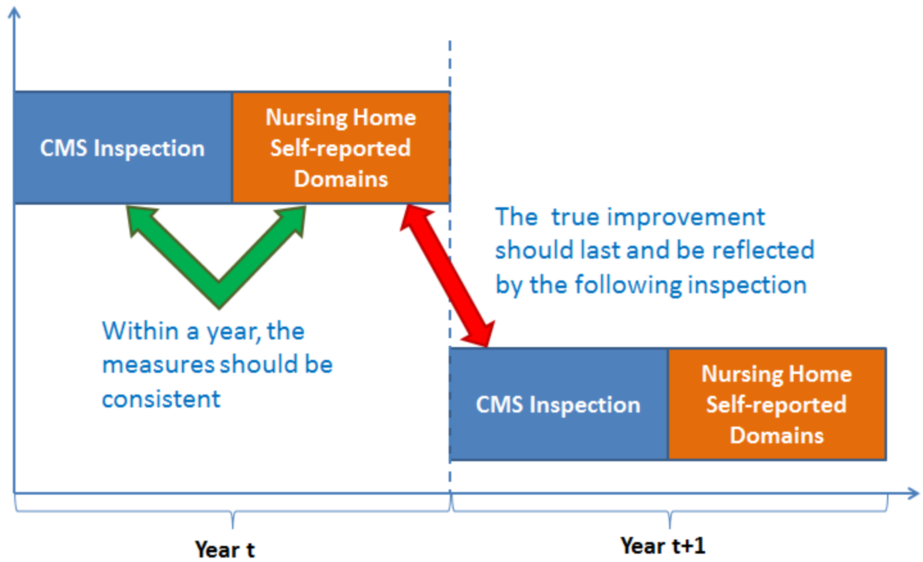
2.5. Inflation Detection and Demonstration

2.5.1 Correlation Analysis

Although the preliminary results show a positive association between the financial incentive and the changes in the star-rating, they do not necessarily indicate inflation in self-reported measures. It is possible that nursing homes gain the additional stars legitimately through their true efforts. To explore the underlying reasons for the changes in ratings, we investigate the correlation between the health inspections and self-reported domains. As illustrated in Figure 2-4, under the assumption that there is no inflation and nursing homes self-reported measures are legitimate, positive correlations are expected between two sets of ratings. First, within the same year, a positive correlation is expected between the star ratings from CMS health inspection and those of nursing homes' self-reported domains. Second, if a nursing home really puts an effort in improving its care quality, these efforts should have a lasting effect and lead to better results in the next year's health inspections and thus there should be a positive correlation between the star ratings from self-reported domains in one year and health inspection ratings in the subsequent year.

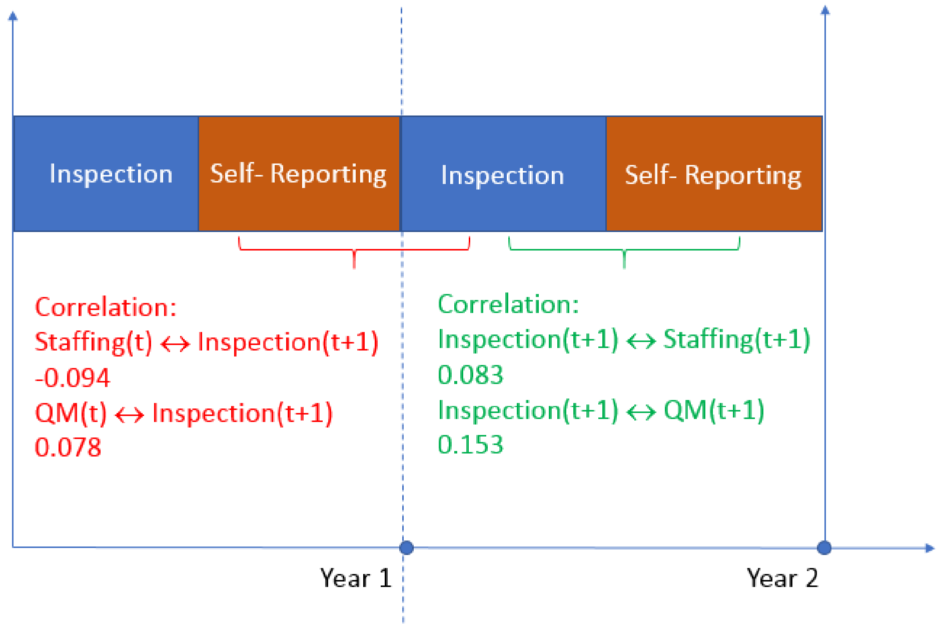
Figure 2-5 shows the two sets of correlations as described above. It can be seen that within the same year, the correlation between Health Inspection and Staffing is only 0.083, while the correlation between Health Inspection and Quality Measures is 0.153. The result clearly indicates inconsistency between the health inspections and self-reported domains within the same year. For the two consecutive years, the correlation between Staffing and the Health Inspection in the following year is -0.094, and the correlation between Quality Measures and the Health Inspection in the following year is 0.078. The result indicates that the self-reported improvements in quality measure and staffing domains have no lasting effect on the next year's health inspection results at all. The correlation analysis serves as a preliminary evidence of potential inflation, and triggers our further analysis.

Figure 2-4. Graphical Representation of Correlation Analysis



Note: If star increase is resulted from legitimate efforts, then a positive correlation is expected between self-reported measures in year 1 and on-site inspections in year 2 (red arrow). A positive correlation is also expected between the self-reported measures and the on-site inspection ratings in the same year (green arrow).

Figure 2-5. Correlation Analysis for 5 Consecutive Years 2009-2013



For the 5-year period we analyzed, it appears that the correlation between Health Inspection and Quality Measures is higher. One interpretation is that the inflation on the Staffing is relatively easier than that of Quality Measures during this period. We do notice that CMS is gradually releasing amendment on the nursing home rating system, and one important policy is to require all nursing homes to report payroll related staffing data since the beginning of 2017. This shows that the inflation on staffing level is also one of the major concerns of CMS, and once payroll related staffing data is reported, the Staffing measure will become more difficult to be inflated.

2.5.2 Complaint-based Analysis

In this section, we conduct further analysis to justify the existence of rating inflation. We identify a quantifiable third-party proxy variable which can serve as an independent measure of service quality, and compare the results with the star ratings given by the rating system. If significant inconsistency exists between the two, then the star ratings are questionable, and rating inflation likely exists. In our method, we use the number of complaints, which has been used as a common measure of the service quality in the literature of service and complaint management in many service industries (E. Anderson, Claes, and Roland 1997; Gardner 2004; Johnson 2001; Roland and Chung 2006). Specifically, we conduct an analysis based on the CDPH complaint data which is independently collected data set of patient complaints of California nursing homes. The combined CMS, OSHPD and CDPH dataset has 3850 records of California nursing homes over the 5 years.

If inflation does not exist, then the overall rating should be consistent with the true service quality, which is reflected by the number of complaints. That is, for nursing homes with the same overall rating, we expect them to have similar service qualities and similar number of complaints. Table 2-8 (a) shows the average number of complaints for nursing homes with different health inspection and overall ratings. In view that larger nursing homes with more patients may get more complaints, we normalized the number of complaints by the size of a nursing home. The normalized results are presented in Table 2-8 (b).

Table 2-8. Average Number of Patient Complaints

(a) Original Complaints

		Overall star rating				
		1	2	3	4	5
Inspection stars	1	7.981	6.989			
	2	6.193	6.271	6.010	8.389	
	3	3.929	3.934	4.633	4.940	4.056
	4		3.923	3.799	3.503	2.826
	5			6.667	2.157	2.423

(b) Normalized Complaints (size=100)

		Overall star rating				
		1	2	3	4	5
Inspection stars	1	6.505	6.909			
	2	5.946	5.687	6.210	8.251	
	3	4.366	4.676	4.860	5.597	4.288
	4		4.218	4.459	4.157	3.547
	5			9.473	2.517	2.921

Note: [a] The blank cells represent the impossible rating transaction according to CMS's rating systemdesign.

[b] The shaded cells represent nursing homes of which the ratings increased after self-reporting. The inflators are among these nursing homes. We denote the shaded and unshaded areas as Area I and Area II, respectively.

For each overall rating level, the nursing homes are divided into two categories: Nursing homes whose star ratings increased after self-reporting and nursing homes whose star ratings did not increase after self-reporting. We denote the upper triangular section as area *I (shaded)* and lower triangular section as area *II*. The shaded area (*I*) includes those nursing homes whose overall rating has increased as a result of their self-reported measures. Area *II* includes those nursing homes whose overall rating either decreased or remained the same after self-reporting. This classification allows us to test the following claims:

Claim 1: If the improvements observed are not resulted from legitimate efforts and inflation does exist, nursing homes with the same overall star rating but different health inspection ratings should have different complaint distributions.

The results of two ANOVA tests are presented in Table 2-9 (a). In the first column, nursing homes with the same overall ratings are grouped by whether or not their star rating increased after self-reporting. In other words, we examine if the shaded and unshaded cells in each column of Table 2-8 have similar distributions. In the second column, we group nursing homes with the same overall rating based on their health inspection ratings. In other words, we examine if all the cells in each column of Table 2-8 have similar distributions. As reported in Table 2-9 (a), all the comparisons are significant and thus the claim that nursing homes with the same overall rating but different inspection ratings have different complaint distributions is supported. The ANOVA test results for the normalized complaints are reported in Table 2-9 (b), which are similar to the results in Table 2-9 (a) and support our conclusion.

Claim 2: If the improvements observed are not resulted from legitimate efforts and inflation does exist, nursing homes with the same inspection rating but different overall ratings should have similar complaint distributions.

The results of two ANOVA tests are presented in Table 2-10 (a). In the first column, nursing homes with the same health inspection ratings are grouped by whether or not their star rating increased after self-reporting. In other words, we examine if the shaded and unshaded cells in each *row* of Table 2-8

have similar distributions. In the second column, we group nursing homes with the same inspection rating based on their overall star ratings. In other words, we examine if all the cells in each row of Table 2-8 have similar distributions. As shown in Table 2-10 (a), we do not observe a significant difference in the number of complaints, although the overall rating can be quite different. The results show that service quality does not improve for nursing homes whose star ratings get improved after self-reporting and thus Claim 2 is also supported. Together with the results obtained for Claim 1, the analysis provides strong evidence of the existence of rating inflation in self-reported measures. The ANOVA test results for the normalized complaints are reported in Table 2-10 (b), which also support our conclusion.

Table 2-9. F Statistics: Comparison in Each Overall Rating a) Original Complaints b) Normalized Complaints (size=100)

(a) F Statistics

		Grouped by Area I vs Area II	*Grouped by inspection ratings*
Overall star rating	1	-	4.61**
	2	7.43***	6.16***
	3	13.05***	5.06***
	4	14.22***	8.35***
	5	5.27**	5.70***

(b) F Statistics

		Grouped by Area I vs Area II	*Grouped by inspection ratings*
Overall star rating	1	-	0.94
	2	7.77**	3.25*
	3	7.2**	2.88*
	4	10.15**	5.33**
	5	3.92*	2.64

Table 2-10. F Statistics: Comparison in Each Inspection Rating a) Original Complaints b) Normalized Complaints (size=100)

(a) F Statistics

		Grouped by Area I vs Area II	*Grouped by overall ratings*
Inspection ratings	1	2.46	2.46
	2	0.12	0.12
	3	0.78	0.78
	4	5.37**	2.00
	5	-	1.91

(b) F Statistics

		Grouped by Area I vs Area II	*Grouped by overall ratings*
Inspection ratings	1	0.5	0.5
	2	1.6	1.46
	3	0.99	0.77
	4	2.51	0.95
	5	-	3.3*

2.6. Prediction Model and Variable Importance Analysis

In this section, we first develop a method which gives a quantifiable estimate of the extensiveness of rating inflation. We then run a variable importance analysis to summarize key characteristics of the likely inflators.

For nursing home that inflates its self-reported measures, the overall rating is driven by two components. The first component is the observable characteristics which are common between cheating and honest nursing homes. The second component is the unobservable inflation coefficient which only pertains to the inflating nursing homes. If we model the overall ratings as a function of observed characteristics, the inflation component is unobserved and omitted from our regression model, thus the estimates of the remaining observed variables will suffer from the omitted variable bias. However, since the overall star ratings of honest nursing homes are only driven by one component of observed characteristics and the inflation component does not exist among the honest nursing homes, our regression estimates for the honest group will not suffer from the omitted variable bias. To develop our inflation prediction model, we first divide the nursing homes into two groups: the honest nursing homes and the remaining, defined as potential inflators. A regression is then run for the honest nursing homes. The obtained regression coefficients from the sample of honest nursing homes are unbiased and reflect the true associations without inflation. These unbiased coefficients are then used to predict the highest possible overall star rating for each nursing home in the suspected inflating group. A nursing home is identified as a likely inflator in our estimation if its actual overall rating is higher than the highest level of its predicted overall rating.

2.6.1 Prediction Model

In our model, the overall star rating is used as the dependent variable, denoted by *OverallRating*. Similar to the variable *StarChange* in the regression model in Section II, *OverallRating* is ordinal and takes values in five levels $\{1,2,\ldots,5\}$ so we employ an ordinal logistic regression model. *OverallRating* is

determined by a set of parameters γ_1, γ_2, γ_3, γ_4, which define the cut points of the five star levels. The model can be written as

$$P(OverallRating \leq k) = \frac{\exp(\gamma_k + \boldsymbol{x}'\boldsymbol{\beta_P})}{1+\exp(\gamma_k + \boldsymbol{x}'\boldsymbol{\beta_P})}, \tag{2.2}$$

where $k \in \{1, 2, 3, 4\}$. The independent variables denoted by vector $\boldsymbol{x}$ are the same as the ones used in equation (2.1). The coefficients of prediction model are denoted by $\boldsymbol{\beta_P}$.

Since we use the coefficients of the honest group as the unbiased baseline, we define the members in this group very strictly to guarantee that there is no evidence of inflation for all nursing homes in the honest group. An honest nursing home is selected based on the following criteria:

1. Its overall star rating does not increase after self-reporting.

2. The number of its patient complaints is strictly lower than the median of its corresponding self-reporting level.

Our logic for selecting the honest nursing homes is as follows: We divide the inflators into two different types. The first type consists of nursing homes which inflate their self-reported measures to achieve higher ratings. For these inflators to be identified, a necessary condition is that they gain additional stars after self-reporting (there can be honest nursing homes who gain the additional stars through legitimate efforts though). In our first criterion, we excluded all the nursing homes whose star rating increased after self-reporting, thus we completely excluded any inflators of this type. The second type of inflators consists of nursing homes which inflate their self-reported measures to avoid losing stars. These nursing homes may have low *Staffing* or low *Quality Measures* that may lead to decreased overall ratings. In our second criterion, we excluded nursing homes whose number of complaints are above the median of its rating level. By using these two criteria, we excluded nursing homes who may lose stars due to their poor services, and guaranteed that the remaining nursing homes deserve staying in that rating level.

Based on the two criteria, we identify the honest (*H*) group, which consists of 1262 nursing home records in 5 years. The remaining 2588 nursing home records are categorized in the potential inflator (*PI*) group. Note that the PI group consists of both the actual inflators and the nursing homes who improve their service qualities through legitimate efforts. In the following, we estimate the proportion of the actual inflators in the *PI* population.

We run the ordinal logistic regression in equation (2.2) on the sample of honest nursing homes (H group) to obtain the unbiased estimates of each coefficient. The regression results for the honest group is reported in Table 2-11. Both the 95% and 90% confidence interval are calculated. Using the upper bounds of unbiased coefficient estimates, we then predict the highest possible rating for each of the nursing homes in the *PI* group. A nursing home is classified as an inflator if its actual overall star rating is higher than the highest possible rating predicted through our model. Based on the 95% confidence interval, we can identify 147 inflator records out of the 2588 nursing home records (5.68%) in the PI group. Based on 90% confidence interval, we can identify 219 inflator records (8.46%) in the PI group.

Table 2-11. Estimates of the Honest Group

Variables	*Coefficients*
Incentive	0.102***
	(0.0184)
BedCert	-0.0115***
	(0.00125)
OccuRate	0.863**
	(0.300)
ForProfit	-1.491***
	(0.208)
Chain	-0.430***
	(0.130)
Medicare	-2.853***
	(0.563)
Medicaid	-1.580***
	(0.408)
CouRes	-0.613
	(0.420)
CouFam	0.646***
	(0.122)
MarketShare	2.234
	(8.146)
HHI	0.00889
	(0.0116)
PctgMedicare	-0.330
	(0.758)
PctgMedicaid	-0.0678
	(0.626)
PctgSelfPay	1.582*
	(0.718)
PctgMGD	-2.551***
	(0.793)

2.6.2 Variable Importance Analysis

It is important to understand the key differences between honest nursing homes and the inflators, so that we can focus on these differences in audits and identify the inflators efficiently. In this section, a variable importance analysis is conducted to explore the key characteristics of the inflators. A subset of the data is first constructed by eliminating nursing homes whose status cannot be identified. The eliminated nursing homes are the ones which are neither identified as likely inflators nor identified as honest ones. The remaining dataset consists of 1481 nursing home records, in which 1262 records are for honest nursing homes and 219 records are for the likely inflators identified using a 90% confidence interval. The status of a nursing home is assigned as 0 if it belongs to honest group and 1 if it is a likely inflator. To perform the variable importance analysis, we use the logistic specification presented in equation (2.3).

$$logit(\lambda) = \boldsymbol{x}'\boldsymbol{\beta} \tag{2.3}$$

where λ is the probability of being identified as an inflator, $\boldsymbol{x}$ is the vector of variables that were also used in equations (2.1) and (2.2).

The variable importance analysis results are presented in Table 2-12. Among the variables, we find the variable *BedCert* to be the top in terms of variable importance. The result indicates that when a nursing home's size grows, its probability to game the rating system increases significantly. The percentage of self-paying is also a key variable contributing to being an inflator. As discussed earlier in the incentive definition section, self-paying patients are typically good in financial situations and they contribute significantly to nursing homes' non-healthcare profits. Since non-healthcare pricing is not regulated by CMS, highly rated nursing homes typically charge much higher prices on non-healthcare services than low-rating nursing homes. It is reasonable to believe that many nursing homes with high percentage of self-paying patients are inflating their self-reported measures in order to gain more non-healthcare profits. The results also indicate that nursing homes with family type councils are more likely to be inflators. Another variable *Chain* also has a very high importance. Note that in section 4, our results suggest that nursing homes in chains are less likely to be driven by their financial incentives to

improve star ratings. One possible explanation for these results is that nursing homes in these franchises follow chain-level decisions for self-reporting, which are less sensitive to individual nursing home's financial incentive. It is possible for some chains to inflate self-reported measures throughout their facilities. Besides the variables discussed above, *Incentive* and *ForProfit* are also important factors for being an inflator. For-profits nursing homes are more likely to inflate their self-reported ratings than the non-profits ones, and the higher their financial incentives are, the more likely they will be inflators. This result is consistent with the work of Chesteen et al. (2005). The probability of being an inflator, on the other hand, is less likely to be affected by market competition (e.g., *HHI*, *Market Share*) and certification status (*Medicare*, *Pctg_Medicare*, *Medicaid*, *Pctg_Medicaid*, etc.).

Table 2-12. Variable Importance Analysis

Variables	*Variable Importance*
BedCert	0.405
Pctg_SelfPay	0.288
Cou_Fam	0.274
Chain	0.264
Incentive	0.254
Occu_Rate	0.194
ForProfit	0.182
HHI	0.176
MarketShare	0.171
Pctg_MGD	0.122
Medicaid	0.109
Pctg_Medicaid	0.097
Cou_Res	0.066
Pctg_Medicare	0.043
Medicaid	0.041

2.7. Conclusion

This chapter systematically analyzes CMS's nursing home rating system, demonstrates the existence of inflation, and presents a model to detect likely inflators. We show that nursing homes have strong financial incentives directly related to higher star ratings, which may in turn drive the inflating behaviors. We then develop a systematical method which uses independent third-party measure of patient complaints to demonstrate the existence of rating inflation. An inflation prediction model is then developed, which provides an estimate of the proportion of inflating nursing homes in the current system, and gives a quantifiable evaluation of the system performance. The variable importance analysis is then performed to identify the factors that contribute the most to being an inflator.

Our research provides several contributions. First, to the best of our knowledge, this is the first study that systematically investigates the inflation in the CMS nursing home rating system. It explores the fundamental financial reason for a nursing home to improve the star rating, even by inflating self-reported measures, which links the dots between incentives and observed behavior. Second, we contribute to the theory by developing this systematical method for demonstrating the existence of rating inflation and evaluating inflator proportion. As we discussed earlier, although CMS has implemented minor improvements to its rating system, it is still largely based on self-reported measures and does not address the issue of inflation. Our research demonstrates the shortcoming of the rating system and informs CMS on how to improve its system or how to identify the likely inflators. This study estimates the proportion of likely inflators and summarizes their key characteristics. The results can be used to strategically focus the future audits on the nursing homes which are most likely to be inflators, and help CMS improve the rating system.

This work also has several limitations. First, we are unable to measure the financial incentives for each nursing home at the individual level. This is practically very difficult, since even for the same nursing home at the same rating level, the financial incentive may vary over time depending on various financial situations. To address this limitation, we perform our analysis on an aggregated level and use the

average as a universal incentive for each rating group, leaving the unobserved incentive fluctuations to the nursing homes fixed effects. Second, we do not observe self-reporting inflation directly and can only infer it from an aggregated level. This is a common issue in misbehavior detection research due to the unavailability of individual-level data. We address this limitation by calculating the highest possible rating using the confidence interval and using the most conservative statistics. Third, we are only able to measure patient complaints in numbers, but not in "severeness". For example, a complaint on medical malpractice may have much more impact than a complaint on sanity. Future research can apply text mining techniques to address this limitation.

Chapter 3. Catching Them Red-handed: Optimizing the Nursing Homes' Rating System

3.1 Introduction

Nursing homes constitute an important segment of the U.S. healthcare system. They provide care to 1.5 million patients in America (Fowles et al. 2012). Medicare annually spends more than $49 billion on the services provided in nearly 16,000 nursing homes in the United States (KFF, 2012). 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 widely popular among 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 on-site rating (CMS 2015). The two self-reported measures can significantly affect a nursing home's overall ratings; for example, a nursing home that initially received three stars from the on-site inspection can increase its overall ratings to five stars if it self-reports excellent measures on its quality and staffing. Prior research shows that between six to twelve percent of the nursing homes inflate their self-reported measures as a strategy to gain higher overall star ratings (Han et al. 2016).

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 strategy to detect and deter fraud in the self-reported measures of the domains that are not directly inspected by CMS. This research, as we describe below, seeks to optimize

the inspection strategy and design an audit system for CMS to improve its current nursing homes' rating system.

Currently, CMS only inspects the Health Inspection domain as one of the three domains. It is not yet known if inspecting this domain is the optimal strategy. In this research, we examine other domains which CMS can conduct inspection on in order to either minimize the percentage of nursing homes that 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 degree of income tax evasion in the US (Cebula 2012). Despite these potential 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 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, on the other hand, will decide how much to inflate their self-reported measures based on their expected profits under different CMS auditing policies.

This chapter presents major insights for improving the inspection system and designing an audit strategy. We develop a novel graph-based model and optimize the selection of inspection domains by minimizing the percentage 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 the domain that CMS is currently inspecting is optimal in term of minimizing the difficulty to detect inflators. We develop conditions on the audit parameter settings and justify our findings through simulation. Our

results indicate that CMS should apply a moderate audit policy in order to balance the tradeoff between the audit's net budget and its efficiency.

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.

The chapter proceeds as follows. In Section 3.2, we review the related literature on audit systems to detect and deter fraudulent behavior. In Section 3.3, we use a novel graph-based 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 percentage of nursing homes that can inflate, and minimizing the difficulty for detecting inflators. In Section 3.4 we design the audit strategy. We consider nursing homes' reaction to different audit policies, and derive conditions on the parameter settings of the audit. We then conduct a simulation of the audit process to examine our analytical results. Finally, Section 3.5 concludes the chapter.

3.2 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. 2012, Abbasi etal. 2015, Wright et al. 2010, Cecchini et al. 2010, Mayzlin et al. 2012, Duggan et al. 2000, Bai et al. 2010, Jocob et al. 2003). The fundamental approach of all fraud detection methods is to pinpoint "abnormal patterns" embedded in the data. We divide these methods into different streams based on the methods used 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 et al. (2002) and Ma et al. (2013) develop a mechanism to detect fake ratings in which a rater is considered dishonest on evaluating an entity, if the evaluation score from all other raters falls in the rejection area of 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) proposes 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) uses a graphical method to analyze the relationship among raters, ratings and entities. Mukherjee et al. (2013) designs a rating fraud detection model which identifies fake ratings by calculating their deviation from the majority. The above studies rely on the majority rule for fraud detection and the assumption that the population provides a consistent evaluation of a certain subject. 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 et al. (2010) proposes 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 et al. (2010) proposes 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) detect suspicious ratings using the previously identified honest ratings as a filter to explore dissimilarities. Teacy et al. (2006) evaluates the trustworthiness of a rater by comparing her ratings with the other previously identified honest raters. Liu et al. (2014) utilizes difference between local and global ratings to identify fake ratings. Han et al. (2015) develop a method

which first identifies honest nursing homes according to a set of restrictive criteria, and then builds a prediction model based on the identified honest nursing homes to detect inflators.

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 only be important for certain groups of patients but not others. In such cases, each type of 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 above methods usually require data on the characteristics of those who are more likely to commit fraud. However, there is currently no audit system to catch the inflators, and 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. (2015) collect 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 12.5% 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. In this chapter, we first set up a model to identify the domain which CMS should inspect in order to eliminate the possibility of inflation, then look into the design of audit system.

3.3 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 the truthful nursing home that pursue genuine efforts to improve their ratings. The selection of inspection domain determines which domains are left self-reported, and has prolonged effect on the reliability of CMS's ratings. Due to limited resources, CMS only inspects one of the three domains and has to rely on nursing homes to self-report their performance in the other two domains. CMS could potentially divert its resources to inspect other domains than the one that it currently does. That is, instead of conducting health inspections, CMS could conduct inspections on quality measures or staffing levels. In this section, we examine alternative inspection strategies under different objectives to identify which of the three domains CMS should inspect in order to ensure that the percentage 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 seek for the best possible rating they can achieve. This implies that all nursing homes can inflate their ratings on domains that are not inspected. Since CMS currently does not have an effective audit mechanism and inflating nursing homes are almost never caught, this assumption can be viewed as the *worst-case scenario*. The assumption is relaxed when nursing homes' reaction for punishment is incorporated to design an audit strategy in the next section. We formulate CMS's inspection strategy in two models within a graph-based framework and then test our models using CMS's historical data.

3.3.1 The Graph-based Framework

We undertake a graph-based approach to model the combination of the ratings of nursing homes' in the three domains and the potential of rating inflation. As shown in Figure 3-1, each node represents a possible rating combination, and is depicted by (A, B, C) in which A, B and C represent the star rating in each of the three domains of *health, staffing* and *quality measures*. As discussed earlier, the overall ratings increase by one star only if the performance in either of the two self-reported domains is equivalent to five stars. As a result, we can transform the five-star ratings into a 0-1 setting in which "1" represents being rated 5 stars and "0" represents being rated four stars or less. Each node thus shows one possible combination of the three measures. For example, node (000) shows the nursing homes that are not rated five stars in any of the three measures. For each node, the corresponding probability of each rating combination is listed in the parenthesis under it.

For each domain, a rating of five stars leads to one additional star in the overall ratings, and results in moving from the current node to the corresponding node in the next level. Such moves are represented by the arcs in Figure 3-1, which are indexed based on the combination of its origin and destination. For example, arc 14 denotes the arc from node 1 to node 4, and indicates that nursing homes at node (000) can have 5 stars in domain A and move to node (100). For any two nodes connected by an arc, the ratings on only 1 domain changes. If a domain is selected for inspection, nursing homes cannot longer inflate their ratings of that domain and the corresponding arcs will be removed. For example, if domain C is inspected, nursing homes will not be able to move to the next rating level by inflating measure C, and therefore arcs 14, 25, 37 and 68 will be removed. An extreme case is for CMS to inspect all the three domains. In this case, all the arcs are removed, and the nodes become isolated. In other words, no nursing home will be able to inflate its star ratings. In the following, we formulate and solve the inspection problem with two objective functions: minimizing the percentage of nursing homes that can inflate, and minimizing the difficulty to detect inflators. Table 3-1 shows all the notations and symbols used in the problem formulation.

Figure 3-1. The Graph-based Framework

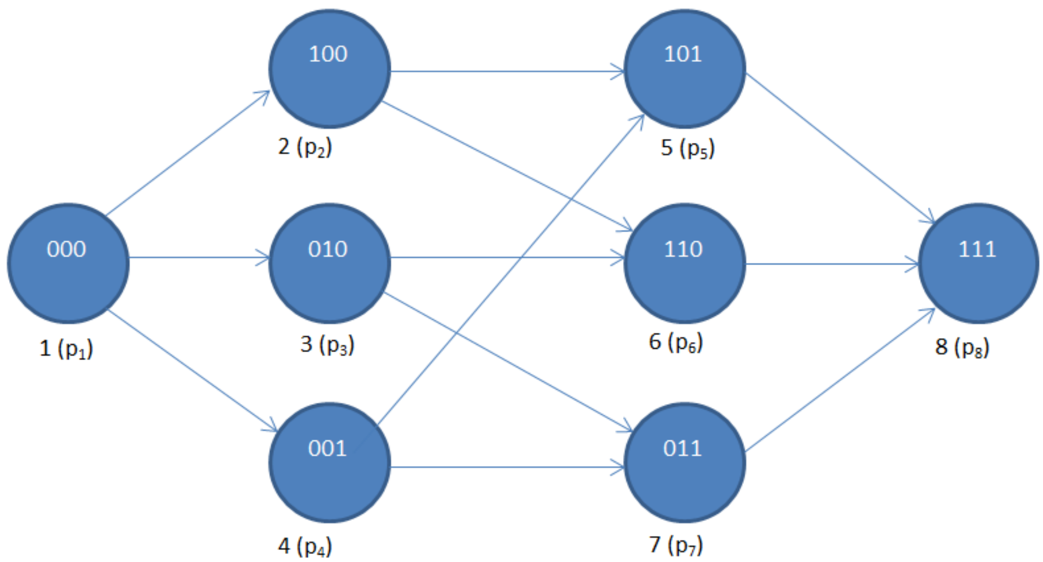
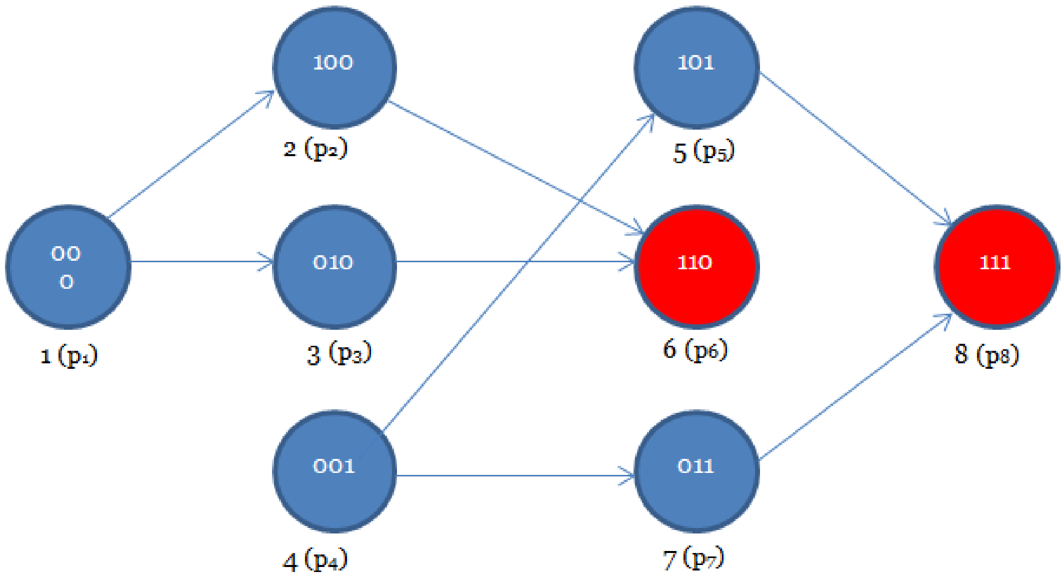


Figure 3-2. The Graph Topology



Suppose domain C is inspected, then arcs 14, 25, 37 and 68 will be removed. The graph is divided into two sub-graphs, with node 6 and node 8 be the ending nodes for each sub-graph

Table 3-1. Notations and Symbols used in the Inspection Problem

i, j	Node index
k	Inspection measure index
π_i	The percentage of nursing homes showing the rating combination associated with node i. $\Sigma_i \pi_i$ =1
x_{ij}	Arc connectivity variable for arc ij. x_{ij}=1 if node i and node j are connected by arc ij
y_k	Inspection measure decision variable. y_k=1 if measure k is selected for inspection
z_i	Flow indicator variable for node i. z_i=0 if no flow is coming out of node i
c_k	The cost to inspect measure k
C	CMS's total inspection budget
v_{ij}	Ending node indicator variable. v_{ij}=1 if node j is node i's ending node, i.e., nursing homes at node i, can increase their star ratings and move to node j.
p_j	The probability of showing the rating combination associated with node j. $\Sigma_j p_j$ =1
M	A big positive number used for linear conversion.

3.3.2 Minimizing the Percentage of Nursing Homes That Can Inflate

In this formulation, we minimize the percentage of nursing homes that can inflate. That is, we minimize the percentage of nursing homes that can move to higher rating nodes. Consider a problem with I nodes indexed by i and K domains indexed k. The percentage of nursing homes at node i is denoted by π_i, where π_i can be estimated from CMS's historical data such that $\sum_i \pi_i = 1$. An arc connecting node i to j is denoted by x_{ij} such that $x_{ij} = 1$ *iff* the arc is connected. The domain which CMS inspects is denoted by a binary variable y_k where $k \in \{A, B, C\}$ and $y_k = 1$ *iff* the measure is selected for inspection. For each node, we introduce a binary indicator variable z_i to denote whether arcs originate from it. Thus $z_i = 0$, *iff* $\sum_j x_{ij} = 0$. In other words, at the nodes where $z_i = 0$, the star ratings of a nursing home cannot increase any more. The problem is to choose an inspection domain, such that the percentage of nursing homes that can inflate their ratings is minimized.

$$min\text{:} \sum_i \pi_i z_i \tag{3.1}$$

s.t. $$Budget\ constraint\text{:} \sum_k y_k c_k \leq C, \tag{3.2}$$

$$Ending\ node\ constraint\text{:} z_i = 1, if \sum_j x_{ij} \neq 0, \tag{3.3}$$

$$Inspection\ constraint\text{:} \quad x_{ij} = 1 - y_k, x_{ij} \in \varphi_k, \tag{3.4}$$

where φ_k is the set of arcs affected by CMS decision k.

Equation (3.1) denotes the total percentage of nursing homes that can inflate, which is the objective of this model. Equation (3.2) represents the budget constraints of CMS, where c_k is the cost for inspecting measure k, and C is the overall inspection budget. Equation (3.3) indicates whether there are arcs originating from node i. In other words, if $z_i = 1$ then nursing homes can still increase star ratings at node i. If $z_i = 0$, then nursing homes cannot improve star rating any more, and there is no arc originating from the node and therefore $\sum_j x_{ij} = 0$. Equation (3.4) denotes the relationship between arcs x_{ij} and the inspection domain y_k. For each domain k, φ_k denotes the set of arcs which represent changes in the

domain measures. For example, for domain A, arc 12, 36, 45, 78 represent the changes in measure of domain A from nodes (000), (010), (001), and (011), respectively. If domain A is inspected, $y_A = 1$, and all the x_{ij}s for the 4 arcs will be set at 0.

The logic constraint in equation (3.3) can be converted to a linear constraint and replaced by the following two constraints:

$$Mz_i \geq \sum_j x_{ij}, \tag{3.5}$$

$$z_i \leq \sum_j x_{ij}, \tag{3.6}$$

where M is a large number.

3.3.3 Minimizing the Difficulty of Detecting Inflators

The other objective of the inspection strategy is to facilitate the detection of fraud in the measures of the domains which are left to nursing homes to self-report. That is, we intend to inspect a domain such that inflating nursing homes, after inflating their scores in remaining uninspected domains, show some rare patterns in their rating combination, which makes them easier to be detected. For example, suppose a rating combination, say (001), is very common and occurs with a relatively high probability. If an inflating nursing homes achieves this rating combination after fraudulent self-reporting, it will not draw CMS's attention and its inflation of the scores will be difficult to be detected. On the other hand, if an inflating nursing home achieves a rating combination with a relatively low probability which rarely occurs, say (110), the case will be highly suspicious and relatively easier for CMS to detect.

In this formulation, we minimize the difficulty for detecting inflators. The difficulty is defined to be the product of inflator population (denoted in percentage) and the probability of the rating combination after self-reporting. The probability of showing rating combination j is denoted by p_j, and is estimated from CMS's historical data. For example, if measure k is inspected, and 10% of inflators will show a rating combination with probability of 0.2 after self-reporting, the difficulty to detect inflators when measure k is inspected is defined to be 0.1×0.2=0.02. Intuitively, we try to strategically change the

topology of the graph, such that the inflators are redirected to show some rare rating combination after self-reporting, and can be detected easily.

Following this logic, we formulate the model to minimize the total difficulty for detecting inflators. Besides the variables introduced in Section 3.2, a new binary variable v_{ij} is defined for each node pair, $v_{ij} = 1$, if j is the ending node of i, and 0 otherwise. A flow equation (3.8) is thus added to the set of constraints described in Section 3.2, and the problem formulation changes to

$$\min_{y_k} \sum_i \pi_i (\sum_j p_j v_{ij}) \tag{3.7}$$

s.t. equations (3.2), (3.3), (3.4), and

$$v_{ij} = 1, \text{ if node j is the ending node of node i, } v_{ij} = \{0, 1\}. \tag{3.8}$$

Equation (3.8) defines the relationship between node pairs when different domains are inspected. For example, v_{16} denotes whether nursing homes at node 1 will reach to node 6 after inflation. Suppose domain C is selected for inspection, and arcs 14, 25, 37 and 68 are removed from the graph shown in Figure 3-1, resulting in two unconnected sub-graphs shown in Figure 3-2. After inflating their self-reported scores, inflating nursing homes at node 1 will move to node 6, with rating combination (110), thus node 6 is the ending node for node 1, and $v_{16} = 1$. However, $v_{12} = 0$, since inflators at node 1 will not stop at node 2, but can further inflate to node 6 to achieve a 2-star improvement. In other words, node 2 is not an ending node. Also, $v_{18} = 0$, since when domain C is inspected, inflators at node 1 cannot inflate their scores in domain C, and cannot reach node 8, and thus node 1 and node 8 are not connected.

According to the discussion above, we can decompose constraint (8) into two parts:

Part 1: Node j is an ending node, with $z_j = 0$

Part 2: Node j and node i are connected.

Part 1 can be expressed as

$$v_{ij} \leq z_j. \tag{3.9}$$

To convert part 2 to a linear format, we define connections between two nodes by exploring all the paths connecting them. For example, for node 1 and node 2 to be connected, $x_{12} = 1$ is necessary and sufficient; For node 1 and node 5 to be connected, at least one of the two paths 1→2→5 or 1→4→5 must be connected. We define binary connection variable x_{125} such that x_{125}=1, iff path 1→2→5 is connected, that is, $x_{12} = 1$ and $x_{25} = 1$. The binary variable x_{145} can be defined similarly. Table 3-2 lists the logical constraints needed for describing connection between any pair of the eight nodes shown in Figure 3-1.

The logical constraints are then systematically converted to linear forms. Take connection x_{15} as an example. Equation (3.10) and (3.11) guarantee that $x_{15} = 1$, if at least one of the two paths is connected ($x_{125} = 1$ or $x_{145} = 1$), and $x_{15} = 0$, if neither path is connected.

$$Mx_{15} \geq x_{125} + x_{145}, \tag{3.10}$$

$$x_{15} \leq x_{125} + x_{145}, \tag{3.11}$$

For path x_{125}, the following linear constraints can be added to accurately describe the connection status.

$$x_{125} \geq x_{12} + x_{25} - 1, \tag{3.12}$$

$$x_{125} \leq x_{12}, \tag{3.13}$$

$$x_{125} \leq x_{25}. \tag{3.14}$$

Equation (3.12), (3.13) and (3.14) guarantee that $x_{125} = 1$, if both x_{12} and x_{25} are equal to 1, and $x_{125} = 0$, if any of them is 0. Similarly, x_{145} and other connection relationship can be converted to linear. There linear constraints are used in the formulation of the mathematical solutions.

Table 3-2. Decomposed Node Connection Logical Constraints

		Nodes							
		1	2	3	4	5	6	7	8
Nodes	1	-	x_{12}=1	x_{13}=1	x_{14}=1	x_{12}=1&x_{25}=1 or x_{14}=1&x_{45}=1	x_{12}=1&x_{26}=1 or x_{13}=1&x_{36}=1	x_{13}=1&x_{37}=1 or x_{14}=1&x_{47}=1	x_{1258}=1, or x_{1268}=1, or x_{1368}=1, or x_{1378}=1, or x_{1458}=1, or x_{1478}=1
	2		-	-	-	x_{25}=1	x_{26}=1	-	x_{25}=1&x_{58}=1 or x_{26}=1&x_{68}=1
	3			-	-	-	x_{36}=1	x_{37}=1	x_{36}=1&x_{68}=1 or x_{37}=1&x_{78}=1
	4				-	x_{45}=1	-	x_{47}=1	x_{45}=1& x_{58}=1 or x_{47}=1&x_{78}=1
	5					-	-	-	x_{58}=1
	6						-	-	x_{68}=1
	7							-	x_{78}=1
	8								-

3.3.4 Optimization Results

To test our models, we collect data on 1219 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 3 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, certificate and ownership. The *finances* dataset data 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 setting parameters in our model, for example, the probability of showing each rating combination as listed in Table 3-3, or nursing homes' financial profits.

The above two models are solved in the IBM ILOG CPLEX Optimization Studio (IBM, 2015). Table 3-4(a) shows the values of the objective function when we solve for minimizing the percentage of nursing homes that can inflate. If no inspection is conducted, every nursing home can inflate, thus the objective value will be close to 100%. Note that a small portion of nursing homes that genuinely have gained 5 stars in all three domains, do not need to further inflate their ratings, and thus the total percentage of nursing homes that can inflate is smaller than 100% even when no inspection is conducted. If every measure is inspected, no nursing home can inflate, thus the objective value will be 0.

Table 3-3. 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

Table 3-4. Inspection Measure Selection Results

Inspected Measures	ABC	BC	AB	B	AC	C	A
Objective Value	0	0.808	0.923	0.946	0.966	0.975	0.992

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

Inspected Measures	ABC	A	AC	C	AB	B	BC
Objective Value	0	0.001288	0.009363	0.015885	0.023533	0.042502	0.095107

(b) Minimizing the Difficulty to Detect Inflators

As we defined earlier, the *audit difficulty* is the product of inflating population and the probability of the rating combination at the ending node. For example, if 5% of the nursing homes can inflate their ratings to a particular node, and the probability to observe the rating combination of that node is 10%, then the audit difficulty is defined to be $0.05 \times 0.1 = 0.005$. Since we cannot estimate the cost of inspecting each domain and the total resources available to CMS, we keep the problem setting as close as possible to the CMS's current practice, and focus on the case that only one measure is inspected. As shown in Table 3-4(b), when the objective function is minimizing the percentage of nursing homes that can inflate, inspecting the Staffing domain is the optimal choice for CMS. However, when minimizing the difficulty for detecting the inflators, the current practice of conducing health infections is the optimal choice.

3.4 Audit Strategy

As discussed in Section 3, the current inspection domain is optimal for CMS in terms of detecting inflators. Unfortunately, CMS currently does not have an audit policy for the nursing homes' self-reported measures and the inflators rarely get caught. In this section, we design an audit system for CMS and conduct a one-year audit simulation based on the most recent available data for year 2013. Table 3-5 lists all the notations and symbols used in the audit simulation.

Table 3-5. 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
ΔA_1	The additional profit a nursing home can gain by improving its rating by one star
ΔA_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

Our audit design assumes that CMS continues inspecting the current domain (Health Inspection) to minimize the difficulty for detecting inflators. We assume that CMS publicly announces the following audit policy to all nursing homes:

- Nursing homes whose overall star ratings increase after self-reporting are subject to audit. An audit can distinguish honest from inflating nursing home without any errors. That is, if audited, an inflating nursing home will definitely get caught and an honest nursing will definitely get exonerated.
- Nursing homes whose rating increase 1 star 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 $7 \times (1 + r)$. The expected profit for each nursing home is calculated by using the OSHPD financial data for nursing homes.
- 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 from government financial allocation. We further assume that all penalties collected from caught inflators and the net audit budget B_0 are used in auditing nursing homes within the same year.

3.4.1 Nursing Homes' Reaction to Audit Policy

Given CMS's audit policy, nursing homes which have the intention to inflate will have to decide what to self-report. By using the latest available year (2013) in our dataset, and by applying the method presented in Han et al. (2015) using a 90% CI, we can identify the likely inflators in the population. There was no effective audit on nursing homes' self-reported measures in 2013, thus nursing homes who did not inflate their ratings are identified to be honest, and assumed to have no intention to inflate, no matter what audit

policy is used. The nursing home population is then divided into two groups: the inflators and honest nursing homes. The inflators will react to different audit policies differently, while the honest nursing homes will report the truth, no matter what audit policy is used. By using OSHPD financial data for nursing home, we can also calculate the expected profit for nursing homes in each Health Inspection star rating. We denote the additional profit for inflating one star or two stars as ΔA_1 and by ΔA_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, thus for them, ΔA_1 and ΔA_2 are equal.

Any nursing home reporting a star increase will be subject to CMS's audit. Nursing homes whose overall ratings increases by one star 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. As a result, for a given combination of p_1, p_2 and r, a nursing home considering inflating calculates its expected payoff for the following three choices:

- Being honest: $\text{Payoff}_0 = 0$
- Inflating one measure: $\text{Payoff}_1 = \Delta A_1 \ (1 - p_1) - p_1 r \ \Delta A_1 = \Delta A_1 (1 - \ p_1 \ - \ p_1 r)$
- Inflating both measures: $\text{Payoff}_2 = \Delta A_2 (1 - p_2) - p_2 r \Delta A_2 = \Delta A_2 (1 - \ p_2 \ - \ p_2 r)$

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

$$r < \frac{1-p_1}{p_1}, \text{ or} \tag{3.15}$$

$$r < \frac{1-p_2}{p_2}. \tag{3.16}$$

If either equation (3.15) or equation (3.16) is satisfied, then the nursing home will choose to inflate. If equation (3.15) and (3.16) are both satisfied, then the nursing home compares the expected payoffs, and will inflate two stars instead of one star, if

$$\Delta A_2 (1 - p_2 - p_2 r) > \Delta A_1 (1 - p_1 - p_1 r). \tag{3.17}$$

3.4.2 Objective Functions of Audit Strategy

An increase in the overall star rating is not limited to inflating nursing homes only and 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 will lead to a waste of audit resource, 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 whose ratings increase by 1 and by 2 stars are π_1 and π_2, respectively. In addition to inflating nursing homes, π_1 and π_2 also include honest nursing homes reporting rating improvements. 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 for CMS, both of which are important indicators of the performance of an audit system.

- For targeted auditing probabilities p_1 and p_2, CMS wants to minimize the initial fixed budget (B_0).
- CMS wants to maximize the efficiency of its audits, as the ratio between the percentage of caught inflators and the percentage of audited honest nursing homes. That is CMS wants more resources spent on auditing inflators and less resources wasted on auditing honest nursing homes.

For either objective used, CMS needs to make sure the following budget constraint is satisfied.

$$\pi_1(p_1, p_2, r)p_1 c + \pi_2(p_1, p_2, r)p_2 c \leq B_0 + F(p_1, p_2, r) \qquad (3.18)$$

The left-hand side of equation (3.18) denotes the total costs for auditing selected nursing homes. It includes both honest and inflating nursing homes selected to be audited. The inflators will then get caught and fined, and the collected fines will be added to CMS's initial budget, B_0, as shown on the right-hand side of equation (3.18). For simplicity, in (3.18), we assume the fines are immediately collected and can be used toward auditing more nursing homes within the same year.

3.4.3 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 instead of 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 the expected payoffs; some choose to inflate two measures, some choose to inflate one measure, 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 whose star ratings increase by one star, and p_2 percent of the nursing homes whose star ratings increase by two stars. The caught inflators are fined according to the punishment rate r. CMS keeps the total cost of audit under the initial budget B_0 plus the collected fines. We set the simulation parameters by using the 2013 data, which is the latest available in our dataset. According to the method developed in Han et al. (2015), we pre-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. Increasing the audit probabilities, and the penalties on one hand results in more inflators being caught, and more funds being collected and therefore lowers the level of required initial budget, while 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 which increases the level of required initial budget.

Complete Inflation Determent:

It is of significant importance for CMS to analyze the optimal parameter settings to completely deter inflation. According to equation (3.17), nursing homes will stop inflating when the expected payoff of inflation equals 0. The conditions are expressed in equation (3.19) and (3.20).

$$\Delta A_1(1-p_1) - \Delta A_1 p_1 r = 0, \tag{3.19}$$

$$\Delta A_2(1-p_2) - \Delta A_2 p_2 r = 0. \tag{3.20}$$

Solving equation (3.19) and (3.20), we can obtain the marginal probability for deterring inflation:

$$p_1 = p_2 = \frac{1}{1+r}. \tag{3.21}$$

Equation (3.21) is consistent with equation (3.15) and (3.16), 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 (3.21) also indicates that when p_1, p_2 are set higher than $\frac{1}{1+r}$, CMS's resource will be "wasted" on auditing honest nursing homes who report true improvements, as no nursing home will have any reason to inflate under those audit policies. Table 3-6 lists the simulation results of minimum audit probabilities, p_1, p_2 for a given penalty rate, r, and the corresponding initial budget B_0. The minimums of p_1 and p_2 obtained from the simulation are equal, which is consistent with equation (3.21). When punishment rate r increases, the corresponding minimum p_1 and p_2 needed to absolutely deter inflating decreases. This result shows a tradeoff between audit probabilities and punishment rate to achieve the same level of audit power. From Table 3-6, we also observe that when the penalty rate r increases, the corresponding net budget B_0 decreases. The reason is that when inflation is absolutely deterred, CMS collects no fine from auditing, and has to solely rely on net budget B_0 to support the audit. Under this circumstance, the lower p_1 and p_2 are, the smaller audit work load is required, thus the less net budget is needed. The results also indicate that though increasing audit probabilities and increasing penalty rate can both deter inflation, the way they function is different. When increasing audit probabilities, CMS increases the audit work load at the same time, and the audit cost, which is proportional to the audit work load in our model, will increase. On the other hand, increasing the

punishment rate has no explicit relationship with the increase of audit work load, and can be a more economical way for CMS to deter inflation.

Incomplete Inflation Determent:

Completely deterring rating inflation, though desirable, may not be feasible in reality, since the audit probability is limited by financial budget and human resource, and the maximum punishment rate is typically restricted by law. While certain level of inflation may be inevitable, the problem for CMS changes to find the optimal parameter settings under the given net budget. Since audits impose unnecessarily burdens on honest nursing homes, it is desirable that CMS can 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 within the interval [0, 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 PCIs.

To study the properties of the audit efficiency, we focus on the upper and lower limits of PCI, as the maximum 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., 2015). 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., $\forall B_{0x}, B_{0y} \in \{0, R^+\}$, $B_{0x} \neq B_{0y}$, $\max\limits_{p_1,p_2,r} PCI(PAH, B_{0x}) = \max\limits_{p_1,p_2,r} PCI(PAH, B_{0y}) = \max\limits_{p_1,p_2,0} PCI(PAH)$.

Proposition 2:

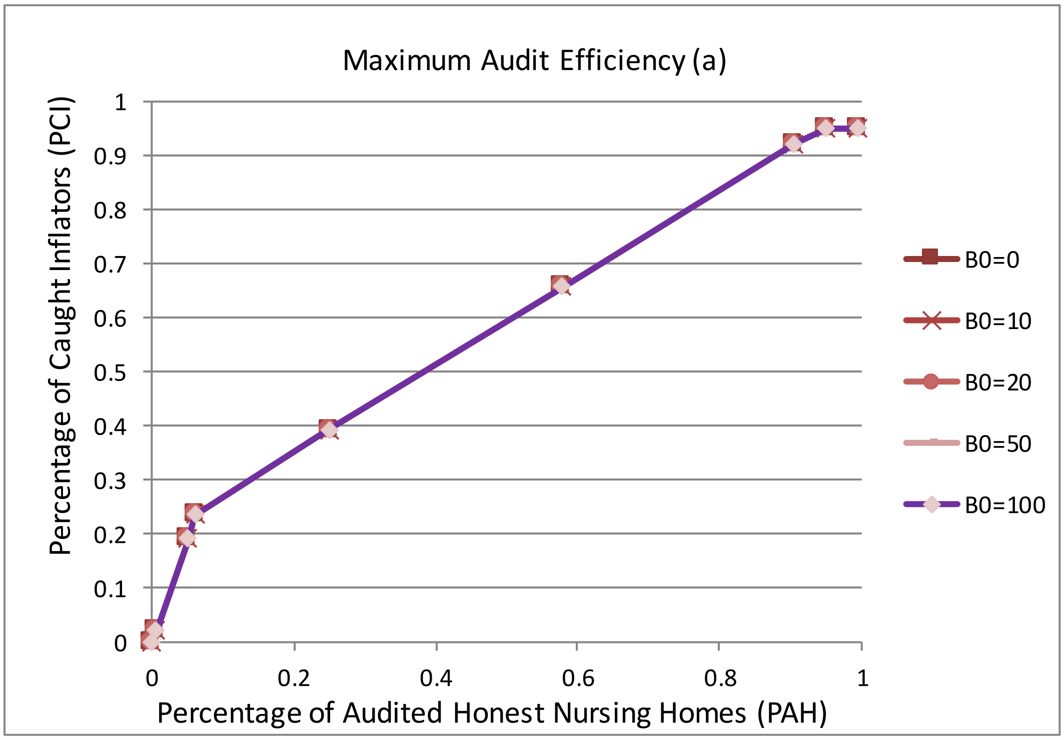
(2.1) Given p_2 and r, the maximum audit efficiency converges to p_2 monotonically when $p_1 \to 1$.

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

(2.3) If $\Delta A_2 \cdot p_2 - \Delta A_1 \cdot p_1 < 0$, or $\frac{p_1}{p_2} > \frac{\Delta A_2}{\Delta A_1} > 1$, then the maximum audit efficiency does not change with respect to r. If $\frac{p_1}{p_2} < \frac{\Delta A_2}{\Delta A_1}$, 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.

Figure 3-3. Efficiency Curves given different net budget B_0



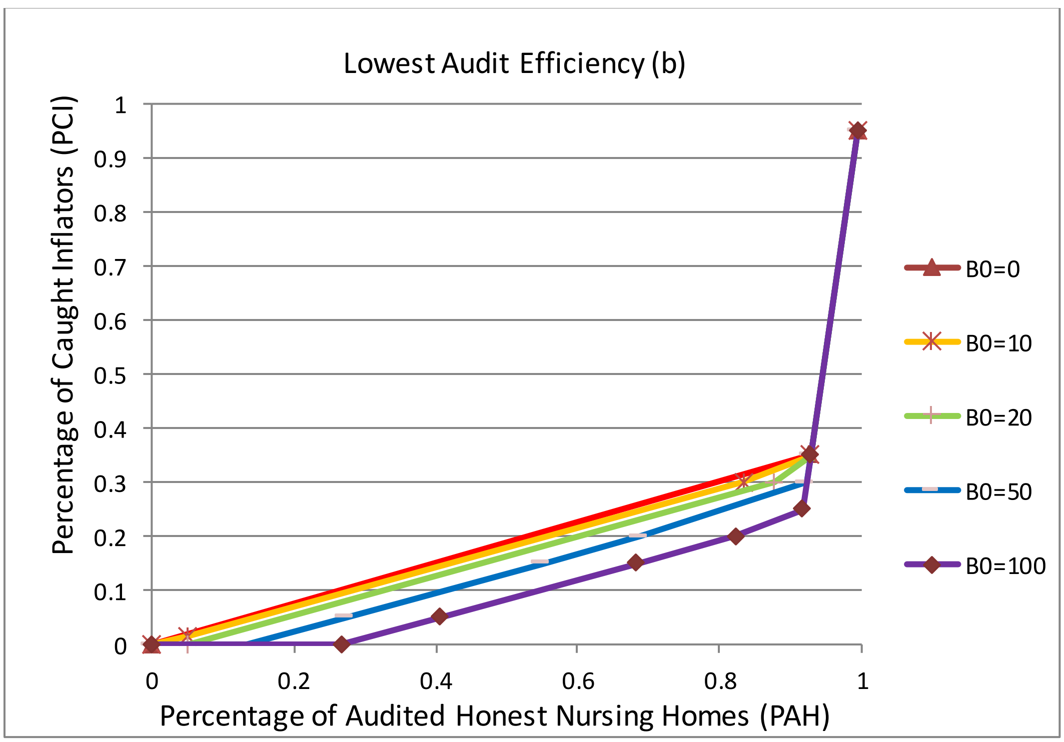


Table 3-6. Optimal Policy Parameters for Deterring Inflation

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

We justify the propositions using simulation results. Figure 3-3(a) shows the maximum audit efficiency CMS can achieve given different net budget B_0. On the contrast, Figure 3-3(b) shows the lowest audit efficiency CMS can achieve.

According to Figure 3-3, the maximum audit efficiency is achieved at r=0, and increasing B_0 cannot lead to an increase in the audit efficiency. When r=0, nursing homes have no consequence when caught inflating, thus all the non-honest nursing homes that may inflate will choose to inflate. As a result, the inflator proportion in the suspect group reaches the maximum, and the audit efficiency 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, but can only put more honest nursing homes under audit. On the other hand, the lowest audit efficiency will become even lower when B_0 increases. This is because for the same percentage of caught inflators (PCI), more honest nursing homes are audited (higher PAH). From Figure 3-3, we also know that the maximum audit efficiency curve is piecewise linear and monotonically non-decreasing. These properties on the maximum audit efficiency curve is also consistent with the results derived in Proposition 2.

Combining the conclusion obtained from Table 3-6 in Section 3.4.3, we reach the following conclusion on audit parameter settings: If CMS has enough net budget to completely deter inflation, then adopting an audit policy with higher penalty rate is more economical than increasing the audit probabilities, and will lead to a lower net budget. However, under high punishment rate, few nursing homes will choose to inflate. As a result, most of the resources will be spent on auditing honest nursing homes, leading to a low audit efficiency. Increasing audit probabilities can also deter inflation, but it will increase the total cost of audit as well. If CMS does not have enough net budget to completely deter inflation, and certain level of inflation is inevitable, the problem is to find an optimal combination of p_1, p_2 and r for the given net budget. Under this circumstance, we recommend CMS to carefully balance the tradeoff between net budget and audit efficiency; Specifically, the punishment rate should be relatively high in order to keep the net audit budget within control; however, the punishment rate cannot be too high

in order to achieve a good audit efficiency, i.e., a good amount of money should be spent on auditing inflators, not on auditing honest nursing homes who made real service improvements.

3.5 Conclusion

In this chapter, we systematically investigate the inspection measure selection and audit design in the CMS's nursing home rating system. For the inspection measure selection problem, we formulate it mathematically by using an innovative graph-based framework, and solve it optimally by using CMS's data. Our result indicates that CMS's current inspection domain is optimal if an effective follow-up audit policy is in place. We then design the audit system for CMS, with nursing homes' different reactions taken into consideration. To study the parameter settings of the audit, we first derive analytical conditions on the parameter settings, and then conduct a simulation to justify the results. We find that CMS has to carefully balance the tradeoff between net audit budget and audit efficiency in order to achieve a good control of self-reporting inflation.

Due to the assumptions and the dataset, this research may have the following limitations. In the inspection problem, it is difficult to accurately measure the cost for inspecting each measure. Our result is obtained based on the assumption that the cost for inspecting each measure is similar and comparable. The real situation may depend on several factors, leading to different cost for inspecting each measure, e.g., regulations for each area may be different, and the cost for investigating each nursing home is case dependent. Our result, however, is a reasonable estimate of the current system, and provides a theoretical framework for optimizing the inspection structure. In the audit design, we assume that the fines can be collected immediately and used in audit within the same year. Since fine collection takes time, and CMS may have different arrangement of the fines, it is possible that not all fines are available to use within the same year. Suppose that only a portion, say α, of the collected fines can be used in the same year for auditing. In this case, the audit can be considered to have an adjusted punishment rate $r'=r\alpha$. The quantitative results we obtained, for example B_0, will be different in this case, but most qualitative results should still hold. In view that no audit is currently conducted for nursing homes' self-reporting, our

design provides a reasonable plan for CMS, which can serve as a benchmark for other audit system designs. The audit simulation is conducted based on the 2013 data of the California nursing home market. For other states or other years, there can be difference in the proportion of inflators. However, we argue that the population of nursing homes in California is big and the case is representative, and the number of inflators in year 2013 is reasonably close to the amount in the current system.

Our result not only provides insights and theoretical support for CMS's current rating system but also provides guidelines for the future audit mechanism design. Moreover, our result has important managerial application on other rating systems sharing similar features. A good example is to apply the results on the Merit-Based Incentive Payment System (MIPS), which rates the physicians based on a combination of inspection and their self-reported measures, and adjusts the payment rates based on the ratings. With the fast development of healthcare IT, justifiable self-reported measures can have a wide application in different rating systems in order to reduce the cost to evaluate these measures. Our research provides a good framework to systematically control the quality of self-reported measures, which guarantees the accuracy of online ratings.

Appendix 3

Proof for Proposition 1:

We first divide the nursing home population into 4 different types, as shown in Figure 3-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 intension to inflate make their decisions based on the expected payoffs of the following three cases:

0. Not Inflate. $\Pi_0=0$

1. Inflate 1 star. $\Pi_1=\Delta A_1 (1-p_1-p_1 r)$

2. Inflate 2 stars. $\Pi_2=\Delta A_2\,(1-p_2-p_2r)$

In our problem, we have $\Delta A_1 \leq \Delta A_2$ for all rating levels.

Given r and p_2, and set $p_1=0$, $\Pi_2 = \Delta A_2(1-p_2-p_2r)$, denote the profit when $p_1=0$ as $\Pi_{10} = \Delta A_1$

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.

The breakpoints is $\Pi_2 = A_2(1-p_2-p_2r) = \Delta A_1(1-p_1-p_1r) = \Pi_1$, i.e.,

$$p_1 *= \frac{\Delta A_1 - \Delta A_2 + \Delta A_2(1+r)p_2}{\Delta A_1(1+r)}$$

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

PCI= p_1, when $p_1 \leq p_1 *$

p_2, when $p_1 > p_1 *$

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

$$p_1 * - p_2 = \frac{\Delta A_1 - \Delta A_2 + \Delta A_2(1+r)p_2 - \Delta A_1(1+r)p_2}{\Delta A_1(1+r)} = \frac{\Delta A_1 - \Delta A_2}{\Delta A_1(1+r)}[1-(1+r)p_2]$$

Since $\frac{\Delta A_1 - \Delta A_2}{\Delta A_1(1+r)} \leq 0$, and $1-(1+r)p_2 > 0$ when $\Pi_2>0$, thus $p_1 * - p_2 \leq 0$, i.e., $p_1 * < p_2$ always holds when $\Pi_2>0$. In other words, PCI is monotonically non-decreasing in p_1.

(2.2) Given r and p_1, and set $p_2=0$, $\Pi_1 = \Delta A_1(1-p_1-p_1r)$, $\Pi_{20} = \Delta A_2 > \Pi_1$, and $PCI_0=p_2$. When p_2 increases, Π_2 decreases.

The breakpoints is $\Pi_2 = \Delta A_2(1-p_2-p_2r) = \Delta A_1(1-p_1-p_1r) = \Pi_1$, i.e.,

$$p_2 *= \frac{\Delta A_2 - \Delta A_1 + \Delta A_1(1+r)p_1}{\Delta A_2(1+r)}$$

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

PCI= p_2, when $p_2 \leq p_2*$

p_1, when $p_2 > p_2*$

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

$$p_2 * - p_1 = \frac{\Delta A_2 - \Delta A_1 + \Delta A_1(1+r)p_1 - \Delta A_2(1+r)p_1}{\Delta A_2(1+r)} = \frac{\Delta A_2 - \Delta A_1}{\Delta A_2(1+r)}[1-(1+r)p_1]$$

Since $\frac{\Delta A_2 - \Delta A_1}{\Delta A_2(1+r)} \geq 0$, and $1-(1+r)p_1 > 0$ when $\Pi_1>0$, 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 $\Pi_{20} = \Delta A_2(1-p_2)$, $\Pi_{10} = \Delta A_1(1-p_1)$,

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

$$\Delta A_2(1-p_2) > \Delta A_1(1-p_1)$$

$$\Delta A_2 - \Delta A_1 > \Delta A_2 p_2 - \Delta A_1 p_1$$

Since $\Delta A_2 - \Delta A_1 \geq 0$, if $\Delta A_2 p_2 - \Delta A_1 p_1 < 0$, then $\Pi_{20} > \Pi_{10}$ holds.

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

$$\Delta A_2(1-p_2-p_2 r) > \Delta A_1(1-p_1-p_1 r)$$

Or

$$\Delta A_2 - \Delta A_1 > (\Delta A_2 p_2 - \Delta A_1 p_1)(1+r)$$

Case 1: If $\Delta A_2 p_2 - \Delta A_1 p_1 < 0$, when r increases, the RHS 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 $\Delta A_2 p_2 - \Delta A_1 p_1 > 0$, three subcases can be discussed.

a: If $p_1>p_2$, then $\Pi_{20} > \Pi_{10}$, the payoff functions are shown in Figure 3-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 3-A.1 (b).

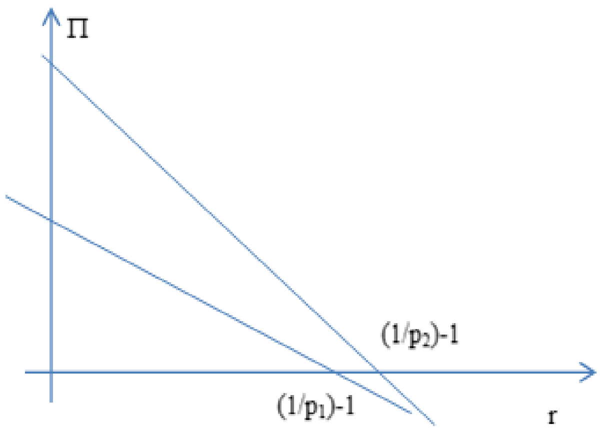
In this case, when $r < \frac{\Delta A_2 - \Delta A_1}{\Delta A_2 p_2 - \Delta A_1 p_1} - 1$, then $\Pi_1 < \Pi_2$, and PCI = p_2.

When $r > \frac{\Delta A_2 - \Delta A_1}{\Delta A_2 p_2 - \Delta A_1 p_1} - 1$, then $\Pi_1 > \Pi_2$, and PCI = p_1.

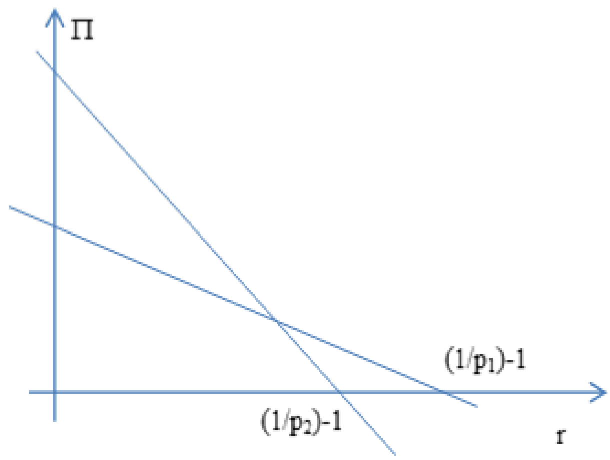
c: If $p_1<p_2$, then when $\Pi_{20} < \Pi_{10}$, the payoff functions are shown in Figure 3-A.1(c).

In this case, $\Pi_1 > \Pi_2$ always holds , and PCI = p_1, and will not change.

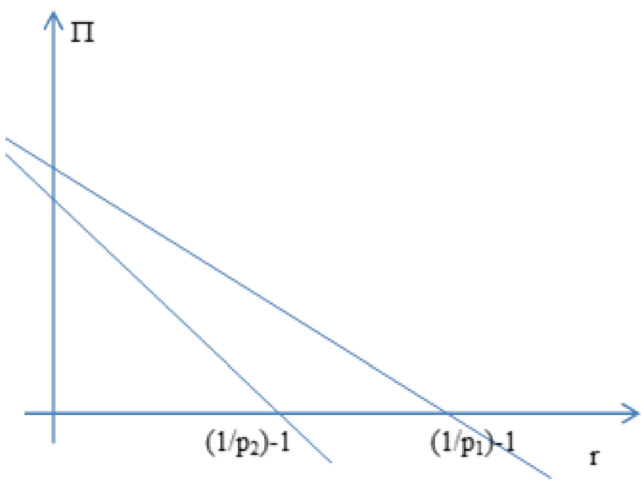
Figure 3-A.1. Three Cases of the Profit Function



(a)



(b)



(c)

Table 3-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.

Chapter 4. Once for Ado: The Impact of Internet of Things Adoption and Subsidization

4.1. Introduction

Nursing homes are an important piece in the U.S. healthcare system. Today there are over 16000 nursing home facilities in the U.S., providing care to 1.5 million residents. They account for 6 percent of the Medicare population but 17 percent of total Medicare spending. The service quality of nursing homes, however, varies significantly. In 2007, the Centers for Medicare & Medicaid Services (CMS) designed and implemented its 5-star nursing home rating system. Since its inception, the CMS rating has become the gold standard in the industry, and has been widely popular among patients, physicians and payers (Thomas, 2014). The CMS 5-star rating system rates nursing homes based on three measures: *on-site Inspection*, *Staffing* and *Quality Measures*. The on-site inspections are conducted by CMS-certified inspectors, while the other two domains are self-reported by nursing homes. The overall star ratings is calculated by using the inspection star rating as a baseline, adding 1 star if any self-reported domain is 5 stars and subtracting 1 star if any self-reported domain is 1 star (CMS, 2015). Obviously, the two self-reported domains can change a nursing home's rating fundamentally. An average nursing home which receives 3 stars in CMS inspection can have an excellent 5-star overall rating, if its two self-reported measures are 5 stars. Cases have been reported in which patients' experiences differ significantly from the CMS star ratings. Some highly-rated nursing homes are sued for substandard care, even causing death of patient due to improper medical treatments. Our previous research has found a significant positive association between nursing homes' star rating change and their financial incentives to achieve high rating (Han 2015), and this association cannot be simply explained by nursing homes' legitimate efforts. In other words, rating inflation does exist in the current system.

In recent years, the emergence of the Internet of Things (IoT) concept has gradually drawn the public attention, and presents new opportunities for research. Internet of Things (IoT) typically refer to a

system of interrelated sensors, objects, mobile devices, mechanical and digital machines that are provided with unique identifiers and are able to transfer data over a network without requiring human-to-human or human-to-computer interaction. Technologies based on IoT concepts have been successfully developed in the healthcare field, especially in the senior community and nursing homes. For example, tracking devices with IoT technologies have been introduced to track the movement of senior patients and detect sudden falls. Although these IoT applications in nursing homes are still relatively new and non-systematical, the idea of applying IoT and related technologies is getting more and more popular, and can be a promising way to improve nursing homes' service quality.

The adoption of IoT technologies in nursing homes could significantly change the way nursing homes self-reports on both domains. Starting July 2016, CMS's requires all nursing homes to report Payroll- based journal (PBJ), and will relate these results with the self-reported staffing level (CMS, 2016a). Meanwhile, in the QM domain, CMS is going to add 5 additional measures (out of the 6 newly posted measures) to the current 9 measures (out of 18 measures) when giving the QM rating (CMS, 2016b). These requirements impose huge burdens on nursing homes' daily operations, leading to additional operational costs. The above staffing and quality measure data, however, can be collected automatically by using devices based on IoT technologies (Stenner 2011, Krawiec 2015). As a result, IoT adoption could possibly release nursing homes from the burdens of self-reporting.

Another potential use of IoT devices is to control self-reporting inflation. When IoT devices are used, the data about the corresponding field which is previously self-reported is now automatically collected by these devices. As a result, these potential inflators will have less room to inflate during self-reporting, leading to more robust data collection.

From CMS's perspective, IoT adoption is not only a promising way to improve nursing homes' service quality, but also serves an alternative way to control rating inflation comparing with the traditional audit method. However, IoT techniques can be costly, which can be a major barrier for IoT adoption. On the other hand, auditing an honest nursing home not only imposes huge burden to the nursing home, but

also wasted limited financial budget and human resources. The fact triggers the interesting question that whether CMS can allocate certain amount of audit budget to subsidize nursing homes for IoT adoption. The subsidization gives nursing homes incentives to adopt IoT techniques, and reduces the ability that potential inflators can inflate, thus IoT subsidization can be a win-win situation for both nursing homes and CMS.

In this chapter, we systematically investigate the impact of IoT adoption on both the CMS rating system and on nursing homes' decisions. The chapter proceeds as follows: In Section 4.2, we analyze the optimal staffing and IoT adoption level decisions for both honest nursing homes and inflators. The results indicate that inflators are more reluctant to adopt IoT than the honest ones, since IoT adoption limits the ability of inflating. We also find diversified staffing level reaction to IoT adoption for nursing homes with different service qualities. We then analyze how the optimal IoT adoption level can be affected by CMS's auditing and subsidization. In Section 4.3, we analyze the problem from CMS's perspective, and obtain insights for CMS on allocating budget between auditing and subsidization. The chapter is concluded in Section 4.4.

4.2 Nursing Homes' Problem Formulation

In this section, we first derive the framework for nursing homes' profit maximization problem. The general problem formulation for a nursing home's problem is presented. We then analyze honest nursing homes and inflators separately, and derive optimal IoT adoption and staffing decisions for both types.

4.2.1 General Formulation

We first build a general analytical model for nursing homes' profit maximization problem[3]. A list of symbols used in this formulation can be found in Table 4-1. As discussed in the literature, the staff-to-

[3] Though not all the nursing homes are for-profit nursing homes, for-profit nursing homes are the majority. According to CMS data, at least 60% of nursing homes are for-profit nursing homes. The remaining nursing homes are non-profit nursing homes or government owned nursing homes, which also pursue surplus income as the for-profit ones, but they distribute the surplus income as dividends rather than profits. In this research, we do not

patient ratio has been proved to be a major factor affecting the service quality of a nursing home (Lu 2015, Konetzaka 2008). In our model, we use s to denote the staff-to-patient ratio[4], i.e., the per patient staffing level. Depending on the type of the nursing home, we assume $s \in (0, \bar{s})$, where $\bar{s}$ is the maximum staffing level nursing homes can have. Besides staffing coverage, IoT adoption level is another key factor affecting the service quality of nursing homes. In our model, we use k to denote the IoT adoption rate, $k \in [0, 1]$, i.e., a nursing home can choose not to use any IoT technology, choose to fully rely on IoT devices to collect data and do self-reporting, or choose to adopt IoT technologies at any level k in order to maximize its profit. The two decisions, staffing level and IoT adoption level, may also affect each other. The service quality of a nursing home is then denoted as $Q(s, k)$.

---

differentiate the difference between for-profit nursing homes and non-profit or government owned nursing homes, and assume all nursing homes maximize its profits as its goal.

[4] To keep the formulation consistent, we use the per patient staffing level, profit/revenue as the calculating unit throughout this dissertation.

Table 4-1 Symbols in the Problem Formulation

R	*Nursing home's revenue*
C_s	*Staffing cost*
C_k	*IoT adoption cost*
U_k	*IoT subsidization*
A	*Highest revenue a nursing home with staffing level s and IoT level k can achieve*
α	*Nursing home revenue vertical position factor (different for each nursing home)*
s	*Nursing home's staffing level*
k	*Nursing home's IoT adoption level*
r	*Punishment rate*
p	*Audit probability*
u	*Subsidization rate*
w	*Staffing wage*
v	*IoT unit adoption cost*
ΔA	*Maximum inflation gain*
C_{Aud}	*Unit audit cost*
N	*Total Population*
N_{II}	*Inflator population*
N_{HI}/N_{HN}	*Population of honest NHs whose rating increase/do not increase*

Besides service quality, nursing homes' vertical position in its competing market can also affect its revenue function. The vertical position of a nursing home can be viewed as the reputation of a nursing home, which is closely related to its CMS star rating. It has been shown in Han et al (2015) that highly rated nursing homes have significantly higher revenues than their poorly rated counterparts. In our formulation, we use α to denote a nursing home's vertical position in the market, $\alpha > 0$. For each nursing home, α is different. We use A to denote the highest revenue a nursing home can achieve. A nursing home's revenue R can then be expressed as a function of its service quality Q and its vertical position α as $R(Q, \alpha)$. Since Q is a function of s and k, we can equivalently express the revenue as $R(s, k, \alpha)$, a function of staffing level s, IoT adoption level k, and vertical position α. Since α is exogenous, we simplify the notation and use $R(s, k)$ to denote the revenue function of a nursing home. In our model, we require R to be increasing in s and k, i.e.,

$$R(s, k + \Delta k) > R(s, k), \tag{4.1}$$

$$R(s + \Delta s, k) > R(s, k). \tag{4.2}$$

Further, we require R to be capped by the most profitable type of patients, i.e., R is concave in both k and s, i.e., $\frac{\partial^2 R}{\partial k^2} < 0,\ and \frac{\partial^2 R}{\partial s^2} < 0,\ and\ \ lim \frac{\partial R}{\partial s} = 0,\ lim \frac{\partial R}{\partial k} = 0$. The assumption is based on the fact that no matter how high a nursing home's staffing level is, or how advanced a nursing home's IoT adoption level is, the revenue cannot go to infinity. There is always a margin patient whose net utility of going to the nursing home is 0, i.e., it is indifferent for the patient to go to the nursing home or not.

The operational costs of nursing homes are divided into two parts. The cost associated with staffing level is denoted as C_s, which meaning refers to wages paid to nurses and administration staff. In our model, re require C_s to be increasing and convex in s, i.e., $\frac{\partial C_s}{\partial s} > 0,\ \frac{\partial^2 C_s}{\partial s^2} > 0$. Another part of cost is associated with IoT adoption, and is denoted as C_k. This part mainly includes the maintenance of IoT devices. Similar to C_s, in our model, we require C_k to be increasing and convex in k.

In this research, we propose CMS to consider subsidizing nursing homes who adopt IoT technologies. We use function U_k of k to denote the amount CMS sets to subsidize IoT adoption. Similar to the cost functions C_s and C_k, we require U_k to be also increasing and convex in k, i.e., $\frac{\partial U_k}{\partial k} > 0$, $\frac{\partial^2 U_k}{\partial k^2} \geq 0$.

A nursing home maximizes its profit by optimizing its staffing level s and IoT adoption k. A general form of a nursing home's profit function is then given in equation (4.3).

$$\pi = R(s,k) - C_s - C_k + U_k, \tag{4.3}$$

where C_s is the staffing cost, C_k is the IoT adoption cost, and U_k is the IoT adoption subsidization which is associated with the IoT adoption level. If a nursing home does not adopt IoT technologies in its operation, then $k = 0$, and the nursing home has complete freedom on self-reporting decisions, though they can be audited after self-reporting. In this case, the nursing home incurs zero cost on IoT technologies, and do not receive any subsidization for CMS, i.e., $C_k = 0$, and $U_k = 0$. If a nursing home fully adopt IoT technologies, and completely rely on IoT technologies to collecting data for the self-reporting measures, then k=1. In this case, the nursing home's IoT cost C_k reaches the maximum, but the subsidization U_k also reaches the maximum.

It is ideally to obtain close form solutions on nursing homes' optimal decision. In the following analysis, we assume the following function forms for revenue function R, staffing cost C_s, IoT adoption cost C_k, and subsidization function U_k. Our functions are in line with the existing literature on nursing homes' staffing problems (Lu et al., 2015).

$$R(s,k) = A - \alpha e^{-\alpha sk}, \tag{4.4}$$

$$C_s = ws, \tag{4.5}$$

$$C_k = vk, \tag{4.6}$$

$$U_k = uk \tag{4.7}$$

in which A denotes the maximum revenue a nursing home can achieve for its rating level, and α is the revenue discount coefficient. It can be easily verified that $R(s, k)$ satisfies all the assumptions we made over revenue function, i.e., $\frac{\partial R}{\partial s} > 0$, $\frac{\partial R}{\partial k} > 0$, $\frac{\partial R}{\partial Q} > 0$, $\frac{\partial^2 R}{\partial Q} < 0$ and $lim_{Q \to \infty} \frac{\partial R}{\partial Q} = 0$. Following most literature, we use a linear form of staffing cost function, which is also close to the reality. We assume a linear form of the IoT cost function, in which C_k is proportional to the IoT adoption level k. We also assume a linear form of the subsidization function U_k, which is proportional to the nursing home's IoT adoption level k. In view that the QM domain consists of a bunch of detailed measures on each of the major aspects on patients' daily life and can be itemized, the linear assumption on IoT cost function and subsidization function is also reasonable.

In the following, we analyze nursing homes' profits in extreme cases of s and k combination. If IoT adoption is k = 1, and s is big, i.e., the nursing home's IoT technology level is advanced and staffing level is very high, then the revenue R approaches A, which is the maximum revenue the nursing home can achieve. If s is very small (close to 0), i.e., the facility is understaffed, then the revenue R approaches A-α, the lower limit of revenue. The case of k = 0 can be analyzed similarly, and it represents the case that the IoT technology level in the nursing home is very low.

The general form of a nursing home's profit maximization problem can then be written as

$$\max_{s,k} \pi = R(s,k) - C_s - C_k + U_k = A - \alpha e^{-\alpha sk} - ws - vk + uk.. \tag{4.8}$$

Since honest nursing homes and inflators have different objectives when adopting IoT, we discuss these two types of nursing homes separately in the following.

4.2.2 Nursing Homes' Optimal Decisions

Honest Nursing Homes: Honest nursing homes looking for service improvement may embrace IoT since these devices and techniques can improve the quality of care they can provide. For an honest nursing home, its profit maximization problem can be written as

$$\max \pi_H = R(s,k) - C_s - C_k + U_k = A - \alpha e^{-\alpha sk} - ws - vk + uk. \tag{4.9}$$

By taking partial derivatives of the profit function with respect to the decision variables s and k, we can solve equation (4.9), and obtain the following result.

$$\frac{\partial \pi}{\partial k} = s\alpha^2 e^{-\alpha sk} - v + u = 0 \tag{4.10}$$

$$k *= \frac{1}{\alpha s} ln \frac{s\alpha^2}{v-u} \tag{4.11}$$

$$\frac{\partial \pi}{\partial s} = k\alpha^2 e^{-\alpha sk} - w = 0 \tag{4.12}$$

$$s *= \frac{1}{\alpha k} ln \frac{k\alpha^2}{w} \tag{4.13}$$

Inflating Nursing Homes: For inflators, they can also enjoy the benefits that IoT can bring to honest nursing homes, but they have additional concerns, since IoT adoption will squeeze the room that they can misreport, and make their inflation more difficult. As a result, the inflating nursing homes have to take different audit policies into consideration, and may have different attitude to IoT adoption comparing with their honest counterparts.

Though CMS currently does not have an audit in place, it has been demonstrated in the previous chapter that audit is necessary in order to control inflation. In this chapter, we keep the audit design consistent with our previous work. In the audit, we assume CMS pre-announces its audit policy, with a proportion p and punishment rate r. If nursing homes report increased star rating in self-reporting, they will be randomly selected based on the announced proportion p. If an inflator is selected for audit, it will be caught. Its illegitimate profit gain will be confiscated, and a fine will be issued against it, which is calculated based on the punishment rate r. The inflators will then calculate their expected payoffs based on the policy announced and decide whether to inflate or not. Define the maximum profit gain through inflation to be ΔA. In our model, ΔA is achieved only when the inflator does not adopt any IoT technology, i.e., k = 0. According to our assumption, when a nursing home fully adopt IoT technologies, the nursing homes' data collection will completely depend on IoT devices, and there will be no room for

the nursing home to inflate, i.e., $\Delta A = 0$, when k = 1. We thus use a linear function to denote a inflating nursing home's illegitimate maximum revenue *A'*.

$$A' = A + (1 - k)\Delta A \tag{4.14}$$

Apparently, A'>A and the illegitimate profit takes values in [0, ΔA].

An important question to ask is that under what condition will the inflators choose to inflate. To provide answer to this question, we compare the expected revenue of an inflator when it is inflating, and the revenue of the same nursing home when it chooses to stay honest, as shown in equation (4.15). An inflator will inflate only when the expected revenue of inflating is higher than that of being honest.

$$p[A - \alpha e^{-\alpha s k} - r\Delta A(1-k)] + +(1-p)\left(A + \Delta A(1-k) - \alpha e^{-\alpha s k}\right) > A - \alpha e^{-\alpha s k}, \tag{4.15}$$

Solving equation (4.15), we can obtain the inflating condition:

$$1\text{-p-pr}>0, \text{ or} \tag{4.16}$$

$$p < \frac{1}{1+r}. \tag{4.17}$$

In other words, the inflators make inflation decisions according to the announced audit policy (p,r) only. If p and r satisfy the condition listed in (4.17), the expected revenue of inflating will exceed the revenue of being honest, and inflator will choose to inflate[5]. The IoT adoption can reduce the extent to which the inflator can inflate, but cannot change the inflators' decisions. In order to avoid trivial cases that all nursing homes are deterred from inflating, our model, we assume that inflating condition (4.17) is always satisfied, i.e., there are always inflators in the system, and they will always choose to inflate. The population of the inflators is assumed to be constant.

For inflators, its profit maximization can be expressed as:

[5] Different from the settings in chapter 3, we use a simplified model in this research in which we do not consider the possibility that some inflating nursing homes may only inflate 1 star, though inflating 2 stars is possible. Such case exists in reality due to the fact that the probability to be audited when inflating 2 stars can be set higher than that of inflating 1 star. There can be other legal costs, which deter nursing homes from inflating to the maximum level though possible. However, in our model, we do not capture these aspects, and assume that the punishment to inflators are purely financial.

$$\max \pi_I = p[A - \alpha e^{-\alpha sk} - r\Delta A(1-k)]$$

$$+(1-p)\left(A + \Delta A(1-k) - \alpha e^{-\alpha sk}\right) - ws - vk + uk \tag{4.18}$$

The optimal staffing level and IoT adoption level of an inflator can thus be calculated.

$$\frac{\partial \pi}{\partial k*} = -\Delta A(1-p-pr) + \alpha^2 e^{-ks} s + u - v = 0 \tag{4.19}$$

$$k* = \frac{1}{\alpha s} ln \frac{s\alpha^2}{v-u+\Delta A(1-p-pr)} \tag{4.20}$$

$$\frac{\partial \pi}{\partial s*} = k\alpha^2 e^{-\alpha sk} - w = 0 \tag{4.21}$$

$$s* = \frac{1}{\alpha k} ln \frac{k\alpha^2}{w} \tag{4.22}$$

Summarizing the results for honest nursing homes and inflators, we have Proposition 1 as stated below:

Proposition 1*: NH's optimal IoT and staffing decisions*

a) For both honest and inflating nursing homes, its optimal staffing level is $s* = \frac{1}{\alpha k} ln \frac{k\alpha^2}{w}$.

b) The optimal IoT adoption level for an honest nursing home is $k* = \frac{1}{\alpha s} ln \frac{s\alpha^2}{v-u}$. *The optimal IoT adoption level for an inflator is* $k* = \frac{1}{\alpha s} ln \frac{s\alpha^2}{v-u+\Delta A(1-p-pr)} < \frac{1}{\alpha s} ln \frac{s\alpha^2}{v-u}$, *if inflation condition is satisfied, i.e., 1-p-pr>0.*

c) The optimal IoT adoption level k and staffing level s* are increasing in* α.

Proposition 1 indicates that when a nursing home inflates its self-reporting to gain more profits, its optimal IoT adoption level is less than the IoT level should it be honest. The result is reasonable, since IoT limits a nursing home' inflating ability. It is also noted that in this model, we require u≤v, so that the natural log function is defined. In other word, the subsidization makes IoT adoption less costly to nursing homes, but nursing homes are not awarded more money then what they spent for adopting IoT. It is also

observed that the optimal IoT adoption level k* and optimal staffing level s* are increasing in α. According to our assumption, the service quality also increases in α, so is the revenue.

4.2.3 Nursing Homes' Staffing Level Reaction to IoT Adoption

In the previous section, we analyzed nursing homes' optimal staffing level and IoT adoption level decisions. The two decisions, however, are not exogenous, and may affect each other. Empirical evidence has been reported in Zhang et al. (2016) and Lu et al. (2015) that the adoption of technologies has different impacts on staffing levels. In this section, we analyze how nursing homes' optimal staffing level will change with respect to changes in IoT adoption levels for honest nursing homes and inflators, and obtain insights on the relationship between optimal staffing level and optimal IoT adoption level.

According to equation (4.13) and (4.22), for both honest nursing homes and inflators, we have

$$\frac{\partial s*}{\partial k} = \frac{1}{\alpha k^2}(1 - ln\frac{\alpha^2 k}{w})>0, \tag{4.23}$$

then we have

$$\alpha < \sqrt{\frac{we}{k}}. \tag{4.24}$$

In other words, an increase in nursing home's IoT adoption level will lead to an increase in staffing levels for nursing homes if $\alpha < \sqrt{\frac{we}{k}}$, and will lead to a decrease in staffing levels for nursing homes if $\alpha > \sqrt{\frac{we}{k}}$. The threshold is $\alpha = \sqrt{\frac{we}{k}}$.

Proposition 2: *The impact of IoT adoption on nursing homes' staffing decision.*

For nursing homes satisfying $\alpha < \sqrt{\frac{we}{k}}$*, IoT adoption will increase its optimal staffing level. For nursing homes satisfying* $\alpha > \sqrt{\frac{we}{k}}$*, IoT adoption will decrease its optimal staffing level. The conditions hold for both honest nursing homes and inflators.*

Since the optimal revenue at optimal staffing level $\frac{1}{\alpha k} ln \frac{k\alpha^2}{w}$ is $R *= A - \alpha e^{-\alpha s*k} = A - \frac{w}{\alpha k}$. We have $\frac{\partial R*}{\partial \alpha} = \frac{w}{k\alpha^2} > 0$. As discussed in the model introduction part, α can be viewed as a vertical position indicator for nursing homes in its competing market, i.e., for highly rated nursing homes with big α, their revenue R is higher.

Proposition 2 thus indicate that for highly ranked nursing homes, increasing IoT will lead to a decreased staffing level s, while for poorly rated nursing homes, increasing IoT will lead to an increased staffing level s. This conclusion is in line with the ones found in Lu et al. (2015).

The different reaction of staffing level to IoT adoption can be interpreted in the following way. IoT adoption improves the efficiency for the staff, thus the marginal quality from more staff increases, which brings in more staff. This complementary effect of IoT adoption on staffing level. On the other hand, since we have R concave in k and $\frac{\partial^2 R}{\partial k^2} < 0$, if service quality is high enough, the increased revenue resulted from additional IoT usage will eventually be overwhelmed by the increasing staff wage, leading to a reduced staffing level when IoT level increases. This is the substitution effect of IoT adoption on staffing level. The two effects dominate each other on nursing homes with different vertical positions in competing markets, leading to a diversified staffing decision on IoT adoption. As a result, adopting IoT may or may not bring additional profits to the nursing homes, depending on the adoption cost and other factors.

4.2.4 Controlling IoT by Auditing and Subsidization

Introducing IoT to nursing homes can not only increase service quality, but is also an alternative way for CMS to control inflation in self-reporting other than using traditional audit method. However, due to investment cost and unforeseen changes in staffing and operations, nursing homes may not have incentives to adopt IoT, and subsidization may be needed in order to align nursing homes' incentives for IoT adoption.

In this section, we derive conditions to obtain insights on how CMS can control IoT adoption by auditing and subsidization, and in turn control inflation. We calculate the partial derivative of the optimal IoT adoption level of honest nursing homes and inflators with respect to the CMS's audit parameter p[6] and subsidization parameter u.

For honest nursing homes, $k* = \frac{1}{\alpha s} \ln \frac{s\alpha^2}{v-u}$, thus

$$\frac{\partial k*}{\partial u} = \frac{1}{\alpha s(v-u)} > 0 \tag{4.25}$$

since v-u>0, and

$$\frac{\partial k*}{\partial p} = 0. \tag{4.26}$$

For inflators, $k* = \frac{1}{\alpha s} \ln \frac{s\alpha^2}{v-u+\Delta A(1-p-pr)}$, thus

$$\frac{\partial k*}{\partial u} = \frac{1}{\alpha s(v-u+\Delta A(1-p-pr))} > 0 \tag{4.27}$$

and

$$\frac{\partial k*}{\partial p} = \frac{\Delta A(1+r)}{\alpha s(v-u+\Delta A(1-p-pr))} > 0, \tag{4.28}$$

since v-u>0, and 1-p-pr>0.

Proposition 3*. CMS's control of IoT adoption by adjusting u and p*

a) CMS can affect honest nursing homes' optimal IoT adoption level by subsidization. The optimal IoT level is monotonically increasing in u, with gradient $\frac{1}{\alpha s(v-u)}$. *CMS cannot affect honest nursing homes' IoT adoption by adjusting audit probability p.*

[6] According to the inflating condition, the relationship between p and r determines the audit policy. In practice, adjusting audit probability p is usually easier than adjusting punishment rate r. Similar case can be found in tax fraud detection, in which IRS has adjusted auditing probability several times over the years, but the punishment rate is relatively stable. As a result, we consider audit probability p as a decision variable that CMS can adjust accordingly, but assume r is preannounced and do not change.

b) CMS can affect inflators' optimal IoT adoption level by both subsidization and audit. The optimal IoT level is monotonically increasing in both u and p. The gradients are $\frac{1}{\alpha s(v-u+\Delta A(1-p-pr))}$ and $\frac{\Delta A(1+r)}{\alpha s(v-u+\Delta A(1-p-pr))}$, respectively.

c) For IoT adoption, honest nursing homes are more sensitive to subsidization than inflators.

According to Proposition 3, the rate for controlling inflators' IoT adoption level by adjusting p is $\Delta\mathrm{A}(1 + \mathrm{r})$ times as much as controlling IoT adoption by u. In practice, if ΔA is big, i.e., the maximum profit gain from inflation is big, then controlling IoT adoption by p is more effective. If ΔA is very small, and $\Delta\mathrm{A}(1 + \mathrm{r}) < 1$, then controlling IoT adoption by subsidization is more effective.

The results also indicate that inflators are less sensitive to subsidization than their honest counterparts. The difference is big when the term $\Delta prof(1 - p - pr)$ is significantly larger than 0. If ΔA is very small, or if $p \approx \frac{1}{1+r}$, such that $1 - p - pr$ is close to 0, then difference is not significant.

4.3. Allocating Resources between Auditing and Subsidization

In this section, we analyze the problem from CMS's point of view, and summarize insights on allocating limited budget between auditing and subsidization.

Suppose there are totally N nursing homes. The population of honest nursing homes and inflators are denoted as N_H and N_I, respectively. The population of honest nursing homes whose rating increased after self-reporting and inflators is denoted as N_{HI}. The population of inflators (whose ratings also increase after self-reporting) is denoted as N_{II}. The population of honest nursing home whose ratings do not increase after self-reporting is denoted as N_{HN}. We assume that the inflate condition is satisfied, i.e., $\mathrm{p} < \frac{1}{1+\mathrm{r}}$, thus all inflators will inflate to the maximum amount they can, i.e., except for the measures self-reported by the IoT technologies they adopt, they will inflate all other measures. Thus we have N_{II}=N_I, and N_{IN}=0, i.e., there is no inflator who is deterred by the audit policy and decides to stay honest.

Suppose x_i is a binary variable which denotes the audit status of the ith nursing home, x_i=1, if the nursing home is audited, and x_i=0, if the nursing home is not audited.

4.3.1 The Damage of An Inflator

The damage of an inflator is normalized to 1, and is defined as $d_i=(1-k_i)(1-x_i)$. If an inflator is audited, and x_i=1, then d_i=0, and the inflator makes no damage to the rating system. If an inflator is not audited, and accept IoT technologies to level k_i, $k_i \in [0, 1]$, then the damage to the system is $1-k_i$. It can be seen that if the inflator's IoT adoption rate k_i is high, i.e., close to 1, then even it is not audited, its damage to the system is still very small, or close to 0. On the other hand, if an inflator's IoT adoption level is low, i.e., close to 0, then if it is not audited, it will make a severe damage to the system, i.e., close to 1.

4.3.2 Two-level Structure

The problem has a 2-level structure, with CMS minimizing inflator damage to the system by determining audit probability p and subsidization amount Sub at the higher level, and nursing homes, both honest ones and inflators, maximining profits by selecting optimal IoT adoption level k and staffing level s at the lower level, as shown in Figure 4-1. Note that the optimal staffing level or each nursing home is not directly related to the CMS's problem, but the IoT adoption level ki for each nursing home i directly affect CMS's subsidization budget.

The CMS's problem is to minimize the damage of all inflators to the system, which can be written as follows.

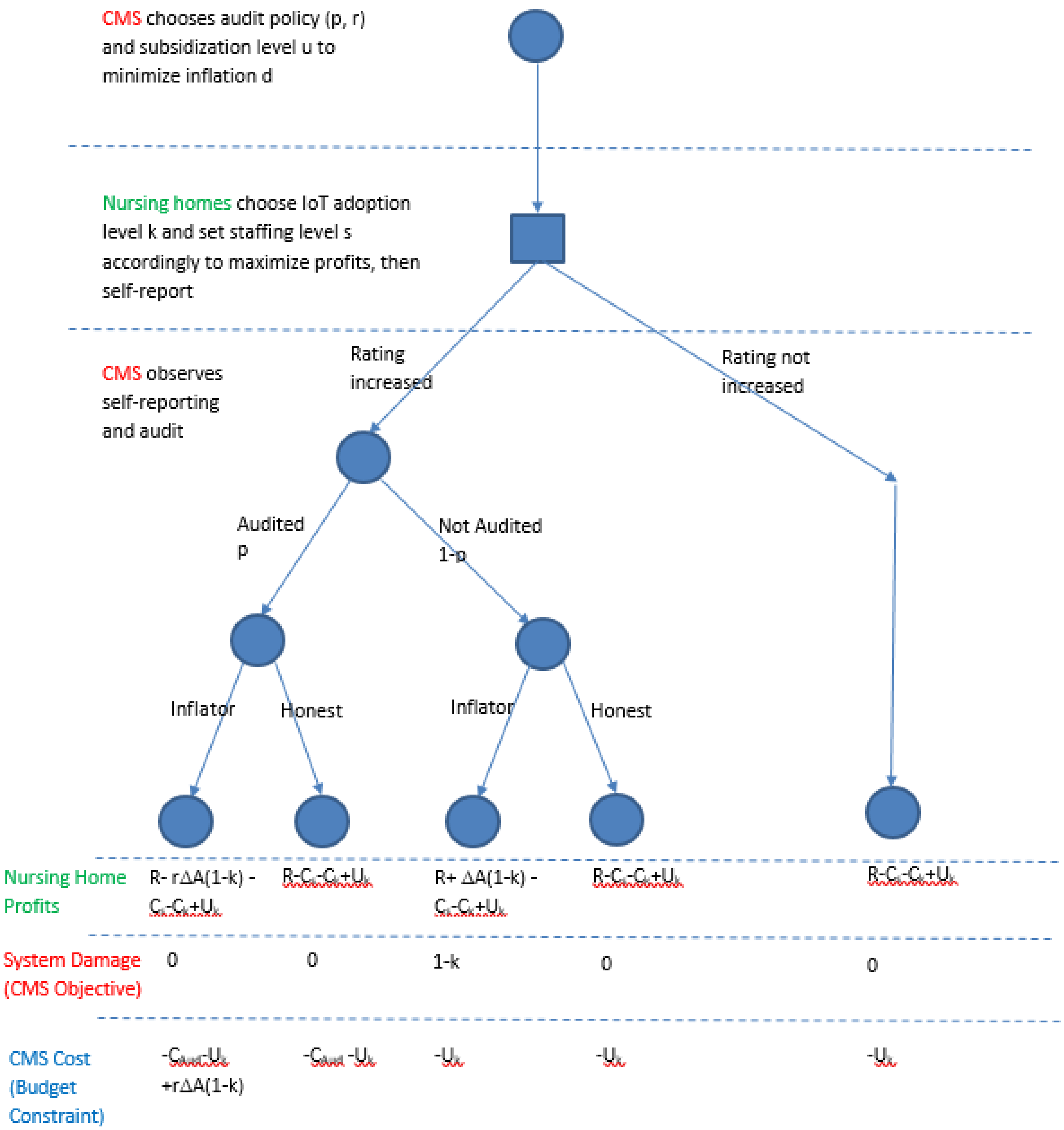
$$\min_{p,u} \sum_{i=1}^{N_{II}} (1-k_i)(1-x_i) \tag{4.29}$$

s.t. $$\sum_{i=1}^{N} k_i u + (N_{II} + N_{HI}) p C_{Aud} \le B_0 + N_{II} pr \Delta A \tag{4.30}$$

$$\sum_{i=1}^{N_{II}} x_i = pN_{II} \tag{4.31}$$

and the KKT conditions (4.20) describing optimal k_i for N_{II} inflators, and the KKT conditions (4.11) for N_H honest nursing homes.

Figure 4-1 A 2-level Problem Structure



4.3.3 Decomposition and Individual Subproblem

To analyze the CMS's overall problem directly is very complicated. Notice that CMS's objective is to minimize the damage of the inflators within the budget. If an inflating nursing home is audited and caught, its damage to the system reduces to 0. If an inflating nursing home is not audited, its damage to the system is 1-k, and IoT adoption level k is the only factor restricting its damage. In other words, CMS wants to limit the damage of the inflators who are not audited to a low level by subsidizing IoT.

For each inflator, it has a probability p to be audited, and probability 1-p to be not audited. The expected damage can then be calculated as

$$\mathrm{E}[\mathrm{d_i}] = \mathrm{p} * 0 + (1-\mathrm{p})(1-\mathrm{k_i}), \tag{4.32}$$

where k_i is its optimal IoT adoption level which is determined in the nursing home's profit maximization problem. The CMS's problem can then be decomposed and written as to minimize the expected damage for each inflator by selecting the right audit policy (p, r), and subsidization amount u. The individual subproblem can be written as

$$\min_{\mathrm{p,u}} \mathrm{E}[\mathrm{d_i}] = \mathrm{p} * 0 + (1-\mathrm{p})(1-\mathrm{k_i}) = (1-\mathrm{p})(1-\mathrm{k_i}) \tag{4.33}$$

s.t. $$\textstyle\sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{k_i}\,\mathrm{u} + (\mathrm{N_{II}} + \mathrm{N_{HI}})\mathrm{p}\mathrm{C_{Aud}} \leq \mathrm{B_0} + \sum_{\mathrm{i=1}}^{\mathrm{pN_{II}}} \Delta\mathrm{A}\,(1-\mathrm{k_i})\mathrm{r}, \tag{4.34}$$

and KKT condition (4.20).

The budget constraint (4.34) can be written as

$$Hu + Jp \leq \mathrm{B_0}, \tag{4.35}$$

where H and J are functions of k.

$$H = \textstyle\sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{k_i} = N\bar{k}, \text{ and} \tag{4.36}$$

$$J = (N_{II} + N_{HI})C_{Aud} - r\Delta\mathrm{A}(1-\overline{k^I})N_{II} \tag{4.37}$$

The problem is highly nonlinear and very difficult to be analyzed quantitatively. We instead summarize qualitative insights which can help CMS allocate its budget. To analyze the problem, we plot it in the plane spanned by u and p, as shown in Figure 4-2.

We are interested in the square area where $p \in [0,1]$, and $u \in [0,v]$. According to Proposition 3, k increases in u monotonically, thus 1-k decreases in u monotonically. Similarly, we can also know that 1-k decreases in p monotonically, and (1-p)(1-k) decreases in p faster. For a given damage level d, the level curve function can be calculated as

$$d = (1-p)(1-k) = (1-p)(1-\frac{1}{\alpha s} ln \frac{s\alpha^2}{v-u+\Delta A(1-p-pr)}), \tag{4.38}$$

or equivalently,

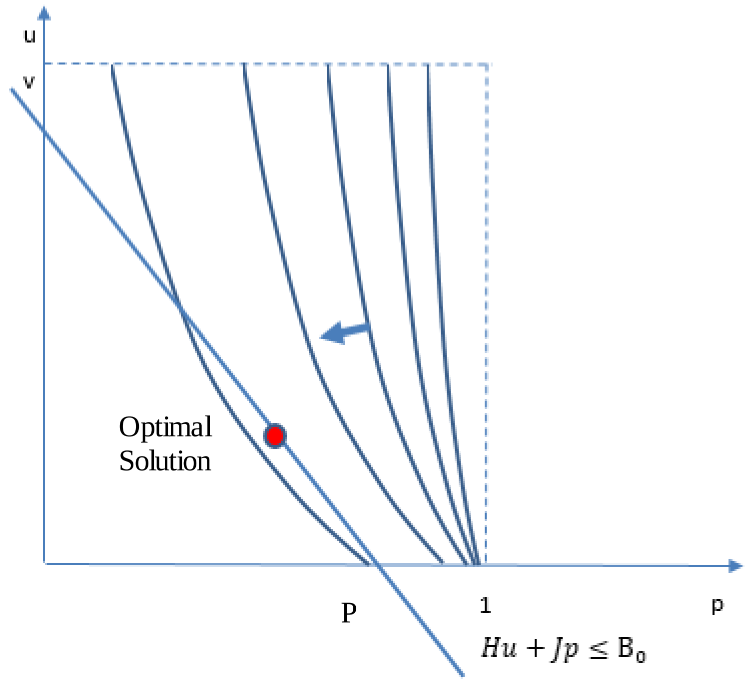
$$u = v + \Delta A(1-p-pr) - s\alpha^2 e^{(\frac{d}{1-p}-1)\alpha s}. \tag{4.39}$$

Thus we have

$$\frac{du}{dp} = -\Delta A - \Delta Ar + s^2\alpha^3 d \frac{1}{(1-p)^2} e^{(\frac{d}{1-p}-1)\alpha s}. \tag{4.40}$$

It can be seen that when p increase, the third term increases, but the first two terms are negative and unchanged, thus the level curve is convex on the plane spanned by u and p, and can be plotted as shown in Figure 4-2. The gradient is also plotted which is orthogonal to the level curve, as shown in Figure 4-2 as an arrow. In this problem, the budget constraint is binding, and the optimal solution is achieved on the budget constraint, as shown in Figure 4-2 as a red dot.

Figure 4-2 The CMS's Damage Minimization Problem (plotted on p and u)



When $H = N\bar{k}$ is big, the optimal solution moves towards the intersection of the budget constraint and horizontal axis p, i.e., CMS should allocate more budget in audit rather than in subsidization. The following cases are included in this scenario:

Proposition 4.1. *Cases that CMS should allocate more budget in audit than in subsidization.*

i) The total population N is big.

ii) The average level of IoT adoption is already high ($\bar{k}$ is large).

When $J = (N_{II} + N_{HI})C_{Aud} - r\Delta A(1 - \bar{k_I})N_{II}$ is big, the optimal solution moves towards the intersection of the budget constraint and vertical axis u, i.e., CMS should allocate more budget in subsidization than in audit. The following cases are included in this scenario:

Proposition 4.2. *Cases that CMS should allocate more budget in subsidization than in audit*

i) The population of confounding honest nursing homes is high (N_{HI} is large)

ii) Unit audit cost C_{Aud} is high

iii) Punishment rate r is low

iv) The additional profit ΔA that nursing homes can gain is small

v) The average IoT adoption level $\bar{k_I}$ for inflators is small. (This may be difficult to be observed directly, but can be partially reflected in the overall average IoT adoption level $\bar{k}$)

4.3.4 Under What Condition Shall CMS Subsidize

It is important to analyze the conditions under which CMS shall consider using subsidization. From Figure 4-2, we can easily find that when the optimal solution is not the corner point P, CMS should consider subsidization. Since we have shown the convexity of the level curves, the condition can be restated as follows: When the slope of the level curve at point P is higher (slope is negative) than the slope of the budget constraint, there will be a tangent point in the first quadrant, which is the optimal solution to the problem.

Proposition 5. Conditions Under Which Subsidization is Necessary

CMS should consider subsidizing IoT when inequality $\left.\frac{du}{dp}\right|_{u=0} > -\frac{J}{H}$ is satisfied.

Solving the inequality directly, however, is complicated, but we can obtain insights by analyzing some extreme cases. Consider the case in which the optimal IoT adoption level for all inflators is k =0. This can happen when r is small, i.e., when nursing homes won't be punished a lot for inflating. In this case, d = 1-p, and the inequality (4.41) reduces to

$$-\Delta A - \Delta A r + \frac{s^2\alpha^3}{1-p} > -\frac{(N_{II}+N_{HI})C_{Aud} - r\Delta A N_{II}}{N\bar{k}} \tag{4.41}$$

$$\Delta A < \frac{s^2\alpha^3 N\bar{k} + (N_{II}+N_{HI})(1-p)C_{Aud}}{((1+r)N\bar{k} - rN_{II})(1-p)} \tag{4.42}$$

Equation (4.42) can be interpreted as follows: If the gains from inflation is large, the inflators will choose to inflate and refuse to adopt IoT. In this case, CMS should not consider subsidization, and the optimal solution is the corner point P, i.e., to allocate all the budget on auditing. CMS should only consider subsidization when the inflating gain ΔA is under certain threshold.

4.4. Conclusion

In this chapter, we create an analytical model to systematically investigate the impact of IoT adoption on the nursing home rating system and their operations decisions. We derive conditions for inflating and honest nursing homes to show their optimal staffing level decisions and IoT adoption level decisions. We find that highly rated nursing homes will have a decreased staffing level when IoT is adopted, but poorly rated nursing homes will have their staffing level increased after IoT adoption. Since IoT devices can automatically collect data which is originally self-reported by nursing homes, IoT adoption reduces the possibility of misreporting. However, IoT technologies can be costly and CMS may consider subsidizing IoT adoption and push it through, and use it as an alternative way to control rating inflation other than the traditional audit method. We derive close conditions to show how CMS can affect nursing homes' IoT

adoption by auditing and subsidization, respectively. We also obtain insights on how CMS should allocate limited resources between auditing and subsidization.

Chapter 5. Conclusion and Future Research

This dissertation consists of three essays which addresses a series of issues in CMS's 5-star Nursing Home Compare system. In the first essay, we conducted an empirical study to demonstrate the existence of rating inflation. We find a significant association between the changes in a nursing home's star rating and its profits, which points to a financial incentive for nursing homes to inflate the ratings. By using the number of patients' complaints as a proxy of the true service quality, we are able to demonstrate the existence of rating inflation. A prediction model is then developed, which provides a quantitative evaluation on the system's performance. In the second essay, we look into methods to improve the rating system. By using an innovative graph-based method, we are able to formulate the problem in a linear form, and solve it based on CMS data. The results show that when minimizing the difficulty to catch inflators, the current inspection domain is the optimal choice, and an effective audit system is necessary. An audit system is then designed for CMS, and insights are obtained on setting optimal audit parameters. In the third essay, we analyze how technologies, particularly IoT, can affect the rating system and nursing homes' operational decisions. A game theoretical model is developed with CMS minimizing inflation damage at the higher level and nursing homes maximizing profits at the lower level. We propose that IoT adoption subsidization can be an alternative method for CMS to control inflation, and we analyze how CMS can affect nursing homes' IoT adoption level by auditing and subsidization. We also obtain insights on CMS's budget allocation between auditing and subsidization.

The results in this dissertation pinpoint the key issues in CMS's nursing home rating system, and provide a guideline for CMS's rating system reform. Some of the results can also be extended to other rating system sharing similar features. For example, the physician rating system developed by CMS also uses a combination of inspection and self-reported measures to generate ratings, thus the inflation detection method we developed in chapter 2 can be applied to this system easily.

References

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

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

Abbasi, A., Zahedi, F. M., Zeng, D., Chen, Y., Chen, H., and Nunamaker Jr., J. F. "Enhancing Predictive Analytics for Anti-Phishing by Exploiting Website Genre Information**."** *Journal of Management Information Systems***,** 31, 4 (2015), 109–157**.,**

Anderson, David, Guodong (Gordon) Gao, and Bruce Golden. "Life Is All about Timing: An Examination of Differences in Treatment Quality for Trauma Patients Based on Hospital Arrival Time." *Production and Operations Management* 23,12 (2014): 2178–90. doi:10.1111/poms.12236.

Anderson, Eugene, Fornell Claes, and Rust Roland. "Customer Satisfaction, Productivity, and Profitability: Differences between Goods and Services." *Marketing Science* 16,2 (2014): 129–45.

Andritsos, Dimitrios A., and Sam Aflaki. "Competition and The Operational Performance of Hospitals: The Role of Hospital Objectives." *Production and Operations Management* 24,11 (2015): 1812–32. doi:10.1111/poms.12416.

Bai, Xue, Ram Gopal, Manuel Nunez, and Dmitry Zhdanov. "On the prevention of fraud and privacy exposure in process information flow."*INFORMS Journal on Computing* 24, no. 3 (2012): 416-432.

Bauer, Thomas, and Sinning Mathias. "An Extension of the Blinder-Oaxaca Decomposition to Nonlinear Models." *AStA Advances in Statistical Analysis* 92,2 (2008): 197–206.

Berg, Katherine, Vincent Mor, John Morris, Katharine Murphy, Terry Moore, and Yael Harris. "Identification and Evaluation of Existing Nursing Homes Quality Indicators." *Healthcare Financing Review* 23.4. (2002)

Biemer, Paul P., and S. Lynne Stokes. "The optimal design of quality control samples to detect interviewer cheating." *Journal of Official Statistics* 5.1 (1989): 23-39.

Blendon, R.J., Brodie, M., Benson, J.M., Altman, D.E. and Buhr, T., "Americans' views of health care costs, access, and quality." *Milbank Quarterly, 84*,4 (2006), pp.623-657.

Carman, J.M., "Consumer Perceptions Of Service Quality: An Assessment Of T." *Journal of retailing, 66*,1 (1990), p.33.

Castle, N.G., "Nursing home caregiver staffing levels and quality of care a literature review." *Journal of Applied Gerontology, 27*,4 (2008), pp.375-405.

Cayirli, Tugba, and Emre Veral. "Outpatient Scheduling in Health Care: A Review of Literature." *Production and Operations Management* 12, 4 (2003): 519–49. doi:10.1111/j.1937-5956.2003.tb00218.x.

Cebula, Richard J. "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).

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

Centers for Medicare and Medicaid Services (CMS). "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*, (2015).

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

Centers for Medicare and Medicaid Services. "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,* (2015).

Centers for Medicare and Medicaid Services. "Staffing Data Submission PBJ." *https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Staffing-Data-Submission-PBJ.html,* (2016a).

Centers for Medicare and Medicaid Services. "Further Improvements to the Nursing Home Compare Five-Star Quality Rating System." *https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/Downloads/Improvements-NHC-April-2016.pdf,* (2016b).

Centers for Medicare and Medicaid Services. "Nursing Home Compare 3.0: Revisions to the Nursing Home Compare 5-Star Quality Rating System." (2015)

Charlene, Harrington, Steffie Woolhandler, Joseph Mullan, Helen Carrillo, and David Himmelstein. "Does Investor Ownership of Nursing Homes Compromise the Quality of Care?" *American Journal of Public Health* 91,9 (2001): 1452–55.

Chen, Hong, Qu Qian, and Anming Zhang. "Would Allowing Privately Funded Health Care Reduce Public Waiting Time? Theory and Empirical Evidence from Canadian Joint Replacement Surgery Data." *Production and Operations Management* 24,4 (2015): 605–18. doi:10.1111/poms.12260.

Chesteen, Susan, Berit Helgheim, Taylor Randall, and Don Wardell. "Comparing Quality of Care in Non-Profit and for-Profit Nursing Homes: A Process Perspective." *Journal of Operations Management*,

Operations Management in Not-For-Profit, Public and Government Services, 23,2 (2005): 229–42. doi:10.1016/j.jom.2004.08.004.

Chow, Vincent S., Martin L. Puterman, Neda Salehirad, Wenhai Huang, and Derek Atkins. "Reducing Surgical Ward Congestion Through Improved Surgical Scheduling and Uncapacitated Simulation." *Production and Operations Management* 20,3 (2011): 418–30. doi:10.1111/j.1937-5956.2011.01226.x.

Committee on Nursing Home Regulation. *Improving the Quality of Care in Nursing Homes*. Institute of Medicine. (1986)

Dabholkar, P.A., Thorpe, D.I. and Rentz, J.O.. "A measure of service quality for retail stores: scale development and validation." *Journal of the Academy of marketing Science, 24*,1 (1995), pp.3-16.

David, Grabowski, Joseph Angelelli, and Vincent Mor. 2004. "Medicaid Payment and Risk-Adjusted Nursing Home Quality Measures." *Health Affairs* 23,5 (2004): 243–52.

David, Grabowski, Zhanlian Feng, Orna Intrator, and Vincent Mor. "Recent Trends in State Nursing Home Payment Policies." *Health Affairs* w4. (2004)

Dellarocas, C. "Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior." *Proceedings of the 2nd ACM Conference on Electronic commerce,* (2000), 150 -157.

DellaVigna, Stefano, and Eliana La Ferrara. "Detecting Illegal Arms Trade." *American Economic Journal: Economic Policy* 2,4 (2010): 26–57.

Dobson, Gregory, Sameer Hasija, and Edieal J. Pinker. "Reserving Capacity for Urgent Patients in Primary Care." *Production and Operations Management* 20,3 (2011): 456–73. doi:10.1111/j.1937-5956.2011.01227.x.

Duggan, Mark, and Levitt Steven. "Winning Isn't Everything: Corruption in Sumo Wrestling." *National Bureau of Economic Research* w7798. (2000)

Duhigg, Charles. “Washington Scrutinizes Nursing Homes.” *The New York Times*, November 16. http://www.nytimes.com/2007/11/16/business/16care.html. (2007)

Eiken, Steve, Kate Sredl, Lisa Gold, Jessica Kasten, Brian Burwell, and Paul Saucier. “Medicaid Expenditures for Long-Term Services and Supports in FFY 2012.” *Bethesda, MD: Truven Health Analytics.* (2014) http://www1a.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Long-Term-Services-and-Supports/Downloads/LTS-Expenditures-2012.pdf.

Engelberg, Joseph, Parsons Christopher, and Tefft Nathan. “Financial Conflicts of Interest in Medicine.” (2014) *SSRN* 2297094.

Fathil, Firhana M., and James M. Schmidtke. "The Relation between Individual Differences and Accountants' Fraud Detection Ability." *International Journal of Auditing* 14.2 (2010): 163-173.

Fairlie, Robert. “An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models.” *Journal of Economic and Social Measurement* 30,4 (2005): 305–16.

Fennell, Mary, Zhanlian Feng, Melissa Clark, and Vincent Mor. “Elderly Hispanics More Likely to Reside in Poor-Quality Nursing Homes.” *Health Affairs* 29,1 (2010): 65–73.

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

Gardner, Everette. “Dimensional Analysis of Airline Quality.” *Interfaces* 34,4 (2004): 272–79.

General Accounting Office. “Nursing homes: Public reporting of quality indicators has merit, but national implementation is premature.” (2002). Washington, DC: GAO.

Gokula, R.R.M., Hickner, J.A. and Smith, M.A. “Inappropriate use of urinary catheters in elderly patients at a midwestern community teaching hospital.” *American journal of infection control*, *32*,4 (2004), pp.196-199.

Grabowski, D.C. and O'Malley, A.J. "Use of telemedicine can reduce hospitalizations of nursing home residents and generate savings for Medicare." *Health Affairs, 33,2* (2014), pp.244-250.

Güneş, Evrim Didem, Egemen Lerzan Örmeci, and Derya Kunduzcu. "Preventing and Diagnosing Colorectal Cancer with a Limited Colonoscopy Resource." *Production and Operations Management* 24,1 (2015): 1–20. doi:10.1111/poms.12206.

Han, Xu, Yaraghi Niam, Ram Gopal, "An Empirical Investigation on the Prevalence of Manipulation in Nursing Homes' Rating System." *Informs CIST 2015,* Philadelphia (2015).

Harrington, C., Olney, B., Carrillo, H. and Kang, T. "Nurse Staffing and Deficiencies in the Largest For-Profit Nursing Home Chains and Chains Owned by Private Equity Companies." *Health services research, 47*(1pt1) (2012), pp.106-128.

Harrington, C., Zimmerman, D., Karon, S.L., Robinson, J. and Beutel, P. "Nursing home staffing and its relationship to deficiencies." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 55,5* (2000), pp.S278-S287.

Helm, Jonathan E., Shervin AhmadBeygi, and Mark P. Van Oyen. "Design and Analysis of Hospital Admission Control for Operational Effectiveness." *Production and Operations Management* 20,3 (2011): 359–74. doi:10.1111/j.1937-5956.2011.01231.x.

Hyattsville, Md: U.S. Dept. of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

IBM, "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).

Jack, Eric P., and Thomas L. Powers. "Volume Flexible Strategies in Health Services: A Research Framework." *Production and Operations Management* 13,3 (2004): 230–44. doi:10.1111/j.1937-5956. 2004. tb00508. x.

Jacob, Brian, and Levitt Steven. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *National Bureau of Economic Research* w9413 (2003).

Jindal, N., and Liu, B. "Opinion Spam and Analysis." *Proceedings of First ACM International Conference on Web Search and Data Mining* (2008). Stanford, California, USA.

Johnson, Robert. "Linking Complaint Management to Profit." *International Journal of Service Industry Management* 12,1 (2001): 60–69.

Josang, A., and Ismail, R. "The beta reputation system." *Proceedings of the 15th Bled Electronic Commerce conference*. (2002)

Justice, Scott C., Eric L. Hopper, and Ken C. Obrist. "Fraud prevention system and method." U.S. Patent No. 6,516,056. 4 Feb. (2003).

Kaiser Family Foundation. "Reading the Stars: Nursing Home Quality Star Ratings, Nationally and by State," White Paper. (2015)

Kaiser Family Foundation (KFF), "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

Kane, R.L. and Kane, R.A., "What older people want from long-term care, and how they can get it." *Health Affairs, 20,6* (2001), pp.114-127.

Kim, H., Kovner, C., Harrington, C., Greene, W. and Mezey, M.. "A panel data analysis of the relationships of nursing home staffing levels and standards to regulatory deficiencies." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, p.gbn019.(2009)

Klimek, Peter, Yuri Yegorov, Rudolf Hanel, and Stefan Thurner. "Statistical detection of systematic election irregularities." *Proceedings of the National Academy of Sciences* 109, no. 41 (2012): 16469-16473.

Konetzka, R.T., Yi, D., Norton, E.C. and Kilpatrick, K.E.. "Effects of Medicare payment changes on nursing home staffing and deficiencies." *Health services research*, *39* (2004), pp.463-488.

Konetzka, R.T., Grabowski, D.C., Perraillon, M.C. and Werner, R.M. "Nursing home 5-star rating system exacerbates disparities in quality, by payer source." *Health Affairs*, *34,5* (2015), pp.819-827.

Konetzka, R. Tamara, Sally C. Stearns, and Jeongyoung Park. "The staffing–outcomes relationship in nursing homes." *Health services research*43, no. 3: 1025-1042, (2008)

Krawiec, RJ, Nadler, Jessica, Tye, Elan & Jarboe, Jennifer, "No appointment necessary: How the IoT and patient-generated data can unlock health care value." *Deloitte University Press: http://dupress.com/articles/internet-of-things-iot-in-health-care-industry/,* (2015).

LaGanga, Linda R., and Stephen R. Lawrence. "Appointment Overbooking in Health Care Clinics to Improve Patient Service and Clinic Performance." *Production and Operations Management* 21,5 (2012): 874–88. doi:10.1111/j.1937-5956.2011.01308. x.

Li, Y., Cai, X., Ye, Z., Glance, L.G., Harrington, C. and Mukamel, D.B. "Satisfaction with Massachusetts nursing home care was generally high during 2005–09, with some variability across facilities." *Health Affairs*, *32,8* (2013), pp.1416-1425.

Long, J. S. "Regression Models for Categorical and Limited Dependent Variables." Thousand Oaks, CA: Sage Publications. (1997)

Li, Jing, Kuei-Ying Huang, Jionghua Jin, and Jianjun Shi. "A survey on statistical methods for health care fraud detection." *Health care management science* 11, no. 3 (2008): 275-287.

Lim, E., Nguyen, V., Jindal, N., Liu, B., and Lauw, H. "Detecting Product Review Spammers using Rating Behaviors." *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, (2010). Toronto, Canada

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

Liu, S., Zhang, J., Miao, C., Theng, Y., and Kot, A. "An Integrated Clustering-Based Approach to Filtering Unfair Multi-nominal Testimonies." *Computational Intelligence*, 30, (2014)

Lu, Susan, Rui, Huaxia, Seidmann, Avi. "Is Technology Eating Nurses? Staffing Decisions in NursingHomes." *Informs CIST 2015*, Philadelphia, (2015).

Lu, Susan F., and Lauren Xiaoyuan Lu. "Do Mandatory Overtime Laws Improve Quality? Staffing Decisions and Operational Flexibility of Nursing Homes." *Management Science, Forthcoming*, (2016).

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

Mayzlin, Dina, Dover Yaniv, and Chevalier Judith. "Promotional Reviews: An Empirical Investigation of Online Review Manipulation." *National Bureau of Economic Research* w18340.(2012)

McCoy, Jessica H., and M. Eric Johnson. "Clinic Capacity Management: Planning Treatment Programs That Incorporate Adherence." *Production and Operations Management* 23,1 (2014): 1–18. doi:10.1111/poms.12036.

Medicare, Centers for, Medicaid Services 7500 Security Boulevard Baltimore, and Md21244 Usa. 2016. "Staffing-Data-Submission-PBJ." https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Staffing-Data-Submission-PBJ.html.

Miller, E.A., Tyler, D.A., Rozanova, J. and Mor, V.. "National newspaper portrayal of US nursing homes: periodic treatment of topic and tone." *The Milbank Quarterly, 90,4* (2012), pp.725-761.

Mor, Vincent, Katherine Berg, Joseph Angelelli, David Gifford, John Morris, and Terry Moore. "The Quality of Quality Measurement in US Nursing Homes." *The Gerontologist* 43 (suppl 2): 37–46. (2003)

Mor, V., Zinn, J., Angelelli, J., Teno, J.M. and Miller, S.C.. "Driven to tiers: socioeconomic and racial disparities in the quality of nursing home care." *Milbank Quarterly, 82,2* (2004), pp.227-256.

Mor, V., "Improving the quality of long-term care with better information." *The Milbank Quarterly, 83,3* (2005), pp.333-364.

Mor, V., Intrator, O., Feng, Z. and Grabowski, D.C., "The revolving door of rehospitalization from skilled nursing facilities." *Health affairs, 29,1* (2010), pp.57-64.

Mukherjee, A., Kumar, A., Liu, B., Wang, J., Hsu, M., Castellanos, M., and Ghosh, R. "Spotting Opinion Spammers using Behavioral Footprints." *Proceedings of SIGKDD International Conference on Knowledge Discovery and Data Mining*, Chicago, USA. (2013)

Munroe, D.J., "The influence of registered nurse staffing on the quality of nursing home care." *Research in nursing & health, 13,4* (1990), pp.263-270.

National Center for Health Statistics (U.S.), ed. "The National Nursing Home Survey: 2004 Overview. Vital and Health Statistics." Ser. 13, *Data from the National Health Care Survey*, (2009) no. 167.

Oaxaca, Ronald. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* (1973), 693–709.

OBRA. "Federal Nursing Home Reform Act from the Omnibus Budget Reconciliation Act of 1987." (1987) US Congress.

Perols, Johan. "Financial statement fraud detection: An analysis of statistical and machine learning algorithms." *Auditing: A Journal of Practice & Theory*30.2 (2011): 19-50.

Roland, Rust, and T. Siong Chung. “Marketing Models of Service and Relationships.” *Marketing Science* 25,6 (2006): 560–80.

Rosalie, Kane. “Definition, Measurement, and Correlates of Quality of Life in Nursing Homes: Toward A Reasonable Practice, Research, and Policy Agenda.” *The Gerontologist* 43 (suppl 2) (2003): 28–36.

Roth, Aleda V., and Roland Van Dierdonck. “Hospital Resource Planning: Concepts, Feasibility, and Framework.” *Production and Operations Management* 4,1 (1995): 2–29. doi:10.1111/j.1937-5956.1995.tb00038.x.

Salzarulo, Peter A., Stephen Mahar, and Sachin Modi. “Beyond Patient Classification: Using Individual Patient Characteristics in Appointment Scheduling.” *Production and Operations Management* (2015), December, n/a – n/a. doi:10.1111/poms.12528.

Shugarman, L.R. and Brown, J.A. “Nursing home selection: How do consumers choose? Volume I: Findings from focus groups of consumers and information intermediaries.” *Nursing*, (2006) *10*, p.01.

Shugarman, L.R. and Garland, R. “Nursing home selection: How do consumers choose? Volume II: Findings from the Website content review.” *Nursing*, (2006)

Shwartz, M., Restuccia, J.D. and Rosen, A.K.. “Composite measures of health care provider performance: a description of approaches.” *Milbank Quarterly*, *93*,*4* (2015), pp.788-825.

Smith, David Barton, Zhanlian Feng, Mary Fennell, Jacqueline Zinn, and Vincent Mor. “Separate and Unequal: Racial Segregation and Disparities in Quality Across US Nursing Homes.” *Health Affairs* 26,5 (2007): 1448–58.

Smith, D.B., Feng, Z., Fennell, M.L., Zinn, J.S. and Mor, V.. “Separate and unequal: racial segregation and disparities in quality across US nursing homes.” *Health Affairs*, *26,5* (2007), pp.1448-1458.

Stavropoulou, C., Doherty, C. and Tosey, P. “How Effective Are Incident-Reporting Systems for Improving Patient Safety? A Systematic Literature Review.” *Milbank Quarterly*, *93,4* (2015), pp.826-866.

Stevenson, David, and Studdert David. “The Rise of Nursing Home Litigation: Findings from A National Survey of Attorneys.” *Health Affairs* 22,2 (2003): 219–29.

Stevenson, David. “Is a public reporting approach appropriate for nursing home care?” *Journal of Health Politics, Policy and Law*, *31*(4), 773–810.Stevenson, David, and Grabowski David. 2008. “Private Equity Investment and Nursing Home Care: Is It A Big Deal?” *Health Affairs* 27,5 (2006): 1399–1408.

Stevenson, David, and Grabowski David. “Private Equity Investment and Nursing Home Care: Is It A Big Deal?” *Health Affairs* 27 (5) (2008): 1399–1408.

Stenner, Shane P. "Evaluating the patient-centered automated SMS tagging engine (PASTE): natural language processing applied to patient-generated SMS text messages." *PhD diss., Vanderbilt University*, (2011).

Thomas, Katie. “Medicare Star Ratings Allow Nursing Homes to Game the System.” *The New York Times*, (2014) August 24. http://www.nytimes.com/2014/08/25/business/medicare-star-ratings-allownursing-homes-to-game-the-system.html.

Tsaur, S.H., Chang, T.Y. and Yen, C.H. “The evaluation of airline service quality by fuzzy MCDM.” *Tourism management*, *23,2* (2002), pp.107-115.

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

Walshe, K. "Regulating US nursing homes: are we learning from experience?" *Health Affairs, 20,6* (2001), pp.128-144.

Wang, G., Xie, S., Liu, B., and Yu, P. S. "Identify Online Store Review Spammers via Social Review Graph." *ACM Transactions on Intelligent Systems and Technology* (2011).

Werner, Rachel, and Tamara Konetzka. "Advancing Nursing Home Quality Through Quality Improvement Itself." *Health Affairs* 29 (January) (2010): 81–86.

Werner, Rachel M., R. Tamara Konetzka, and Daniel Polsky. "Changes in Consumer Demand Following Public Reporting of Summary Quality Ratings: An Evaluation in Nursing Homes." *Health Services Research*, (2016) February. doi:10.1111/1475-6773.12459.

Werner, Rachel, Elizabeth Stuart, and Daniel Polsky. "Public Reporting Drove Quality Gains at Nursing Homes." *Health Affairs* 29,9 (2010): 1706–13.

William, Cabin, David Himmelstein, Michael Siman, and Steffie Woolhandler. "For-Profit Medicare Home Health Agencies' Costs Appear Higher and Quality Appears Lower Compared to Nonprofit Agencies." *Health Affairs* 33,8 (2014): 1460–65.

Wright, R. T., and Marett, K. "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.

Zepeda, E. David, and Kingshuk K. Sinha. "Toward an Effective Design of Behavioral Health Care Delivery: An Empirical Analysis of Care for Depression." *Production and Operations Management*, (2016) January, n/a – n/a. doi:10.1111/poms.12529.

Zhang, X. and Grabowski, D.C.. "Nursing home staffing and quality under the nursing home reform act." *The Gerontologist, 44,1* (2004), pp.13-23.

Zhang, Yue, Martin L. Puterman, Matthew Nelson, and Derek Atkins. “A Simulation Optimization Approach to Long-Term Care Capacity Planning.” *Oper. Res.* 60,2 (2012): 249–61. doi:10.1287/opre.1110.1026.

Zhang, Ning, Susan F. Lu, Biao Xu, Bingxiao Wu, Rosa Rodriguez-Monguio, and Jerry Gurwitz. "Health Information Technologies: Which Nursing Homes Adopted Them?." *Journal of the American Medical Directors Association* 17, no. 5: 441-447, (2016).