Do Patients Benefit from Regulatory Stringency?
Evidence from Targeted Nursing Homes

May 9, 2023

Abstract

Certification requirements are a key policy lever to incentivize quality in the nursing home industry, but the effects of regulation has been hard to identify, in part because homes are all subject to the same regulatory requirements and sanctions are imposed in response to care failures. This paper studies the Special Focus Facility program (SFF), which aims to improve quality of care by targeting oversight toward some, but not all, of the industry's worst performing facilities. I leverage capacity constraints in a difference-in-difference framework and show that after two years treated facilities improve an additional 17% compared to untreated candidates, or about two fewer deficient practices. I find performance reverts back *after treatment ends*, which makes alternative explanations unlikely.

Keywords: Regulation, Deficiencies, Minimum Quality Standards, Targeting, Nursing Homes

JEL Classification: I180, L510, K230

1 Introduction

Despite requirements to meet hundreds of minimum care standards, U.S. nursing homes have a history of quality concerns including inadequate staff-resident ratios and infection control practices. Standards are enforced through a deterrence model based on unannounced inspections, during which inspectors observe facility practices and interview staff and residents. Regulators have the power to fine, deny payment, or even decertify facilities who fail to meet minimum care standards.

During annual inspections, 20% of facilities are found to have care practices that inspectors deem as causing "actual harm" to residents, leading residents to have unsupervised falls and accidents, acquire preventable bedsores, and contract potentially dangerous infections (United States General Accounting Office, 1999). Care quality has also been found to persist in the same facilities over time (Grabowski and Castle, 2004; Walshe and Harrington, 2002). However, enforcement remains largely one-size-fits-all; facilities with proven records of excellence receive the same inspection frequency and scrutiny, as those with severe and persistent deficiencies.

When inspectors find quality issues, state health departments recommend enforcement actions, which the Center for Medicare and Medicaid Services (CMS) carries out. Health departments tend to recommend the least stringent sanction available, which in practice means that most facilities receive opportunities to correct problems before sanctions are imposed. A common observation has been that facilities correct problems "just in time" to avoid mandatory termination, only to have the same problem resurface on the next inspection, which suggests the underlying problems have not been addressed.

In a simple model of nursing home markets, firms face a trade off between the quality of care they provide, which is costly, and profits, due to low competitive pressures in geographically segmented markets. Deterrence based regulation aims to raise the cost of failing to meet minimum care standards by adding an expected regulatory cost term to the firm's optimization problem, which is the product of the probability of detection times the cost of the sanction if detected.

The only federal exception to the one-size-fits-all regulatory approach is the Special Focus Facility (SFF) program. CMS started the SFF program in 1998 with the goal of improving quality of care by targeting oversight toward some of the industry's worst performing facilities. The SFF program has since targeted over 1,000 facilities using a three-pronged approach to raise the expected cost of failing to meet minimum care standards. First, by doubling the frequency of inspections to every six months, facilities face increased likelihood that any given care failure is discovered by regulators. Second, by requiring progressively stringent sanctions when deficiencies are discovered, the expected regulatory cost conditional on discovery, are raised as well. Third, the program applies additional scrutiny until facilities complete *two consecutive inspections* without any "harm-level" deficiencies, or, if it fails to achieve this in 2-3 years, risk decertification from Medicaid & Medicare programs, which in practice means the facility would shut down, imposing significant financial losses. Since 1998, the SFF program has doubled the frequency of inspections for over 1000 facilities and imposed swift, progressively harsh penalties, including threat of termination unless the facility complete two consecutive inspections without any findings of harm.

The current literature on enforcement of certification requirements provides mixed evidence for whether firms respond to regulatory stringency by improving care quality, but effects are hard to estimate causally because (a) homes are subject to the same requirements, (b), enforcement actions are endogenously related to quality, and, (c), the most commonly evaluated outcome, inspection results, could reflect both changes in quality but also changing stringency of state inspectors (Walshe, 2001). A final, difficulty, (d), is that inspection results have a natural tendency for mean-reversion, which means that year-to-year within-facility changes can overstate the effects of enforcement actions without a credible counterfactual.

To overcome the identification challenge in (b), this paper leverages the program's capacity constraints, which are capped by state. The cap has meant that while each state has many poor performers, similar on observable characteristics and performance trajectories, the program can only target a fixed number of facilities at any given time, creating close-to-random variation in which firms are treated. I exploit this in a difference-in-difference framework to produce the first causal effects estimates of targeted enforcement on nursing home care standards and health outcomes.

To recover a causal estimate of the SFF program's *additional targeting* I use data from a Freedom of Information Act request of facilities that were *candidates* for the program, but that were not targeted. Restricting comparisons to the most similar candidate facilities in the same state recovers 600 policy experiments (cohorts) of facilities who followed similar paths in the years leading up to treatment. Restricting comparisons to come from within state-years overcomes (c), because facilities are subject to the state's same overall stringency, and comes close to overcoming (d), because any mean-reversion tendency should be similar for treated and untreated facilities.

I combine generalized difference-in-difference specifications with Coarsened Exact Matching to further overcome (d) and find comparisons with parallel trends leading up to treatment. The average targeted facility improve significantly more than their counterparts; two years after treatment, the average targeted facility receive two fewer deficiencies, an additional 17% reduction relative to the average candidate, who improves 48%. These results support the hypothesis that additional oversight benefits patients.

The alternative hypothesis is that additional oversight does not benefit patients; treated facilities would have improved more than candidates irrespective of the SFF-program. This could only be due to a time-varying factor that starts affecting treated facilities (more), just after treatment is assigned. However, this alternative hypothesis gives a clear prediction: we expect targeted facilities to be unaffected when graduating from the program. I find clear evidence that performance deteriorates after treated facilities graduate the SFF program. When considered together, these two findings are hard to reconcile with alternative explanations; the omitted factor(s) would have to both *turn on* and then *off again* disproportionately for treated facilities *at the exact right times*. I hold that this is implausible considering that enrollment and graduation events happen across such vast set of time-space combinations.

A more pressing concern is whether the program has unintended consequences. There is a large literature documenting the difficulty of crafting regulatory interventions to raise quality of care, because quality is multi-dimensional, interventions that raise quality on one margin, will often be, at least partially, offset by on other margins (Bowblis et al., 2012; Bowblis and Lucas, 2012; Chen and Grabowski, 2015; Konetzka et al., 2013). A primary concern is therefore that the program may lead to unintended consequences, such as increased use of physical or chemical restraints (anti-psychotic medications). I find little evidence of this when breaking down survey scores into 15 detailed subcategories, which largely mirror the overall findings.

Section 2 summarize the literature on nursing home enforcement, regulation, and their impacts on quality of care. Section 3 describes the SFF program including changes to the program (over time) that are important to the empirical analysis. Section 4 describes the data and provides summary statistics. Section 5 and 6 describes the empirical framework and estimation techniques and interprets the results across various specifications and data sources. Section 7 concludes.

2 Literature Review

There is a large literature devoted to the documentation of quality of care in nursing homes, variables that are associated with differences in care, and policy aimed at incentivizing better care. I will review the literature that considers regulatory mechanisms to incentivize quality improvements, including two descriptive papers on the SFF program, and then review an important paper by Hackmann (2019) who argues for increasing Medicaid payment rates.

Harrington and Carrillo (1999) were among the first to analyze nationwide deficiency trends for all U.S. nursing homes and found that the number of deficiencies declined by 44% and the number of firms with perfect survey scores doubled from 1991-97. While Harrington and Carrillo (1999) acknowledge studies demonstrating "innovative efforts to reduce the use of restraints, to improve incontinence care, and to make other improvements in care" they are ultimately skeptical of interpreting the decrease in deficiencies as improvements in quality, and offer alternative interpretations such as the industry learning to avoid detection and regulators becoming less vigorous over time. They also cite Toby Edelman, who argued enforcement had been watered down by allowing firms the opportunity to correct deficiencies before applying sanctions, and by changing interpretive guidance such as the term "widespread" to only cover violations that affect all residents in the facility.

Walshe (2001) provides an early survey of the U.S. regulatory system.

> [t]he impact of regulation has not been much researched, in part perhaps because it presents several methodological challenges ... [a]ltough numerous studies have examined the implementation of nursing home regulation and the management of regulatory arrangements, these reports are of limited help in determining what impact regulation has had on nursing home performance and the quality of nursing home care."

He proceeds to describe three methodological/identification challenges to estimate the impact of regulation on performance. First, since (virtually) all nursing homes are regulated, no control group exists to compare regulated homes to, which "means that one can really only study changes in quality over time and attempt to determine whether those changes can be attributed to regulatory interventions". Second, it is challenging to distinguish changes in quality from changes in the regulatory process, using available data that is itself the product of the regulatory process. Third, Walshe notes that the reliability, validity, completeness, and timeliness of the available data has been questioned and suggests caution is needed to in analyzing survey data.

Despite progress in the past 20 years since Walshe (2001) and Harrington and Carrillo (1999) the task of overcoming these identification and methodological challenges remain incomplete. One strain of literature focuses on the first point raised by Walshe; that since all firms are regulated, no firms can be form a control group. This is probably true with respect to federal regulations on the *extensive margin*, as one cannot operate a federally certified nursing home without being subject to the federal participation requirements, However, it is clearly not true on various *intensive margins*.

Another strain of literature considers targeting survey resources toward low-performing facilities more specifically.

> "Federal and state survey efforts [should] focus more on providers that are chronically poor performers by surveying them more frequently than required for other facilities, increasing penalties for repeated violations of standards, and de-certifying persistently substandard providers"
>
> Wunderlich and Kohler (2001)
>
> Committee on Improving Quality in Long-Term Care
>
> Institute of Medicine

The idea was specifically addressed by Grabowski and Castle (2004) who are skeptical of whether enforcement will address the underlying causes of poor quality care

> "Clearly, targeted efforts to penalize low-quality facilities may be effective in the short run, but this proposal raises broader questions in the long run as to

> whether the root causes of persistently low quality will be addressed. That is, if low Medicaid payment rates or a lack of consumer information are the underlying sources of persistent low quality, it is unclear that simply shutting down chronic offenders will address the larger problem. The low-quality nursing home may not persist in a highly regulated environment, but the presence of low-quality care might"
>
> Grabowski and Castle (2004)

Three articles are published specifically on the Special Focus Facility program. Castle and Engberg (2010) provide a descriptive examination where they compare the certification scores and MDS quality measures of facilities that participated in the program during 2007 with those of all other facilities. Castle and Engberg (2010) find that SFF facilities receive on average twice as many deficiencies (12.36 vs 6.91) and quality of care deficiencies (2.80 vs 1.50), as well as nine times as many deficiencies for placing residents under immediate jeopardy of health and safety (0.36 vs 0.04). They also find that residents in SFFs are prescribed more anti-psychotic medications (30.80% vs 25%) and are more frequently put under physical restraints (7.1% vs 5.4%), and conclude that CMS succeeded in targeting facilities of poor quality during the year studied (2007).

Castle et al. (2010) examine whether SFF participation of one facility has spillover effects on other facilities in the same county. They use data from 2007-08 and compare quality provided by firms in a county where one firm had been enrolled in the SFF program to quality by firms in all other counties (excluding counties with only 2 firms, reducing the sample to 123 firms out of the 135 SFF firms in 2007). Castle et al. (2010) find little evidence of spillover effects of the SFF program. Of 22 quality outcomes, they find changes are significantly different in SFF counties on six; Urinary Tract Infections (UTI's) and pressure sores for high risk long-stay, low risk long-stay, and short stay residents improve disproportionately in SFF counties, however, total citations and quality of care citations both worsen (increase). As the authors note, the analysis is limited by only covering one year of SFF facilities. The estimates cannot be given a causal interpretation as quality in SFF-markets are likely to differ for reasons other than SFF assignment, a point the authors note.

States can have standards that are more strict than federal standards, or chose to interpret federal standards more stringently, thus providing another source of treatment and control groups. Bowblis and Lucas (2012) estimates the effects of different regulatory stringency and minimum staffing requirements across states and over time on survey deficiencies facility level deficiencies. Bowblis and Lucas (2012) finds because quality is multidimensional, improvements on one margin can come at the expense of deterioration on another. In particular, they find that higher direct care staffing requirements reduce the use of feeding tubes but increase the use of physical restraints.

Another strain of literature uses instrumental variable techniques to attempt to address endogeneity in regulatory stringency. Mukamel et al. (2012) instruments for statewide differences in regulatory stringency with area 2 of the Economic Freedom Index of North America of 2010; "Takings and Discriminatory Taxation", and finds that stringency increase certain kinds of staff hours per resident (Certified Nurse Aides), but reduce others (Registered Nurses), leading to fewer pressure sores.

Miller and Mor (2008) review regulatory systems for long term care providers including nursing homes, assisted living facilities, home health agencies, and even daycare centers. They emphasize the tension between the regulator's role in policing standards versus that of consulting with providers to improve, and notes that practices vary over time as well as both within and across states. Miller and Mor (2008) use Hurricane Katrina and publicly available data from St. Rita's Nursing Home in New Orleans as a case study of the possibility of using real-time data in identifying residents who were particularly vulnerable. They argue that the publicly available data showed residents had been virtually abandoned years prior to Katrina and that more aggressive oversight from regulators and state officials could have prevented the literal abandonment that followed, resulting in 34 resident drowning.[1]

[1] See Gruneir and Mor (2008) for a different perspective. Gruneir and Mor (2008) points out that the history of nursing home regulatory policy is characterized by a repeating cycle of public scandal, which gar-

Like the IOM report (Wunderlich and Kohler, 2001), Miller and Mor (2008) envision a "smarter" regulatory approach, in part based on the idea of more targeted enforcement. They note

> top performing providers deemed fully immersed in continuous quality improvement might instead be subject to state surveys every two to three years

while providers that fail to make sufficient progress "might be required to undergo more frequent visits by state inspectors." They also favor more explicit incentives including "less/more frequent inspections, lower/higher fines and other penalties" to induce providers to improve quality, both "on their own" and through consulting Quality Improvement Organizations (QIOs).

Hackmann (2019) builds a structural supply and demand model of the nursing home industry in Pennsylvania, estimates the parameters of this model, and simulates the effects of increased Medicaid reimbursement rates and competition. The model quantifies that residents value an additional skilled nurse at $133,000, which exceeds the annual costs of $83,000 (including wages and fringe benefits, both in 2002 dollars). He shows that staffing ratios are inefficiently low in 96% of nursing homes.[2] A simulation finds a 10% increase in the Medicaid reimbursement rate would increase nursing homes skilled nurse hours per resident by 8.7%. In comparison, a new public facility entering rural markets - where gains from additional competition presumably are large as they tend to be served by a handful facilities - would raise staffing ratios by less than 1%.[3]

ners attention until it is met by a regulatory crackdown. Notably, they write that "it may be the adversarial environment within which nursing homes operate that poses the largest barrier to quality improvement" and further voice broader concerns regarding the deterrence model as a channel for enacting quality improvement which they argue "pits the regulatory body against the industry and complicates the development of productive and responsive relationships between the two" which precludes more official involvement of other industry stakeholders such as CMS sponsored Quality Improvement Organizations (QIOs).

[2]While estimates are based on data from just one state, it is unlikely to drive results as Pennsylvania has a slightly higher Medicaid reimbursement rate than the national average.

[3]Hackmann's estimates also show that 45% of the increased reimbursements are kept as profits, while 55% are passed on to consumers through higher staffing ratios (or lower prices). Higher reimbursement rates also lead to a considerable market expansion effect which, in Pennsylvania would increase the cost to taxpayers from $228 million (holding demand constant), to $331 million.

3 Nursing Home Enforcement and the Special Focus Facility (SFF) Program

Nursing homes must provide care that meets federally imposed minimum care standards in order to receive Medicare and/or Medicaid payments. While most requirements are federal, enforcement is carried out separately by each state, typically through departments of health, which conducts surveys and certifies compliance or non-compliance.[4] Compliance with care standards are primarily enforced through unannounced inspections every 9-15 months, where an interdisciplinary survey team consisting of social workers, dietitians, pharmacists, rehab specialists, and at least one registered nurse (RN) investigate the home's care practices (CMS, 2018a).

Examples of commonly violated standards, or deficiencies, include failure(s) to prevent avoidable accidents with adequate supervision or keep areas free from hazards (F-323), maintain an effective infection control program (F-441), or provide care that maintains the dignity and respect for each resident (F-241). All deficiencies are scored according to the severity of harm (or potential harm) posed to residents and the number of residents affected (or with potential to affect). The Social Security Act Section 1819(h)(2)(C) requires any nursing home that does not achieve substantial compliance within 6 months, defined as having a deficiency rated G or higher, to be terminated. The Scope & Severity Table is reproduced in Appendix Table 1.

The regulatory system for Nursing Homes was historically based on a more "informational and cooperative model" where surveys informed providers of failures to meet federal standards. The focus of these surveys were typically with respect to the physical environment and facility management. The Omnibus Budget Reconciliation Act of 1987 (Public Law 100–203) changed the focus to a deterrence model that use penalties to deter firms from committing care failures (Harrington et al., 2004). This model remains in place to-

[4] Some states have additional (or more stringent) requirements than the federal government.

day and is subject to ongoing debate among policy makers. To deter homes from going out of "substantial compliance" and encourage swift return for those that do, states recommend enforcement actions ranging from a directed plan of correction, Civil Money Penalties (CMPs), denial of payment for new admissions (DPNAs), or even termination of provider agreement (decertification) depending on the scope and severity of the deficiencies. CMS, who ultimately imposes the penalties, could in theory override the state's recommendation and impose harsher or softer penalties, but in practice this is rare.

The change to a deterrence model did not eliminate substandard quality care. Proponents argue this is due to states and CMS being overly lenient with enforcement. The average survey reveal about seven deficiencies, but 88% of all deficiencies are rated as posing "no actual harm, with potential for minimal harm", or "more than minimal harm that is not immediate jeopardy", which leads to a maximum possible fine of about $2000 (CMS, 2018b). A common observation is that states frequently recommend the least stringent sanction available, and only after giving the facility an opportunity to correct. This opportunity allows the facility to correct the problem within a certain date, typically a couple months, without any penalty, and CMS rarely override a state's recommendation.

In the cases where fines are imposed, facilities automatically receive an offer of a 35% reduction if they accept the fine, further reducing potential deterrence effects. Therefore, it is not surprising that the threat of fines and penalties appear to have had limited deterrence effects if we consider a firm scenario where providing the required level of care is costly, and where competitive pressures alone might be insufficient to induce the firm to provide the minimum quality level.

In 2016, nursing home expenditures totaled $170 billion across 15,500 facilities, which amounts to about 5% of U.S. health care spending (Hackmann, 2019). A typical home generates about $10,000,000 in annual revenues and pays $0 fines.[5] In comparison, CMS have in recent years collected between $40 and $80 million annually from CMPs, with the

[5] I abstract away from Payment Denials as they are more difficult to quantify.

largest penalty levied being approximately $1.25 million. Most facilities compete in narrowly defined local markets as the median resident chose a facility within four miles of her former residence (Hackmann, 2019). These institutional details suggest, and empirical evidence supports the notion that, for a fraction of nursing homes, the rational profit maximizing strategy (absent regulatory interventions like the SFF program) is to provide care that does not always meet the minimum standards laid out in Chapter 7 of the State Operations Manual (CMS, 2018a). Facilities would rationally expect that not all deficiencies will be discovered, and for those that are, facilities will be given opportunities to correct. Further discounts means that the expected sum of $P_{detection} x Penalty$ is not sufficient for the profit maximizing strategy to be to hire more staff or invest in expensive equipment.

This is what has been observed in practice. A portion of homes consistently fail standard surveys, and improve just enough for their correction dates that the scope and severity gets below G, which let's the facility stay certified, only for subsequent inspections to find more deficiencies, often in the same category. While the deficiency need not be from the exact same care failure, it remains the case that repeated deficiencies in the same cateogry rated G or higher, suggests there are underlying systematic problems that have not been addressed.

The SFF program was established in 1998 as part of the Nursing Home Oversight and Improvement Program with the goal of changing the incentives for facilities with persistent care failures, essentially raising the probability of detecting care failures and the penalties if failures are detected (CMS, 2004). This is put in place by inspecting facilities twice as often. The program requires the state to enforce progressively strict penalties.[6]

Homes "graduate" from the program by showing substantial compliance on two consecutive inspections, which they are expected to do within 18-24 months. If a facility does not achieve this it will be notified that the next inspection will be its last chance to achieve substantial compliance, or it will be subject to termination. SFF slots opens up when homes graduate or are terminated, at which point, the state is required to select a new facility

[6] For more details about this progressive enforcement, see Figure 13 in the Appendix.

from the candidate list within 21 days (CMS, 2017b). I argue these procedures create quasi-random variation in both which firms end up treated, and when treatment occurs.

The program currently targets 88 nursing homes, with a fixed number of program slots for each state ranging from one (29 states) to six in California and Texas.[7] CMS ranks facilities within each state based on a weighted average of its three most recent years of inspection scores, and designates the lowest ranking firms in each state as candidates for the program.[8] Each SFF slot has five candidate slots, so in a state with two SFF slots, the ten lowest ranked facilities are automatically candidates.

The size of the program and distribution of entry and exit periods is shown in Figure 1-3. A feature of the program's design, continually enrolling, graduating, and terminating facilities, means that treatment periods are spread across time and facilities such that no particular periods can be driving the results. The size and geographic distribution of the SFF program has changed at various points, largely due budgetary considerations.

The number of program slots was 135 from February 2005 until October 2010, when this was raised to 167 (CMS, 2004; Casey and Toomey, 2019). The program was reduced to 48 in April 2013 when the Balanced Budget and Emergency Deficit Control Act, commonly known as the "sequester," went into effect. In May 2014 the program was increased to it's current size of 88 facilities (CMS, 2004, 2013; Hamilton, 2014). Treated facilities have typically served a quarterly flow of between 5-15,000 residents over this period.

[7]See the appendix Figure 12 for the program's current geographic distribution.

[8]The most recent year is weighted 50%, last year 33%, and two years prior 17%.

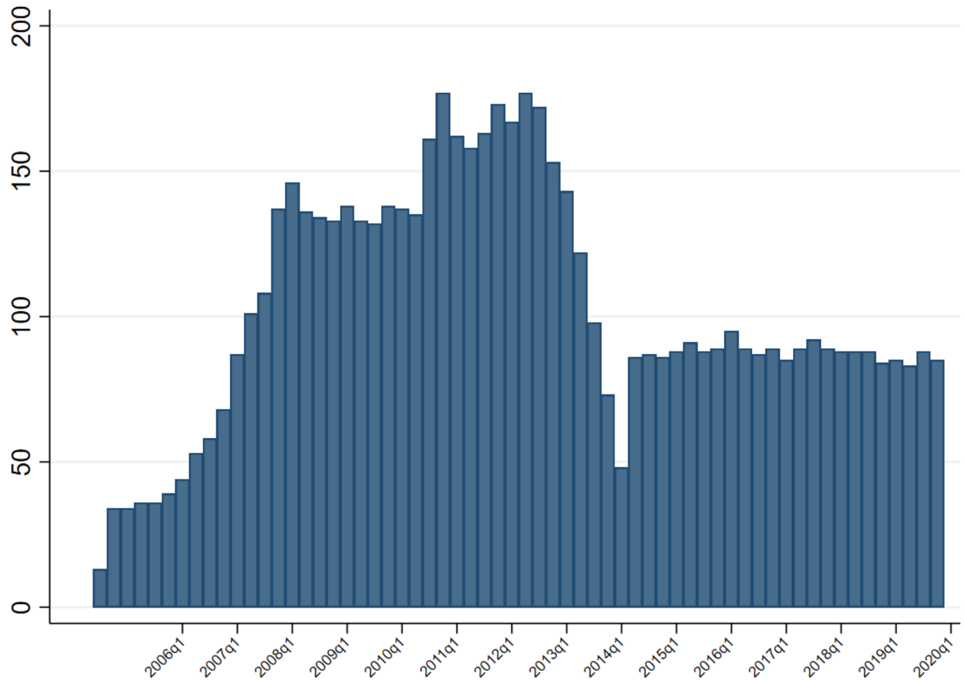

Figure 1: SFF Program Size

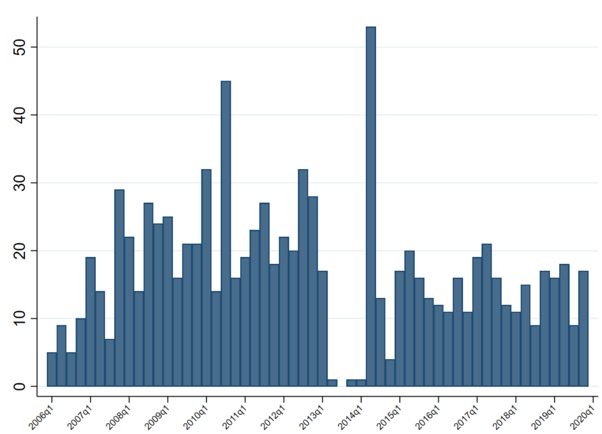

Figure 2: Enrollments

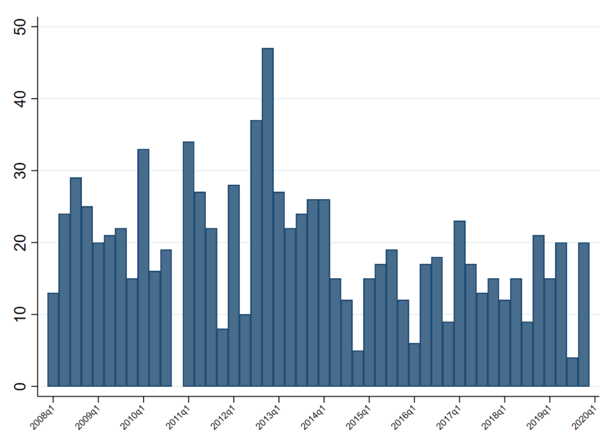

Figure 3: Exits

Note: Figure 1 shows the size of the SFF program from 2008 through 2019. The number of slots grew from 135 to 151 in September 2010, and was reduced to 48 in March 2013 when budget cuts in the Balanced Budget and Emergency Deficit Control Act of 2013 (Sequester) went into effect. The number of slots was increased to 88 in May 2014 which remains in place today. Figure 2-3 shows the quarterly number of entries and exits.

4 Data and Descriptive Statistics

4.1 Nursing Home Compare Firm and Quality Data

My primary data source is Nursing Home Compare (NHC), a website maintained by the Center for Medicare and Medicaid Services (CMS) that provides a comprehensive way to compare nursing homes on dimensions such as care practices, clinical outcomes, inspection results, and firm characteristics.[9] I investigate whether the SFF program benefits patients by looking at inspection results from 2004 through 2019.

4.2 Special Focus Facility Data

The SFF program was started in 1998, but CMS did not not make public which firms were enrolled until February 2008. CMS started releasing the current candidate facilities in June 2019 (Hamilton, 2008). I obtained monthly candidate lists from November 2010 to July 2019 through a Freedom of Information Act Request. During this period I observe close to 850 facilities that have been targeted by the program and about 3,000 facilities that have been candidates.[10]

4.3 Deficiencies and Inspection Scores

The primary way I measure the care provided is through the survey outcomes homes receive from standard health inspections. The inspection process and requirements for participation are standardized and have only undergone minor changes between 1995 and 2019, making

[9] Because facilities receive about 70% of their revenues from Medicaid and/or Medicare payments, it is extremely rare for facilities to operate outside of these programs. As a result, we observe the universe of almost 16,000 nursing homes operating in the U.S.

[10] One potential concern is that facilities can receive new National Provider Identifiers (NPIs) during changes to certification status, ownership, or other legal changes. At best, this would make longitudinal analysis using NPI's noisy. But, if for instance below average firms strategically shut down and re-open, this could bias the analysis. To overcome this potential problem, I utilize a facility-identifier crosswalk produced by the Long Term Care Focus group at Brown University, which allows continued tracking of facilities from 2000-2017. I use address and geographic information to link facilities that received new identifiers in 2018-19(Focus, 2017). This does not appear to be a significant driver; only ten facilities appear to have been treated under different NPIs.

longitudinal analysis of nationwide comprehensive results available for each certified facility.[11] While surveyors might not catch all deficiencies, deficiencies are generally found to be a representative floor of care failures, because facilities have the opportunity to contest deficiencies, thus providing an incentive to eliminate frivolous deficiencies.

Surveyors review the home's compliance with minimum care standards and issue deficiencies when standards are not met. Each standard, referred to as F-Tags, are assigned a score ranging from 0 to 175 points according to the scope and severity of the violation with more points indicating more residents affected or a more severe potential for harm. This scoring system, commonly referred to as the Scope/Severity table, is reproduced in the Appendix as Table 1. The most natural way to evaluate performance is to analyze the overall deficiencies and scores received during each inspection.

Deficiencies are reported with the exact date of issuance, while most other data sources, including the home's SFF status, is generally available on quarterly frequency. To most accurately capture both short- and long-term impacts of the program I aggregate the survey scores to the quarter in which they were issued. This means that on average homes will only have one quarter with a new survey score each year, and will have missing values for the other quarters. Second, in the rare event that a home has a perfect survey, the deficiency score could be missing rather than 0, thus overstating the average scores. This is unlikely to be of concern for two reasons. Facilities receive on average 6-7 deficiencies and a score between 40 and 50, while homes enrolled in the SFF program on average receive twice as many deficiencies and scores that are more than twice as high. Earlier work has shown that about 10% of facilities, the very best in the industry, receive perfect scores Harrington et al. (2000). It is therefore unlikely that facilities that were ever associated with the SFF program would receive a perfect score.

[11] CMS issued a comprehensive overhaul of survey and participation requirements in 2016 to be implemented in three phases; the first phase was implemented in November, 2016 but added only minor changes (Department of Health and Human Services Centers for Medicare and Medicaid Services, 2016). Phase 2, implemented in November, 2017, updated the survey process itself including renumbering and reclassifying the F-Tags, and made the survey itself computer based, etc. CMS (2017a) I believe these changes are small enough that they have little impact even when inspections occur across survey methods.

5 Effects of Targeted Oversight on Nursing Home Quality

To evaluate the effects of targeted enforcement on care standards and health outcomes, I estimate a series of generalized difference-in-difference models. The treatment starts at different times for different facilities and also has variable duration as facilities take different number of inspections to meet (or fail to meet) graduation requirements. To see this graphically see Figure 1- 3. I combine this difference-in-difference framework with the Coarsened Exact Matching (CEM) technique developed by Iacus et al. (2012). This identification strategy has been used when there is (a) variation in treatment timing, (b), a limited number of untreated units that can serve as plausible controls, and (c), the researcher wants to allow some control units to serve as controls for multiple treated units. Similar identification and estimation strategies have been used in by Jeon and Pohl (2017) and Rellstab et al. (2019).

5.1 Sample Selection of Treatment and Control Cohorts

I start with the problem state policy makers face when a Nursing Home has left the SFF program either due to meeting the "graduation" requirements or due to termination. The state policy maker must now enroll a new facility in no more than 21 calendar days, and must choose this facility from the most recent candidate list provided by CMS (CMS, 2017b).[12] This event occurs in 608 unique "state-quarters" from 2011Q1-2019Q3.

I group each treated facility into cohorts based on the state-quarter of the "enrollment-event" and let the pool of facilities eligible for the corresponding control group be those that were on the candidate list during the time of the event, but that were not chosen. I restrict the control groups to facilities that entered no more than two quarters before or after the treated facilities. To make the samples more homogeneous, I also restrict the control groups facilities that were not treated in the past two years. Note that a small minority of facilities can fit the criteria for multiple cohorts, for instance if a state has multiple consecutive

[12] For more information about how this is carried out in practice, see Figure 14 in the Appendix for a "Model Letter" informing newly selected facilities of their enrollment.

quarters with enrollment-events. When this occurs the facility is duplicated (and given an additional identifier), which means the analysis include 608 cohorts where each cohort is an unbalanced panel of a total of 622 treated and 2,720 untreated facilities.

5.2 Difference-in-Differences

To estimate a difference-in-differences model, I define an event-time indicator k, which takes a value of the number of years the facility is away from an enrollment-event q_{it}^k, and interact this with an indicator, D, that is 1 for the facilities that will end up treated. However, when treated facilities leave the program, D is coded 0.[13] βD_{it}^k can therefore be interpreted as the difference-in-differences, k years relative to the treatment-event, except for $k = 4$ which bin together inspections that happen four or more years after enrollment for the minority of facilities that stay in the program for extended periods. Because facilities are selected based on three years of compliance history and typically graduate from the program in 12-24 months, we are most interested in the estimated β's for $k = -4, \ldots, k = 2$. The main estimating equation is

$$Y_{it} = \sum_{k=-15}^{4} \gamma^k q_{it}^k + \sum_{k=-15}^{4} \beta^k D_{it} q_{it}^k + [\alpha_i + \lambda_t + \delta_{st}] + \epsilon_{it} \quad (1)$$

where the first summation term captures the relative inspection-years for all facilities, while the second summation term is the one of interest and captures the effect of treatment k inspection-years relative to treatment assignment. The bracketed terms are used in various combinations, and include fixed effects for each facility, α_i, quarter, λ_t, as well as state-by-year indicators, δ_{st}. (1) is estimated using the within transformation.[14]

Facility indicators are used to absorb differences across facilities that are constant over time, state-by-year indicators to absorb statewide changes in enforcement stringency, and

[13] This is done to prevent $\beta^{k=1,\ldots,k=4}$ to be confounded by treatment reversals, which are estimated in a separate section.

[14] Note that some facilities serve as control units in multiple cohorts. In order to construct the relative-time variables these are duplicated and given "temporary" identifiers.

quarterly time indicators to absorb nationwide shocks that are common to all facilities.

My preferred specifications do not not control for the facility's occupancy rate or staff per resident, due to the concern that these are endogenous. The standard errors are clustered on the facility level to account for within-facility serial correlation, but note that alternative clustering strategies such as by cohort had little impact.

For β^k to have a causal interpretation in (1), it has to be plausible that quality among firms *not chosen* for the program, are an unbiased approximation of what would have happened to quality among chosen firms, had they not been chosen - the standard parallel trends assumption. This assumption is fundamentally untestable, but I follow common practice and test whether the path of quality evolved similarly prior to enrollment by inspecting the β^k's prior to treatment.

Figure 4 and 6 plots the average outcomes for treated and control facilities. Figure 5 and 7 are weighted based on the Coarsened Exact Matching (CEM) procedure described below. The main conclusion is that candidates who end up treated follow similar paths as those who don't from seven years before until to two years before treatment. Some candidates who end up treated have especially pronounced deterioration just prior to treatment, but this is driven by a small number of outliers.

I apply the Coarsened Exact Matching (CEM) technique put forth by (Iacus et al., 2012) to make the treatment and control groups even more homogeneous. CEM applies an exact-matching algorithm that temporarily coarsens pre-treatment variables provided by the user, and assigns units to different strata according to each possible combination of these (coarsened) pre-treatment variables. Within each strata, CEM calculates weights that balance the empirical distribution of each strata, where treated units receive weights = 1. Treated or untreated units that don't meet the common support of any strata are given weights = 0.

I match facilities based on the outcomes at $k = -3$, $k = -2$ and $k = -1$. This drops a small number of facilities, depending on which outcome is examined. For the regressions on deficiencies (scores) 24 treated (14) and 625 (188) control units are dropped because they were outside the common support area. Summary statistics from the last period before treatment are shown in Table 1, which shows that treated and control facilities are comparable in terms of pre-treatment outcomes, patient mix, and firm characteristics even before matching. It is further clear that the matching procedure is able to find units that are even more comparable on pre-treatment outcomes, without distorting the balance on other variables that could influence care quality.

Table 1: **Descriptive Statistics Last Inspection Before Treatment**

	Control mean	Treated mean	Matched Control mean	Matched Treated mean
Health Score	146.2	184.6	182.7	181.4
Deficiencies (#)	12.6	14.3	13.6	14.3
Age	76.7	76.4	76.6	76.3
Black (%)	23.1	22.3	24.0	22.0
Hispanic (%)	6.28	5.29	6.55	5.31
Female (%)	63.5	63.2	63.5	63.1
Acuity Index NCMI (0-4)	1.14	1.14	1.14	1.14
Medicaid (%)	67.7	69.2	67.9	69.2
Medicare (%)	12.1	12.3	12.1	12.3
For profit (%)	0.81	0.85	0.83	0.85
Residents (#)	97.1	95.3	98.1	95.0
Beds (#)	121.0	123.1	122.2	123.1
Rehospitalization Rate	0.20	0.20	0.20	0.20
Successful Discharge Rate	0.45	0.46	0.44	0.46

Note: This table compare treated and control units on inspection results, resident, and facility characteristics the last period before treatment.

I then plot the average outcomes by treatment status, where the plots on the right are weighted. The CEM-weights make treated and control units follow each other even more closely at $k = -3$ $k = -2$ and $k = -1$, which we expected since these were used for matching. However, it is reassuring that this procedure does not seem to have come at the expense of fundamentally changing the paths from $k = -7$ to $k = -4$. This reinforces the premise of the research design that there are facilities for which treatment appear conditionally random. Following common practice we can then test for pre-treatment differences in trends by inspecting the estimated $\beta^{-7}, ..., \beta^{-1}$'s and 95% confidence intervals from estimating (1) while applying the discussed CEM-weights.

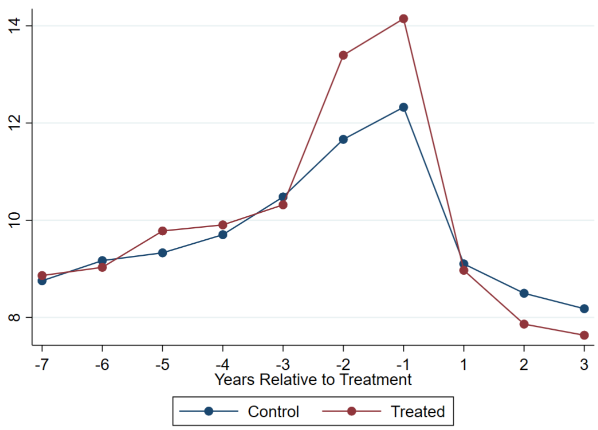

Figure 4: Health Deficiencies

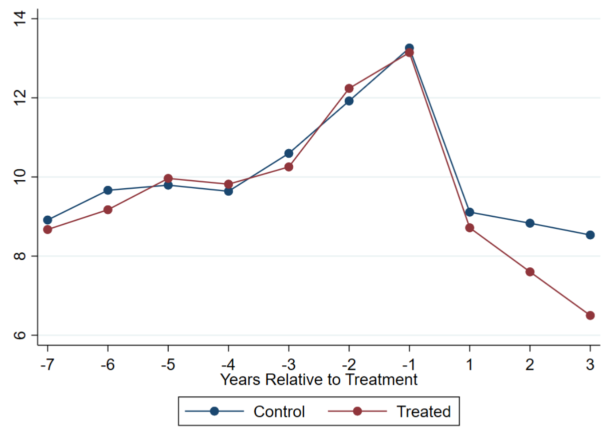

Figure 5: Health Deficiencies CEM-Weighted

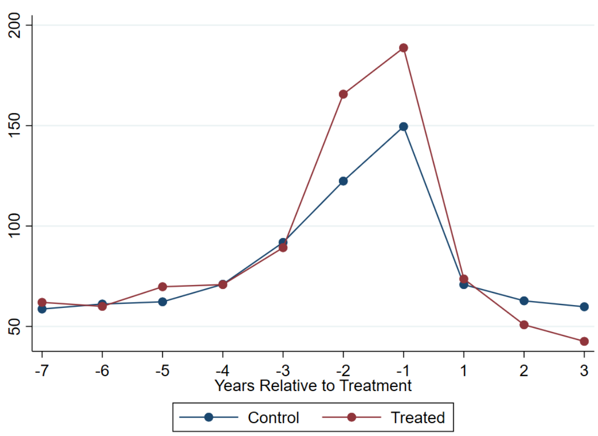

Figure 6: Health Score

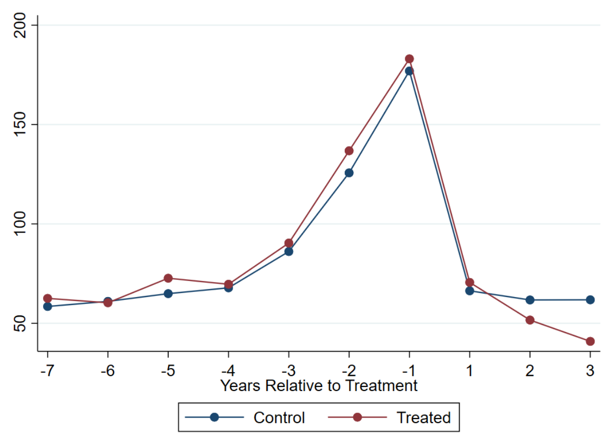

Figure 7: Health Score CEM-Weighted

Note: These figures plot the average outcomes for treated (red) and untreated (blue) facilities. Figures on the left are raw averages, while figures on the right are weighted by the weights from the CEM matching procedure. A lower score or fewer number of deficiencies represents a better survey result.

6 Effects of Enforcement on Deficiencies and Inspection Scores

Figure 8 and 9 plots the β^k's from (1) on the health survey score for four health surveys prior to, and three following, the enrollment event showing two clear patterns. First, treated and control facilities are on similar paths leading up the (placebo) treatment. The parallel trend assumption appears to be credible, as both treated and untreated facilities experience a similar large performance swing leading up to treatment.[15] The second pattern is that treated facilities improve survey results more than untreated candidates.

Some researchers have cautioned against matching on pre-treatment outcomes while simultaneously using unit fixed effects to difference out permanent differences Chabé-Ferret (2017). I therefore include various specifications in Tables 2 and 3, and pay most attention to the results in columns 4 and 5 which include the CEM weights.

The estimated treatment effect indicates that treatment caused an additional improvement of approximately 1-2 deficiencies, or an additional 7-20%, for facilities enrolled 1-3 years after treatment begun, compared to untreated candidates. The difference is typically statistically significant by the third year, i.e. it appears to get larger the longer facilities stay in the program. This is surprising because facilities that remain in the program three years after enrollment are those that have not yet been able to graduate, and thus one might think this selection would lead the coefficient to shrink.

The estimates on the overall health score are somewhat noisy, but tell a similar story. Treated facilities receive on average 27 fewer survey points (S.E. = 15.1) three years after treatment, or an additional improvement of about 15%.

[15] As a rough measure of this, the coefficients on γ^k, which captures the shared event-time between treated and untreated facilities, are typically 10 times larger during the periods leading up treatment compared to the β^ks.

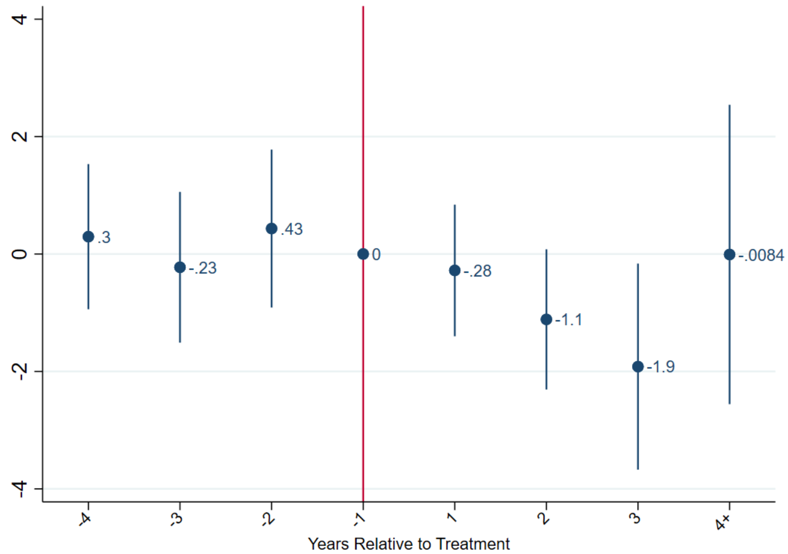

Figure 8: Treatment Effect on Health Deficiencies

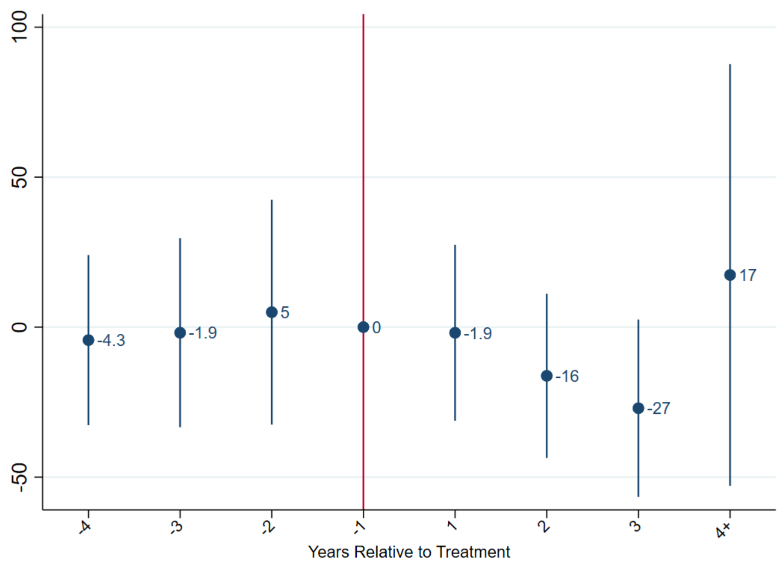

Figure 9: Treatment Effect on Health Scores

Note: These figures plot event-study difference-in-difference coefficients on health survey scores and deficiencies from (1). A lower score/number of deficiencies represents a better survey result.

Table 2: **Main Results for SFF Program on Deficiencies and Health Scores**

	(1) Deficiencies b/se	(2) Deficiencies b/se	(3) Deficiencies b/se	(4) Deficiencies b/se	(5) Deficiencies b/se
k=-4	-0.21	0.34	0.24	0.30	0.24
	0.40	0.41	0.40	0.63	0.62
k=-3	0.15	0.34	0.25	-0.23	-0.12
	0.42	0.42	0.40	0.65	0.65
k=-2	1.15**	1.16**	1.09**	0.43	0.31
	0.43	0.43	0.42	0.69	0.68
k=-1	0	0	0	0	0
	.	.	.	.	.
k=1	-1.27***	-0.96**	-1.02**	-0.28	-0.59
	0.34	0.34	0.33	0.57	0.55
k=2	-2.20***	-1.62***	-1.82***	-1.11	-1.97**
	0.39	0.39	0.39	0.61	0.60
k=3	-3.38***	-2.70***	-3.17***	-1.92*	-3.99***
	0.65	0.66	0.65	0.89	0.94
k=4+	-2.38**	-1.40	-1.13	-0.0084	-1.75
	0.82	0.81	0.73	1.30	1.26
Facility	Yes	Yes	Yes	No	Yes
CEM	No	No	No	Yes	Yes
Quarterly	No	Yes	No	No	No
State x Year	No	No	Yes	No	No
Mean of Y	9.35	9.35	9.35	9.72	9.72
N	63759	63759	63752	32310	32309

Note: Standard errors are clustered on the facility level. * 0.05 ** 0.01 *** 0.001

Table 3: **Main Results for SFF Program on Deficiencies and Health Scores**

	(1) Health Score b/se	(2) Health Score b/se	(3) Health Score b/se	(4) Health Score b/se	(5) Health Score b/se
k=-4	-14.6	-8.8	-10.9	-4.3	-3.3
	9.8	9.8	9.7	14.5	14.4
k=-3	-12.3	-9.5	-11.9	-1.9	-1.3
	10.5	10.6	10.4	16.1	16.1
k=-2	26.4*	25.7*	23.7	5.0	2.6
	12.7	12.7	12.6	19.1	19.1
k=-1	0.0	0.0	0.0	0.0	0.0
	.	.	.	.	.
k=1	-25.2**	-20.9*	-23.0*	-1.9	-3.8
	9.1	9.2	9.1	15.0	14.9
k=2	-36.0***	-27.6**	-31.7***	-16.2	-22.3
	8.8	8.9	9.1	14.0	13.9
k=3	-52.5***	-43.1***	-49.1***	-27.0	-46.2**
	9.9	10.2	10.5	15.1	15.6
k=4+	-29.8*	-22.5	-18.8	17.4	2.7
	14.9	14.8	13.4	35.8	40.8
Facility	Yes	Yes	Yes	No	Yes
CEM	No	No	No	Yes	Yes
Quarterly	No	Yes	No	No	No
State x Year	No	No	Yes	No	No
Mean of Y	75.1	75.1	75.1	78.6	78.6
N	63759.0	63759.0	63752.0	37729.0	37728.0

Note: Standard errors are clustered on the facility level. * 0.05 ** 0.01 *** 0.001

6.1 Treatment Reversals

The results discussed above indicates that the SFF program cause targeted facilities to improve certification results more than untreated candidates, but it remains possible that targeted facilities would have improved more, irrespective of the SFF program. Given the regression specification in (1), this could only be due to some time-invariant factor that starts affecting targeted facilities (more) just as the facility becomes targeted. This could happen if treated facilities, who before the matching procedure receive slightly worse survey results, decide on their own and irrespective of the SFF program to improve care standards. While the goal of the matching procedure was to eliminate this possibility, it remains possible that this was not achieved.

This and other alternative explanations predict that facilities will be unaffected by treatment ending. I test this by estimating

$$Y_{it} = \tau_{it} Reversal + \alpha_i + [\kappa_c + \lambda_t + \delta_{st}] + \sum_{k=-17}^{3} \gamma^k q_{it}^k + \sum_{k=-17}^{3} \beta^k D_i q_{it}^k + \epsilon_{it} \tag{2}$$

where every variable is the same as before but where I now include an indicator, Reversal, that turns on for treated facilities when treatment reverses, and is zero otherwise. I code Reversal to zero for untreated facilities so that the fixed effects still control for previously mentioned shocks. I no longer code the relative time as missing after treatment reverses. The results from (2) are found below in Table 4-5. I also estimate a version of (2) where Reversal is coded as years from graduation events, which is plotted in Figure 10- 11. Both specifications indicate that treatment reversals are followed by significant performance deterioration that are of comparable size to the preceding two years of differential improvements.

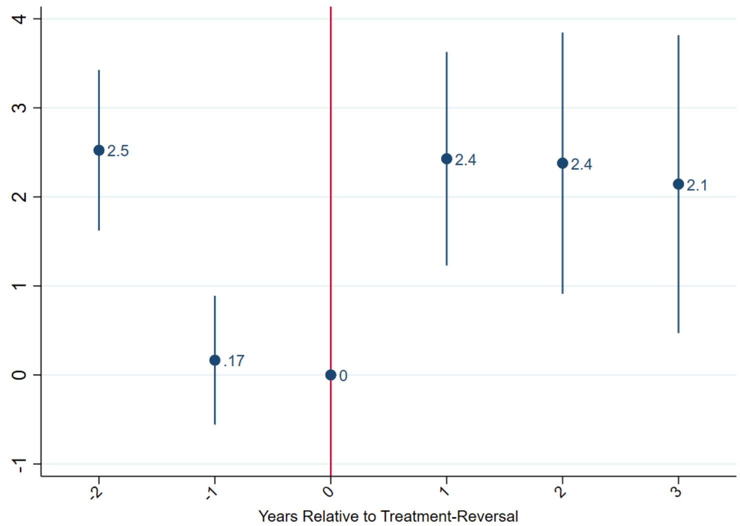

Figure 10: Treatment Reversal Effect on Deficiencies

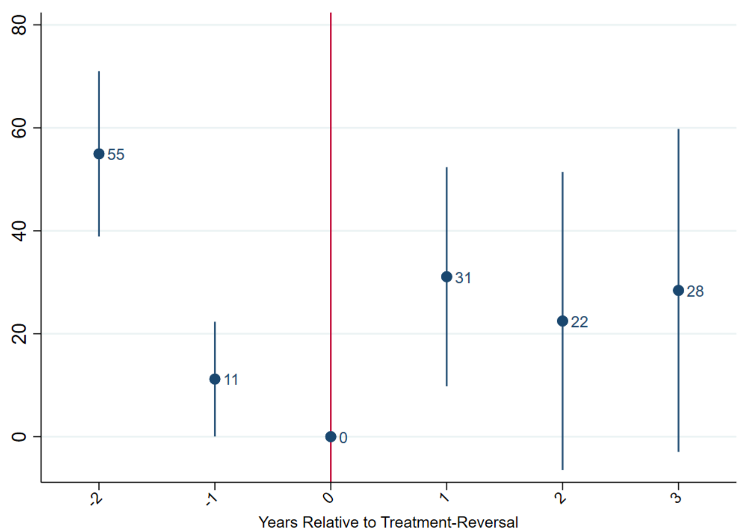

Figure 11: Treatment Reversal Effect on Health Score

Note: Figure above plot event-study coefficients from regressing health survey deficiencies and scores on treatment reversals (2). A lower number represents a better survey result.

Table 4: **Effect of Treatment Reversal on Health Scores**

	(1) Health Score	(2) Health Score	(3) Health Score	(4) Health Score	(5) Health Score
Treatment Reversal	21.66*	20.49*	23.03**	22.47**	21.69**
	(11.31)	(11.56)	(11.58)	(10.76)	(10.95)
Facility	Yes	Yes	Yes	Yes	Yes
Quarterly	No	Yes	Yes	No	Yes
Cohort	No	No	Yes	No	No
State x Year	No	No	No	Yes	Yes
N	52166	52166	52166	52157	52157
Mean of Dep. Variable	74.50	74.50	74.50	74.51	74.51

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: **Effect of Treatment Reversal on Deficiencies**

	(1) Deficiencies	(2) Deficiencies	(3) Deficiencies	(4) Deficiencies	(5) Deficiencies
Treatment Reversal	2.291***	2.122***	2.243***	2.385***	2.390***
	(0.633)	(0.646)	(0.654)	(0.599)	(0.600)
Facility	Yes	Yes	Yes	Yes	Yes
Quarterly	No	Yes	Yes	No	Yes
Cohort	No	No	Yes	No	No
State x Year	No	No	No	Yes	Yes
N	52166	52166	52166	52157	52157
Mean of Dep. Variable	9.497	9.497	9.497	9.497	9.497

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7 Discussion

This paper investigated whether a program that provides targeted enforcement of under-performing nursing homes causes the homes to improve care practices and ultimately the clinical outcomes of residents. The evidence I find is consistent with this hypothesis. Facilities that get additional oversight improve more than comparable facilities that don't. Facilities that receive *more* treatment appears to improve more, and finally, improvements revert after treatment ends, a set of findings that are hard to reconcile with explanations other than oversight being the cause.

References

John R Bowblis and Judith A Lucas. The impact of state regulations on nursing home care practices. *Journal of Regulatory Economics*, 42(1):52–72, 2012. ISSN 0922680X. doi: 10.1007/s11149-012-9183-6. URL https://link.springer.com/content/pdf/10.1007/s11149-012-9183-6.pdf.

John R. Bowblis, Stephen Steven Crystal, Orna Intrator, and Judith A. Lucas. Response to regulatory stringency: The case of antipsychotic medication use in nursing homes. *Health Economics (United Kingdom)*, 21(8):977–993, 2012. ISSN 10579230. doi: 10.1002/hec.1775. URL https://onlinelibrary-wiley-com.mutex.gmu.edu/doi/pdf/10.1002/hec.1775https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.1775.

Robert Casey and Patrick Toomey. Families' and Residents' Right to Know: Uncovering Poor Care in America's Nursing Homes. Technical report, 2019. URL http://www.startribune.com/senior-home-residents-are-abused-and-ignored-across-minnesota/450623913/.

Nicholas G Castle and John Engberg. An examination of special focus facility nursing homes. *Gerontologist*, 50(3):400–407, 2010. ISSN 00169013. doi: 10.1093/geront/gnq008. URL https://academic.oup.com/gerontologist/article-abstract/50/3/400/576194.

Nicholas G Castle, Kristen Sonon, and Jenya Antonova. The impact of special focus facility nursing homes on market quality. *Gerontologist*, 50(4):519–530, 2010. ISSN 00169013. doi: 10.1093/geront/gnq006. URL https://academic.oup.com/gerontologist/article-abstract/50/4/519/743000.

Sylvain Chabé-Ferret. Should We Combine Difference In Differences with Conditioning on Pre-Treatment Outcomes? Working Paper. 2017. URL https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/wp/2017/wp{_}tse{_}824.pdf.

Min M Chen and David C. Grabowski. Intended and Unintended Consequences of Minimum Staffing Standards For Nursing Homes. *Health Economics (United Kingdom)*, 24(7):822–839, 2015. ISSN 10991050. doi: 10.1002/hec.3063. URL https://onlinelibrary-wiley-com.mutex.gmu.edu/doi/pdf/10.1002/hec.3063.

CMS. Improving Enforcement via the Special Focus Facility Program for Nursing Homes Background, 2004. URL https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/SurveyCertificationGenInfo/Downloads/SCletter05-13.pdf.

CMS. FY2013 Sequestration Adjustments for Survey & Certification (S&C) 13-23-ALL, 2013.

CMS. SC18-04-LTC, 2017a. URL https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/SurveyCertificationGenInfo/Downloads/Survey-and-Cert-Letter-18-04.pdf.

CMS. CMS S&C: 17-20-NH, 2017b.

CMS. Chapter 7 - Survey and Enforcement Process for Skilled Nursing Facilities and Nursing Facilities, 2018a. URL https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/som107c07pdf.pdf.

CMS. Calculation of CMP Adjustments, 2018b.

CMS Department of Health and Human Services Centers for Medicare and Medicaid Services. Reform of Requirements for Long-Term Care Facilities. *Federal Register: Rules and Regulations*, 81(192):68688–68872, 2016. URL https://www.gpo.gov/fdsys/pkg/FR-2016-10-04/pdf/2016-23503.pdf.

Long-Term Care Focus. Shaping long term care in America Project at Brown University funded in part by the National Institute on Aging (1P01AG027296), 2017.

David C Grabowski and Nicholas G Castle. Nursing Homes with Persistent High and Low Quality. *Medical Care Research and Review*, 61(1):89–115, 2004. ISSN 10775587. doi: 10.1177/1077558703260122.

Andrea Gruneir and Vincent Mor. Nursing Home Safety: Current Issues and Barriers to Improvement. *Annual Review of Public Health*, 29(1):369–382, 2008. ISSN 0163-7525. doi: 10.1146/annurev.publhealth.29.020907.090912. URL http://publhealth.annualreviews.org.

Martin B Hackmann. Incentivizing better quality of care: The role of medicaid and competition in the nursing home industry. *American Economic Review*, 109(5):1684–1716, 2019. ISSN 19447981. doi: 10.1257/aer.20151057. URL https://doi.org/10.1257/aer.20151057.

Thomas Hamilton. Special Focus Facility (SFF) Program Survey Scoring Methodology. Technical report, Center for Medicaid and State Operations/Survey and Certification Group, 2008. URL http://www.cms.hhs.gov/certificationandcomplianc/downloads/sfflist.pdf.

Thomas E Hamilton. Fiscal Year (FY) 2014 Post Sequester Adjustment for Special Focus Facility (SFF) Nursing Homes. Technical report, 2014.

Charlene Harrington and Helen Carrillo. The Regulation and Enforcement of Federal Nursing Home Standards, 1991-1997. *Medical Care Research and Review*, 56(4):471–494, dec 1999. ISSN 1077-5587. doi: 10.1177/107755879905600405. URL http://journals.sagepub.com/doi/10.1177/107755879905600405.

Charlene Harrington, David Zimmerman, Sarita L. Karon, James Robinson, and Patricia Beutel. Nursing home staffing and its relationship to deficiencies. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 55(5):S278–S287, 2000. ISSN 10795014. doi: 10.1093/geronb/55.5.S278. URL https://academic.oup.com/psychsocgerontology/article-abstract/55/5/S278/536413.

Charlene Harrington, Joseph T Mullan, and Helen Carrillo. State nursing home enforcement systems. *Journal of Health Politics Policy and Law*, 29(1):43–73, 2004. ISSN 0361-6878. doi: 10.1215/03616878-29-1-43. URL https://muse.jhu.edu/article/52238.

Stefano M Iacus, Gary King, and Giuseppe Porro. Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20(1):1–24, 2012. ISSN 10471987. doi: 10.1093/pan/mpr013. URL http://gking.harvard.edu/cem;.

Sung Hee Jeon and R. Vincent Pohl. Health and work in the family: Evidence from spouses' cancer diagnoses. *Journal of Health Economics*, 52:1–18, 2017. ISSN 18791646. doi: 10.1016/j.jhealeco.2016.12.008. URL http://dx.doi.org/10.1016/j.jhealeco.2016.12.008.

R. Tamara Konetzka, Daniel Polsky, and Rachel M Werner. Shipping out instead of shaping up: Rehospitalization from nursing homes as an unintended effect of public reporting. *Journal of Health Economics*, 32(2):341–352, 2013. ISSN 01676296. doi: 10.1016/j.jhealeco.2012.11.008. URL http://dx.doi.org/10.1016/j.jhealeco.2012.11.008.

Edward Alan Miller and Vincent Mor. Balancing regulatory controls and incentives: Toward smarter and more transparent oversight in long-term care. *Journal of Health Politics, Policy and Law*, 33(2):249–279, 2008. ISSN 03616878. doi: 10.1215/03616878-2007-055. URL https://read.dukeupress.edu/jhppl/article-pdf/33/2/249/359261/JHPPL332{_}06{_}Miller.pdf.

Dana B Mukamel, David L Weimer, Charlene Harrington, William D Spector, Heather Ladd, and Yue Li. The effect of state regulatory stringency on nursing home quality. *Health Services Research*, 47(5):1791–1813, 2012. ISSN 00179124. doi: 10.1111/j.1475-6773.2012.01459.x. URL https://onlinelibrary-wiley-com.mutex.gmu.edu/doi/pdf/10.1111/j.1475-6773.2012.01459.x.

Sara Rellstab, Pieter Bakx, Pilar García-Gómez, and Eddy van Doorslaer. The kids are alright - labour market effects of unexpected parental hospitalisations in the Netherlands. *Journal of Health Economics*, 69, 2019. ISSN 18791646. doi: 10.1016/j.jhealeco.2019.102275.

United States General Accounting Office. GAO Report to Congressional Requesters NURSING HOMES Additional Steps Needed to Strengthen Enforcement of Federal Quality Standards. Technical report, 1999.

Kieran Walshe. Regulating U.S. Nursing Homes: Are We Learning From Experience? *Health Affairs*, 20(6):128–144, nov 2001. ISSN 02782715. doi: 10.1377/hlthaff.20.6.128. URL https://doi.org/10.1377/hlthaff.20.6.128https://www-healthaffairs-org.mutex.gmu.edu/doi/pdf/10.1377/hlthaff.20.6.128https://www.healthaffairs.org/doi/pdf/10.1377/hlthaff.20.6.128.

Kieran Walshe and Charlene Harrington. Regulation of Nursing Facilities in the United States. *The Gerontologist*, 42(4):475–487, 2002. ISSN 0016-9013. doi: 10.1093/geront/

42.4.475. URL https://academic.oup.com/gerontologist/article-abstract/42/4/475/686614.

Gooloo S. Wunderlich and Peter O. Kohler. *Improving the Quality of Long-Term Care*. 2001. ISBN 0309503809. URL http://www.nap.edu.

8 Appendix

Table 1
Health Inspection Score: Weights for Different Types of Deficiencies

Severity	Scope		
	Isolated	**Pattern**	**Widespread**
Immediate jeopardy to resident health or safety	**J** 50 points* (75 points)	**K** 100 points* (125 points)	**L** 150 points* (175 points)
Actual harm that is not immediate jeopardy	**G** 20 points	**H** 35 points (40 points)	**I** 45 points (50 points)
No actual harm with potential for more than minimal harm that is not immediate jeopardy	**D** 4 points	**E** 8 points	**F** 16 points (20 points)
No actual harm with potential for minimal harm	**A** 0 point	**B** 0 points	**C** 0 points

Note: Figures in parentheses indicate points for deficiencies that are for substandard quality of care.
Shaded cells denote deficiency scope/severity levels that constitute substandard quality of care. See the Electronic Code of Federal Regulations (https://www.ecfr.gov/cgi-bin/text-idx?SID=9c4d022241818fef427dc79565aba4b5&mc=true&node=pt42.5.488&rgn=div5#se42.5.488_1301) for a definition of substandard quality of care.
* If the status of the deficiency is "past non-compliance" and the severity is Immediate Jeopardy, then points associated with a 'G-level" deficiency (i.e., 20 points) are assigned.
Source: Centers for Medicare & Medicaid Services

The number of SFF slots and candidates list for each State (effective May 1, 2014).

State	Required SFF Slots	Size of Candidate List	State	Required SFF Slots	Size of Candidate List
Alabama	1	5	Montana	1	5
Alaska	-	-	Nebraska	1	5
Arizona	1	5	Nevada	1	5
Arkansas	1	5	New Hampshire	1	5
California	6	30	New Jersey	2	10
Colorado	1	5	New Mexico	1	5
Connecticut	1	5	New York	3	15
Delaware	1	5	North Carolina	2	10
District of Columbia	-	-	North Dakota	1	5
Florida	3	15	Ohio	5	20
Georgia	2	10	Oklahoma	2	10
Hawaii	1	5	Oregon	1	5
Idaho	1	5	Pennsylvania	4	20
Illinois	4	20	Rhode Island	1	5
Indiana	3	15	South Carolina	1	5
Iowa	2	10	South Dakota	1	5
Kansas	2	10	Tennessee	2	10
Kentucky	1	5	Texas	6	30
Louisiana	1	5	Utah	1	5
Maine	1	5	Vermont	1	5
Maryland	1	5	Virginia	1	5
Massachusetts	2	10	Washington	1	5
Michigan	2	10	West Virginia	1	5
Minnesota	2	10	Wisconsin	2	10
Mississippi	1	5	Wyoming	1	5
Missouri	3	15	Total	88	435

Figure 12: Program Distribution

Progressive Enforcement Table

Surveys After SFF Selection	<u>No</u> Deficiencies cited at a Scope & Severity of "F" or Greater	Deficiencies at "F" or above (no improvement)	Immediate Jeopardy
1st Standard Survey	Complete 2nd Standard Survey	Immediately recommend remedy (CMP or DPNA at a minimum)	Recommend remedy and proceed to termination if not corrected.
2nd Standard Survey	Graduate (if 2 surveys with no deficiencies above "E")	Recommend more stringent remedy. Must be in substantial compliance at 6 months or face termination.	Recommend remedy and proceed to termination if not corrected.
3rd Standard Survey	If a facility has deficiencies at E or below on the 3rd Standard Survey after selection (but is not able to graduate due to findings at F or above on 2nd Standard Survey or LSC deficiencies greater than F), Schedule 4th Standard Survey.	If a facility has deficiencies at G or above at the 3rd Standard Survey, Triage- (1) Schedule a 4th standard survey or (2) Issue a termination notice	Recommend remedy and proceed to termination if not corrected.
4th Standard Survey	Graduate (if 2 consecutive surveys with no deficiencies above "E")	Triage - either (1) schedule 5th standard survey, or (2) issue a termination notice	Recommend remedy and proceed to termination if not corrected.
5th Standard Survey	Graduate (if 2 consecutive surveys with no deficiencies above "E")	Issue termination notice (timing may be extended but not beyond statutory timeframes).	Recommend remedy and proceed to termination if not corrected.

Figure 13: Progressive Enforcement

Appendix A

MODEL LETTER TO PROVIDER SELECTED AS A "SPECIAL FOCUS FACILITY"

IMPORTANT NOTICE – PLEASE READ CAREFULLY

(Date)

Nursing Home Administrator Name
Facility Name
Address
City, State, ZIP Code

Dear (Nursing Home Administrator)

Because of your facility's poor compliance history for the past three years, you have been selected as a Special Focus Facility (SFF) program. The purpose of this letter is to notify you of this designation and to explain what this designation means for your nursing home.

What Does This Mean?
You will be subject to two standard surveys per year instead of the one required by law. You can expect that we will be closely monitoring your facility with the desire that your facility can attain and maintain compliance.

How Does A Facility Get Removed From the SFF?
A nursing home may be removed from the SFF program when it demonstrates at two standard surveys that it has no deficiencies cited at a scope and severity level of "F" or greater and no intervening complaint-related cited at "F" or greater. A nursing home may also be removed through a termination action if it fails to make significant improvements in the 24 months (3 standard surveys) following its selection as a SFF.

Robust Enforcement for Lack of Significant Progress: CMS will impose an immediate sanction on a SFF that fails to achieve and maintain significant progress in correcting deficiencies on the first and each subsequent standard survey after a facility becomes a SFF. Enforcement sanctions will be of increasing severity. These will include a Civil Money Penalty and/or a Denial of Payment for New Admissions.

If, after 24 months and four surveys subsequent to being selected as a SFF, you fail to have made significant progress, a notice of termination from participation in Medicare and Medicaid will be issued. CMS will consider a facility's status and progress as a SFF in setting a reasonable assurance period before a home can reapply to participate in Medicare.

Can This Be Appealed?
Your selection as a SFF cannot be appealed. However, you still have the right to informal dispute resolution (see 42 Code of Federal Regulations §488.331) and to appeal the

Figure 14: Selection Letter

noncompliance that led to a remedy through an Administrative Law Judge of the Department of Health and Human Services. Specific requirements for requesting a formal hearing are contained in the notice of the imposition of the remedy.

It is our intent that you take the designation of a special focus facility seriously. We can help. We can refer you to helpful resources, including help from the (Name of State Quality Improvement Organization).

We are also sending a copy of this notice to (name of nursing home owner) and (name of mortgagee) to give them notice of the designation of SFF for your facility.

If you have any questions, please contact (name, title, address, phone number, fax number and e-mail address of appropriate survey agency official.)

Sincerely yours,

(Name and Title)

Cc: CMS Regional Office
(Name of Quality Improvement Organization)
(Name of Owner)
(Name of Mortgagee, if applicable)

Figure 15: Selection Letter