Predictably Unpredictable: Inspections and Quality in Health Care*

Ashvin Gandhi
UCLA & NBER

Andrew Olenski
Lehigh University

Maggie Shi
UChicago & NBER

March 21, 2025

Abstract

Inspections are a common tool used by principals to monitor and induce effort from agents. However, even unannounced inspections often occur on a routine basis, making their timing predictable. We study inspections in the nursing home context, leveraging detailed, high-frequency administrative data on facility staffing and patient health outcomes to study how the predictability of inspection timing shapes effort and quality. We find that nursing homes "slack off" during periods when they are unlikely to be inspected and ramp up efforts as inspections become more likely. Patient survival and other positive health outcomes follow a similar pattern, suggesting that efforts in anticipation of inspection meaningfully impact facilities' quality of care. To understand how the inspection regime affects care, we estimate a simple dynamic model of nursing home effort that we use to simulate counterfactual inspection regimes. We find that even without increasing the number of inspections, simply reducing the predictability of inspection timing can meaningfully increase effort and save lives.

*We are very grateful to Liran Einav, Amy Finkelstein, Joshua Gottlieb, Ariel Pakes, and Yu Fu Wong for helpful conversations, as well as seminar audiences at Whistler Junior Health Economics Summit. Michelle Gard, Sangwook Suh, Fuman Xie, and Yan Bo Zeng provided excellent research assistance. We thank Mohan Ramanujan and Carla Tokman for providing excellent assistance in accessing and managing the data. This project received financial support from Arnold Ventures.

1 Introduction

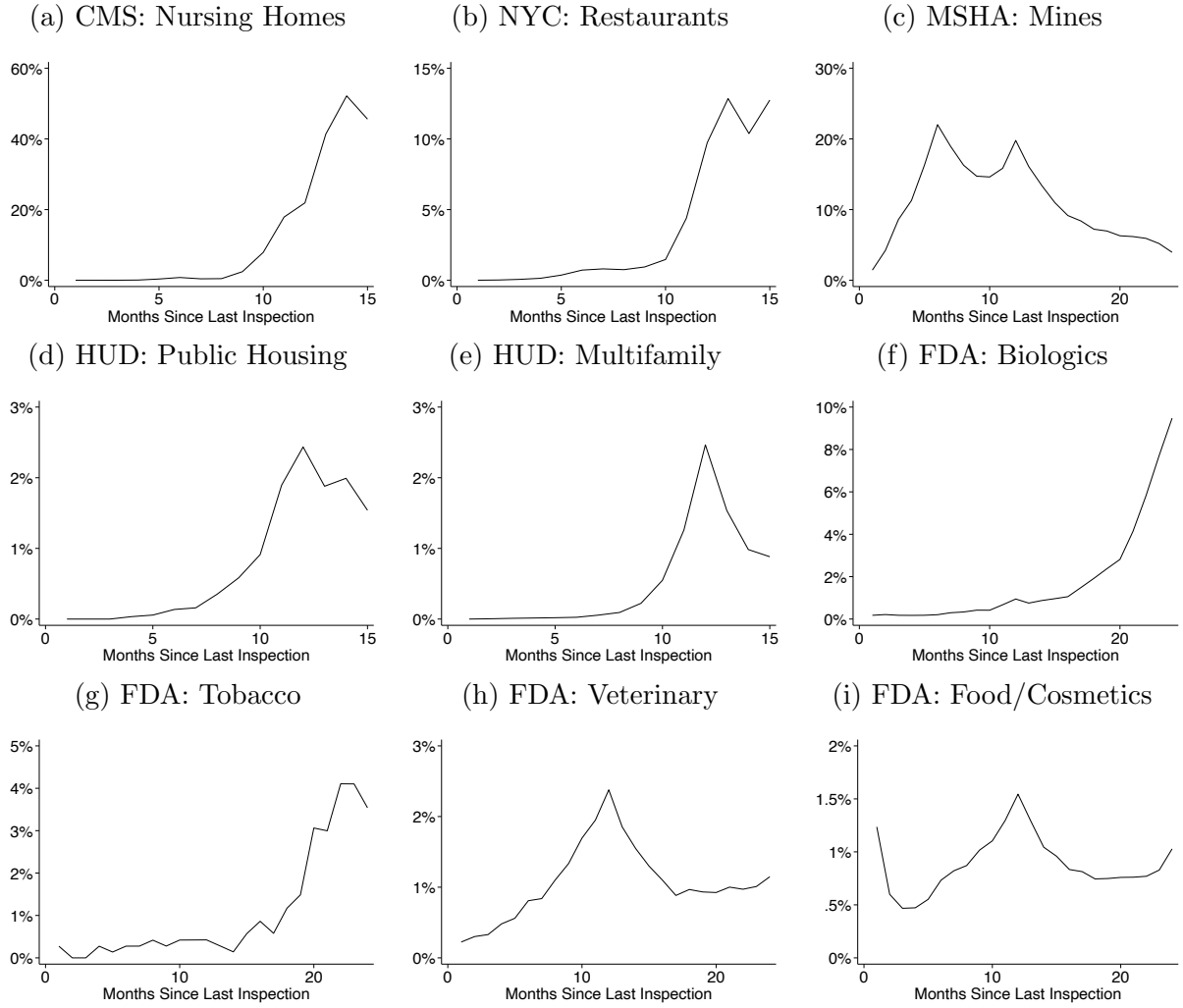
Inspections are a key tool for ensuring compliance from regulated entities or contracted agents across a variety of settings, including product quality, environmental protection, worker safety, and public health. Inspections can improve compliance both directly, by providing information to identify and correct violations, as well as indirectly, as even the threat of an inspection can disincentivize shirking. While the exact timing of inspections is often unannounced to avoid "window-dressing" to cover up poor compliance, many inspection regimes also require that inspections occur at regular frequency. This makes the timing of inspections highly cyclical and therefore *predictable* based on the time elapsed since the previous inspection. Figure 1 shows that this feature is common across inspection regimes in a variety of contexts. In each case, time elapsed since last inspection predicts whether another inspection is likely imminent.

The benefits of unpredictability in inspection timing depend crucially on the extent to which firm effort and compliance dynamically respond to inspection risk (Lazear, 2006; Varas et al., 2020; Ball and Knoepfle, 2023). This paper studies firm responses to dynamic inspection risk and the corresponding benefits of unpredictability in a high-stakes healthcare setting. Specifically, we examine the U.S. nursing home industry, where facilities are required by the Centers for Medicare and Medicaid Services (CMS) to be inspected approximately every year. We use rich, high-frequency administrative data to examine the dynamic relationship between inspection risk, nursing home staffing levels as a proxy for firm effort, and patient health outcomes. These patterns then inform a simple dynamic model that we use to assess how changing features of inspection regimes such as their predictability and frequency would affect overall nursing home effort and compliance.

Our empirical analysis centers on how inspection hazard, effort, and compliance vary throughout the inspection cycle, or the period between standard health inspections. In our primary specification, we examine how these outcomes evolve as a function of the time elapsed since the previous inspection. We first confirm that inspection timing is distinctly cyclical. Hazard rates are relatively low immediately after an inspection, but begin to rise sharply approximately 35 weeks into a new cycle, before "resetting" after an inspection occurs.

We then consider how nursing home effort evolves over the inspection cycle. Effort to improve compliance could involve a variety of endogenous responses that may be difficult to observe, like conducting on-the-job training, performing practice inspections, having more hands-on management, or enforcing safety standards more stringently. We use payroll-based staffing ratios—the number of care hours per resident-day—as a proxy for facility effort,

Figure 1: Inspection Hazard Rates Across Regulatory Settings



Notes: This figure plots inspection hazard rates across different regulatory settings. They correspond to (a) standard inspections of nursing homes conducted by state survey agencies on behalf of CMS (2013-2019), (b) restaurant food safety inspections conducted by the New York City Department of Health and Mental Hygiene (2021-2025), (c) mine safety inspections conducted by the Mine Safety and Health Administration (1978-2004), inspections of public housing properties (d) and multi-family assisted properties by the Department of Housing and Urban Development (2018-2021), and inspections of manufacturers of (f) biologics, (g) tobacco goods, (h) veterinary drugs, and (i) food and cosmetics (2009-2024).

under the assumption that staffing decisions are correlated with these harder-to-observe compliance behaviors. Furthermore, although staffing ratios may not capture every facet of a facility's operations, labor is *the* critical input in this industry (National Academies of Sciences Engineering and Medicine, 2022). Indeed, the empirical linkage between staffing levels and patient outcomes is so strong that staffing is itself used as a key quality measure by both regulators and researchers.[1] Hence, this measure not only provides a high-frequency and reliably-reported proxy for overall compliance effort, but also directly reflects the central role of labor levels in our context. Accordingly, throughout the paper we refer to "staffing" and "effort" interchangeably.

We find that once a facility is no longer subject to follow-up visits for violations from their previous inspection, nursing homes maintain low staffing levels as the likelihood of the subsequent inspection is low. However, as the threat of the next inspection rises, staffing levels improve accordingly. Our results indicate that staffing levels start to increase approximately 35 weeks after the last inspection – this is roughly 20 weeks before the next one, given an average inspection cycle length of 55 weeks. There is a steady ramp-up in staffing several months in advance of the next inspection, which suggests we are primarily capturing a response to gradually increasing inspection risk rather than last-minute "window-dressing" ahead of a known inspection date.

Our main measure of nursing home compliance is patient survival. We find that patient survival rates follow the same pattern as staffing: once the follow-up re-visits from the previous inspection are over, patient survival dips in the middle of the inspection cycle and then begins to rise again at 40 weeks since the last inspection. Patient mortality is about one percent lower 55 weeks from the inspection than it is in the mid-point of the inspection cycle at 30 weeks. As a result, there is a strong positive relationship between staffing and patient survival rates: a one standard-deviation increase in effort is associated with 3 percent lower patient mortality rate. We find similar patterns when examining secondary health outcomes, including measures of the activities of daily living, the inappropriate use of antipsychotics (a form of chemical restraint) and the use of any physical restraints.

We conjecture that the relationship between inspection hazard and patient health reflects the causal effect of nursing home compliance effort in high inspection risk weeks, but these patterns could also be driven by patient selection if nursing homes respond to inspection risk by admitting healthier patients and discharging sicker ones. Differentiating between a causal effect and a patient composition effect is key to informing the counterfactuals in our structural model. If the empirical relationship is primarily driven by differences in selective admission or discharge practices, then we would incorrectly conclude that raising the

[1] This perspective is why the rich payroll-based staffing data we use are made publicly available.

inspection probability would reduce mortality, when it is merely shifting patient composition.

We assess this patient composition channel in two ways. First, we examine whether observable patient risk varies across the inspection cycle. We make use of a wide variety of health and utilization measures about patients prior to their admission, which we aggregate into a single predicted survival index. While actual patient survival evolves across the inspection cycle, we find no movement in predicted survival, suggesting that facilities are not adjusting their patient composition based on observable risks. Furthermore, we ensure that our results on actual patient survival are robust to controlling for this measure of observable risk.

Second, we employ an instrumental variables approach to isolate the causal effect of inspection risk-induced effort on patient health. We construct a time-varying, distance-based instrument of nursing home effort by combining our staffing estimates with a patient's location and the timing of their admission. Intuitively, the instrument captures where in the inspection cycle the nursing homes closest to the patient's home zip code were during the patient's stay. For a given patient zip code and week, we use the coefficients from the staffing regression to construct expected effort in each facility, based on each facility's time since their last inspection. We then aggregate this across all nearby facilities, weighted by the admission probability based on the patient's home zip code. The resulting measure is agnostic to the *actual* facility a patient is admitted to and hence purged of the potential impact of patient selection. Using this measure, we show that being exposed to inspection-induced nursing home effort causally reduces patient mortality. The largest effects on patient survival are within the first 10 weeks since admission, the period in which mortality risk is the greatest.

Guided by our empirical evidence on the relationship between inspection risk, staffing levels, and patient outcomes, we develop a simple dynamic model that captures how a nursing home chooses its level of effort over time in anticipation of a potential inspection. In each period, the facility weighs the rising probability of an inspection against the costs of exerting higher effort. By linking these evolving incentives to the observed patterns in our data, we can estimate the model parameters in a transparent way. We then use the estimated model to perform counterfactual policy experiments, simulating how alternative inspection regimes—such as making inspections less predictable or more frequent—would shift nursing home effort and therefore alter patient health outcomes. This structural approach complements our reduced-form evidence by providing a clearer sense of the mechanisms driving cyclical patterns in staffing and patient mortality and quantifying the potential long-run gains from changing the design of inspections.

While an extensive empirical literature has demonstrated that increasing the frequency

(Olken, 2007; Levine et al., 2012; Blundell et al., 2020; Shi, 2024), targeting (Duflo et al., 2018; Johnson et al., 2023), and visibility (Jin and Leslie, 2003; Dranove and Jin, 2010; Johnson et al., 2023) of inspections can improve compliance, this paper is the first to empirically examine a key dimension of inspection policy design: their predictability. Recent theoretical work has characterized conditions under which random or predictable inspection timing is optimal, and shown that randomly-timed inspections are preferable when the principal's goal is to maintain compliance (Ball and Knoepfle, 2023) and to deter agent shirking (Varas et al., 2020).

The main challenge in studying inspection predictability empirically is that in most settings, day-to-day shirking and compliance is difficult to measure in the absence of an inspection.[2] We overcome this by focusing on the nursing home context, where firm effort and compliance are measured both at high frequency and objectively. The staffing data are derived from auditable payrolls and were designed specifically to ensure accurate reporting (Geng et al., 2019; Gandhi et al., 2021). Our main compliance measure, patient mortality, is not manipulable and reported by CMS rather than nursing homes themselves. These rich data thus capture the near-universe of nursing home patients and staff at any given point in time. This enables us to observe how facilities perform *even* when they are not currently under inspection, helping us overcome this limitation in the empirical literature on inspection design.

This paper also contributes to our understanding of inspections as a policy tool to improve healthcare quality. Regular facility inspections—also known as accreditation or recertification surveys—are widely used across of a variety of healthcare settings, including hospitals, home health agencies, dialysis facilities, clinical laboratories, and hospices (CMS, 2024). But despite their prevalence, the literature on how inspections shape provider behavior and care quality is relatively limited. Existing work has largely focused on responses in the weeks immediately surrounding inspections (Barnett et al., 2017; Geng et al., 2019). While we indeed confirm sharp changes in nursing home effort during the inspection week, potentially driven by short-run tip-offs or inspector presence, the main focus of our paper is the gradual ramp-up in compliance in the several months preceding an inspection, which we take as reflecting nursing home responses to increased inspection hazard. Leveraging this variation allows us to understand how quality of care fares under more or less predictable inspection regimes.

[2]One exception is Zou (2021), which uses satellite-based measures of air quality and finds evidence of non-compliance on unmonitored days under the Clean Air Act. Sites in that context face an inspection with certainty every 6 days and the shirking occurs precisely on days they are not monitored. In contrast, the nursing home inspection cycle spans over a year, inspection dates are not known with certainty far in advance, and we detect a ramp-up in nursing home compliance almost half a year before their next inspection, suggesting it is a response to the increased risk of inspection.

Our findings from this exercise contribute more generally to the literature on policies which aim to improve healthcare quality by subjecting providers to scrutiny, like denials, anti-fraud litigation, and prior authorization (Dillender, 2018; League, 2022; Macambira et al., 2022; Brot-Goldberg et al., 2023; Leder-Luis, 2023; Eliason et al., 2024). We show how changing the *timing* of when these tools are used can improve overall compliance, even when holding their frequency fixed. Finding ways to optimize how these tools are implemented is highly policy-relevant, given that they are costly for both regulators and providers (U.S. Senate Committee on Aging, 2023; Dunn et al., 2024).

Finally, our study also adds to the growing literature on the economics of the nursing home industry. This literature has primarily focused on quality of care (Grabowski et al., 2008; Einav et al., 2022; Cronin and Evans, 2022; Olenski and Sacher, 2024), staffing and its effects on patient health (Lin, 2014; Matsudaira, 2014; Friedrich and Hackmann, 2021; Ruffini, 2022; Furtado and Ortega, 2023; Grabowski et al., 2023), issues surrounding access-to-care (Ching et al., 2015; Gandhi, 2023; Olenski, 2023; Hackmann et al., 2024), ownership and organizational structure (Grabowski et al., 2013; Gandhi et al., 2020; Gupta et al., 2024; Gandhi and Olenski, 2024), and on reimbursement design (Grabowski, 2001; Hackmann, 2019; Gandhi et al., 2024). While a large literature has studied the public report card system used to provide information about nursing home quality (Grabowski and Town, 2011; Werner et al., 2012; Cheng, 2023; Konetzka et al., 2021), very little is known about the inspection process that generates a key input to these scores. We contribute by examining the inspection process, an important yet under-studied component of the regulatory environment.

2 Background

2.1 Nursing Home Industry

This paper considers staffing levels, patient outcomes, and inspection behavior in skilled nursing facilities, colloquially referred to as nursing homes. The nursing home industry is large — it accounts for approximately one percent of U.S. gross domestic product — and composed of a mix of for-profit and non-profit operators. Nursing homes provide a continuum of care that spans two broad categories of patients. First, they offer *post-acute*, short-stay rehabilitative services for individuals transitioning from a hospital stay, largely funded by Medicare. Second, they deliver *long-term* custodial care for residents who require daily assistance with activities of daily living (ADLs) over an extended period. Such long-term care is primarily financed privately or by Medicaid.

The hallmark issue facing the nursing home industry is pervasively low quality-of-care.

Patients routinely face harm during their stays. Avoidable falls, pressure ulcers ('bed sores'), and inappropriate chemical/physical restraints are such common occurrences that the government reports the relative frequency of such incidents as part of facilities' public report cards. Such harm owed to neglect and abuse is commonly attributed to inappropriately low levels of nurse and nurse aide staffing levels. Most facilities are under-staffed relative to expert-proposed benchmarks (Geng et al., 2019), which has been linked to a variety of adverse health outcomes, including death (Konetzka et al., 2008; Friedrich and Hackmann, 2021).

Consequently, providers are subject to extensive regulation at both the state and federal levels. Nearly all facilities are certified to provide care and receive public reimbursement by the Centers for Medicare & Medicaid Services (CMS). As part of this certification process, CMS mandates annual (or near-annual) inspections for facilities receiving Medicare or Medicaid reimbursement. These inspections are conducted by state agencies, who have some flexibility over how inspections are scheduled and carried out, but must follow federal guidelines in what to evaluate and how often.

Information on both nursing home inputs and outputs is uniquely rich, even for the health care sector. On the patient side, all patients, regardless of payer, must receive detailed clinical assessments at regular intervals (on admission, at discharge, and at least quarterly in between). These assessments record a wide array of health indicators and diagnoses, enabling researchers and regulators to measure patient outcomes comprehensively at any point in time. On the input side, facilities submit auditable payroll-based reports of staffing on a daily basis, providing a detailed view of how labor fluctuates over time.

2.2 Role of Inspections

Inspections (formally referred to as "surveys") are a primary regulatory tool used by federal and state authorities to monitor quality within nursing homes. State survey agencies conduct unannounced inspections—typically about once a year, but no more than every 15 months—to check each CMS-certified facility's compliance with these federal standards (U.S. Congress, 1987). In practice, facilities are inspected on average every 54 weeks and the majority of inspections happen in a fairly tight window around this point, as illustrated in Appendix Figure C.1 – 80.4 percent of inspections occur between 45 to 65 weeks since the last inspection. During an inspection, a team of state surveyors reviews care practices, interviews residents and staff, and observes conditions to identify any deficiencies in care or safety. Identified violations ("deficiencies") are documented and rated by scope and severity, from minor isolated issues to widespread harm. Appendix Figure C.2 shows that 86 per-

cent of inspections find a deficiency and the average deficiencies number of deficiencies per inspection is 7.6.

Inspection findings are tied to enforcement actions: facilities must correct all cited problems and face penalties for serious non-compliance. Common enforcement actions include mandated corrective action plans with follow-up visits, civil monetary fines, denial of Medicare/Medicaid payments for new admissions, and in extreme cases suspension or termination of the home's certification, which effectively shuts down the facility. In this way, inspections directly serve as both a monitoring mechanism and an enforcement lever to uphold care standards in nursing homes.

Inspection results are also made publicly available and are an important resource for patients and families evaluating nursing homes. CMS publishes each facility's inspection reports as well as a 5-star report card rating, which aggregates information on inspection results, staffing levels, and patient quality measures. When constructing this composite star measure, inspection results receive the greatest weight, owed to CMS's belief that they provide the strongest signal of quality. Survey results indicate that although patients are not fully aware of the availability of these data, those who are become responsive to negative inspection reports (Konetzka and Perraillon, 2016).

Nursing homes operators are keenly aware of the inspection process and are widely believed to adjust their behavior in anticipation of new inspections. Although annual inspections are unannounced by law, in practice they are fairly predictable, typically within a 9-15 month window. One patient advocacy organization notes "repeated anecdotal reports indicate that facilities...move staff and supplies into facilities in anticipation" of coming inspections (Center for Medicare Advocacy, 2023), a phenomenon also reported in popular news media (Thomas, 2014; Rau, 2018).

Indeed, we find suggestive evidence that facilities that appear to face earlier-than-anticipated inspection tend to have worse inspection outcomes. Inspection outcomes are characterized by deficiency scores, defined as the sum of all deficiencies identified during an inspection, weighted by their severity (CMS, 2008).[3] Thus, a higher deficiency score indicates worse performance during the inspection. Even after accounting for facility, state-year, and calendar week fixed effects, we find that deficiency scores decline as the number of weeks since a facility's last inspection increases. Similarly, deficiency scores are negatively associated with the difference in cycle length between the current and previous inspection cycles. That is, facilities that are inspected earlier in a given cycle and those that are inspected earlier

[3]This score is used to assign facilities to the Special Focus Facility (SFF) program, which subjects the worst-scoring facilities to more frequent inspections. We exclude observations during which a facility is participating in the SFF program for this analysis, though we consider these observations in our main analyses.

than they were in the prior cycle both tend to have worse outcomes (see Appendix Figure C.3). Taken together, these findings imply that early, unexpected inspections—when a facility has not yet had sufficient time to ramp up compliance efforts—lead to worse inspection outcomes.

3 Data and Sample Construction

3.1 Data

Nursing Home Inspections. Our primary source for inspection dates is the CMS Quality, Certification, and Oversights Reports (QCOR) website, which provides detailed data on federal and state inspections across various healthcare facilities. We scrape information on all surveys of nursing homes conducted between 2010 and 2019. For each survey, QCOR reports the survey end date; whether it was a standard initial or recertification survey or in response to patient complaints; whether it marked the first visit of an inspection cycle or a follow-up revisit to check on previously-identified deficiencies; the survey's focus (health or fire safety); and operational details, such as the number of surveyors present and average surveyor hours.

While nursing home inspection dates are also published directly by CMS (CMS, 2025b), we found that the QCOR data was more complete and included inspections that were not reported in the publicly-available CMS data. In supplementary analyses, we also use survey start dates acquired from a Freedom of Information Act request for inspections between 2015 and 2023. To study inspection outcomes, we supplement the QCOR data with public inspection deficiencies published by CMS (CMS, 2025a).

Staffing. We measure staffing using the Payroll-Based Journal (PBJ) data from 2017-2019, which provides daily administrative micro-data on direct care staff in all CMS-certified nursing homes. These data include information on hours worked, licensing type (registered nurse (RN), licensed practical nurse (LPN), or certified nursing assistant (CNA)), employee or contractor status, and a daily measure of the facility's patient census. The PBJ data are derived from facility payroll records which are automatically submitted to CMS; accordingly, these data are widely considered to more reliable and precise than self-reported staffing measures (Geng et al., 2019; Gandhi et al., 2021). Our main staffing measure is constructed using total weekly staffing hours per resident-day. For our primary results, we then standardize this measure at the facility-level using the mean and standard deviation for each facility in the middle of the inspection cycle (between 15 and 30 weeks from the last inspection). Ac-

cordingly, our primary staffing results can be interpreted as standard deviations away from the mean for each facility's "baseline" mid-cycle period, when they face minimal inspection risk.

Patient Characteristics and Health Outcomes. Our analysis of patient health outcomes relies primarily on the 2013-2019 Minimum Data Set (MDS), which captures regular assessments from all CMS-certified nursing homes. The MDS provides near-universal detailed clinical and demographic data on residents, assessed at admission, discharge, and regular intervals. These data are widely used by policymakers and researchers to evaluate care quality and inform reimbursement decisions. We use the MDS to construct a weekly panel of nursing home stays, allowing us to track patient outcomes over the inspection cycle. We augment our MDS panel data with patient characteristics and mortality derived from the 2013-2019 Medicare Master Beneficiary Summary File (MBSF) enrollment files, which includes patient demographics, chronic conditions, home zip codes, and death dates for all Medicare beneficiaries covered by traditional Medicare or Medicare Advantage in a given year. We also use the 2013-2019 Medicare Provider Analysis and Review (MedPAR) files, which contain information about hospital and nursing home stays covered by Medicare, to provide measures of from recent hospitalizations that inform our sample construction of nursing home stays.

Table 1: Summary Statistics

Facility Effort Sample: 2017-2019					
	Mean	Std. Dev	50%	25%	75%
Inspection Characteristics					
Weeks between inspections	54.8	9.9	55	49	60
Average deficiencies	9.04	9.46	6	2	14
Share with deficiency	79.8%	-	-	-	-
Staffing Characteristics					
Total hours	2236.0	1420.5	1955.7	1297.8	2808.7
Hours per resident-day	3.76	0.95	3.19	3.63	4.14
N Inspection cycles = 40,654	N Facilities = 15,692		N Facility-weeks = 2,255,867		
Patient Health Outcomes Sample: 2013-2019					
Inspection Characteristics					
Weeks between inspections	54.4	9.9	55	48	60
Average deficiencies	7.72	7.94	6	2	10
Share with deficiency	86.0%	-	-	-	-
Patient Characteristics					
Age	79.3	11.2	80	72	87
Predicted survival rate at admission	0.87	0.1	0.88	0.82	0.94
Share white	79.1%	-	-	-	-
Share female	61.1%	-	-	-	-
Patient Health Outcomes					
ADL score (negative)	-7.65	4.79	-8	-12	-4
4-week survival rate	95.9%	-	-	-	-
Share not on anti-psychotics	82.8%	-	-	-	-
Share without pressure ulcers	90.0%	-	-	-	-
Share without physical restraints	98.7%	-	-	-	-
Share without fall or fracture	96.8%	-	-	-	-
N Inspection cycles = 100,594	N Facilities = 16,428				
N Stays = 16,942,300	N Patient-weeks = 351,646,731				

Notes: Patients are limited to patients in MBSF with a valid predicted survival rate at admission. Antipsychotic use and ADL scores are carried forward between assessments, while all other health outcomes are reported conditional on being assessed.

3.2 Sample Construction

To identify facility inspection cycles, we restrict our QCOR sample of inspections to standard first visit health inspections and set this as the beginning of each inspection cycle. For our patient-level analyses, we restrict to nursing home residents who are enrolled in Medicare (i.e., those with corresponding enrollment records in the MBSF), who account for 94 percent of stays. A patient will be in our sample as long as they are a Medicare beneficiary, regardless

of whether any part of their stay was covered by Medicare.

In order to assign patients to nursing homes, we construct patient stay spells using MDS admission and discharge assessments. We identify a stay as starting when a patient is observed receiving an admission assessment and ending with a discharge assessment where return is not anticipated, when a patient dies, or if another nursing home reports an admission assessment for the patient. If a stay does not have a discharge assessment, death, or is interrupted by another admission, then we set the end date of the stay to be 150 days after the last observed assessment, following CMS practice (CMS, 2022).

4 Empirical Analysis

Our empirical analyses consist of several specifications, each relying on variation in inspection risk over time. In our primary specification, we examine outcomes throughout the inspection cycle – the period between standard health inspections. To rule out effects driven by endogenous patient selection across the inspection cycle, we also develop an instrumental variables approach which couples the inter-temporal variation in inspection risk with exogenous shifters to patient preferences over facilities. Finally, to ensure robustness to changes in sample composition in the later weeks of a cycle, we also estimate a stacked event study design centered around the inspection (discussed in Appendix A).

Broadly, our analyses span two sets of outcomes and time periods. Measures of staffing are aggregated to the facility-week level, whereas all health outcomes are measured at the patient-week level. The staffing measures – restricted by the availability of the payroll data – are defined only for calendar years 2017-2019. However, we observe patient-level outcomes extending back to 2013, and so all analyses of health outcomes are conducted over the window 2013-2019.[4]

4.1 Inspection Cycle Approach

Design. Our primary empirical design includes indicator terms for the number of weeks that have elapsed since the beginning of the current inspection cycle (i.e., the "weeks since last inspection"). This approach is summarized by equation (1), which models an outcome of interest for facility j in calendar week t:

$$Y_{jt} = \sum_{w \in [0,30)}^{(30,60]} \beta_w \, \mathbb{1}(\text{Week Since Inspection}_{jt} = w) + \alpha_j + \lambda_t + \varepsilon_{jt}, \tag{1}$$

[4]Our patient-level results are robust to choice of time period; restricting this sample to 2017-2019 to overlap with the PBJ data generates highly similar estimates.

where $\mathbb{1}(\text{Week Since Inspection}_{jt} = w)$ is an indicator for whether w weeks have elapsed since facility j's most recent inspection as of calendar week t. To account for time-invariant differences in outcomes across providers, we also include facility fixed effects α_j. To account for aggregate fluctuations across time, we include calendar-week fixed effects λ_t. Y_{jt} is a facility-level outcome, such as standardized hours per resident day. Because the omitted week from inspection is $w = 30$, the parameters of interest – the β_w coefficients – can be interpreted as the difference in outcome relative to the middle of an inspection cycle, after accounting for baseline differences across facilities and across time.

In addition, because many of our outcomes are measured at the patient-level, we consider an individual-level analogue given by equation (2):

$$Y_{it} = \sum_{w \in [0,30)}^{(30,60]} \beta_w \, \mathbb{1}(\text{Week Since Inspection}_{j(i),t} = w) + \gamma X_i + \alpha_{j(i)} + \lambda_t + \varepsilon_{it}, \tag{2}$$

where Y_{it} is a patient-week level measure, such as 4-week survival and $j(i)$ denotes the facility of patient i.[5] The only substantive modification from equation (1) is the inclusion of patient-level controls X_i. A wide array of clinical information is measured for nursing home patients. In the interest of parsimony, we aggregate this information into one measure of predicted mortality using demographics, utilization, and diagnoses prior to admission. Specifically, we use a high-dimensional fixed effect regression to predict 90-day mortality after admission using age, race, sex, chronic conditions noted in the initial admission assessment, pre-admission hospital stay diagnosis group (DRG) and length of stay, initial payer, and whether the patient qualifies for end stage renal dialysis coverage. Appendix Figure C.4 plots a binscatter between predicted and actual 90-day mortality; the relationship is monotonic and has an R^2 of 0.09.

For all our analyses, we compute two-way cluster-robust standard errors by facility and by calendar week. This choice accounts for two sources of error correlation. First, there may be serial correlation within a facility over time (e.g., persistent facility-level factors). Second, there may be correlated shocks across different facilities in the same calendar week (e.g., seasonality). This second dimension is particularly relevant if inspection cycles are not uniformly distributed over the calendar year, so that many facilities enter similar phases of their inspection cycle at the same point in calendar time.

Results. Figure 2 documents the results of the inspection cycle analysis on our two primary outcomes: nursing home effort as proxied by staffing (panel a) and patient survival (panel b). Panel (a) plots the staffing coefficients by weeks since last inspection as well as the inspection

[5] Formally, i denotes a patient stay, as a few patients are observed across multiple stays.

hazard, defined as the probability that a nursing home receives their next standard health inspection conditional on having not received one until that point in the current cycle. The inspection hazard is relatively flat up until week 40 and then subsequently rises rapidly, reflecting the predictability of the inspection cycle.

Nursing home staffing displays a spike in the week of inspection, which is consistent with related work investigating staffing during the narrow window surrounding inspection dates (Geng et al., 2019). Once the initial inspection is over, staffing follows a "U-shaped" pattern: it remains elevated but declining for about 20 weeks after the inspection. This aligns with the period of the inspection cycle in which nursing homes are subject to inspector revisits to follow up on any deficiencies identified in the initial inspection. Staffing then remains flat during the period of minimal new inspection risk. Eventually, as the inspection hazard increases and the next inspection becomes imminent, staffing levels gradually rise again, until a new inspection occurs and the cycle repeats itself. This pattern is also observed with two alternative measures of staffing: non-standardized hours per resident-day and log total hours (see Appendix Figure C.6), suggesting that these estimates are robust to our choice of staffing measurement.

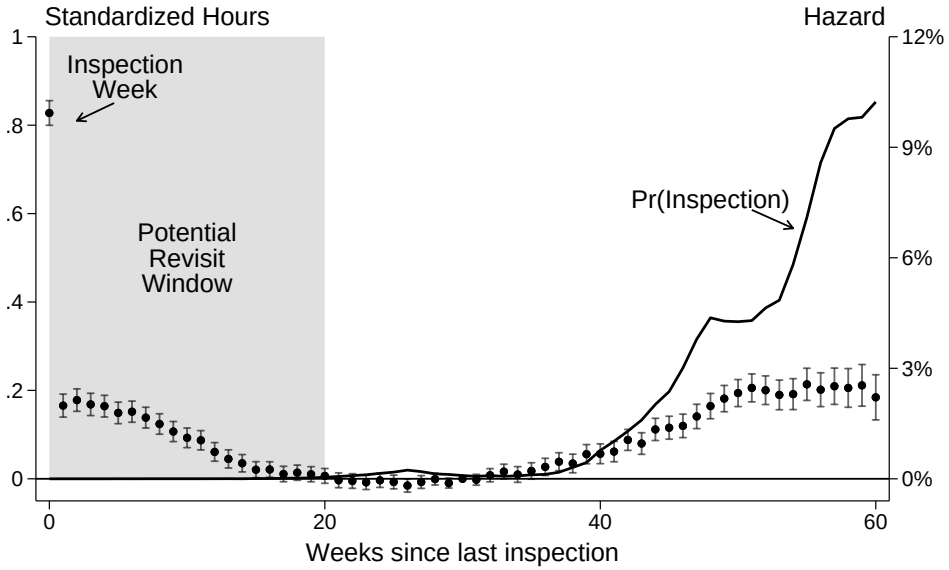
Panel (b) plots the estimates from our patient-level specification using 4-week survival as the dependent variable. The pattern of the coefficients also follows a "U-shape" similar to the staffing coefficients: survival is elevated in the weeks immediately following an inspection, dips in the middle of the cycle, and then rises again as the next inspection approaches. To identify which patients drive this pattern, we split all residents into "short-stay" (defined by CMS as those within 90 days of admission, whose care is primarily rehabilitative) and "long-stay" (the remainder, whose care is primarily custodial), and re-examine survival rates within in each group. Short-stay patients drive the overall survival effects, and we find only minimal changes in survival across the inspection cycle for long-stay patients (see Appendix Figure C.7). This heterogeneity is consistent with differences in baseline risk: 4-week mortality is highest in the early weeks of a stay and gradually declines thereafter (Appendix Figure C.8).

Of course, there are many operational routines that may change across the inspection cycle which may improve patient survival. Facilities may retain more qualified or experienced staff, the staff that are present may adhere more strictly to protocols, management practices may improve, and so on. Changes in these routines — which we conceive of as a nursing home's "effort" — are likely to be well-approximated by changes in a nursing home's overall staffing level, given how crucial this input choice is for the industry (see Section 2). Accordingly, for the remainder of the paper, we will use the terms "effort" and "staffing" interchangeably.

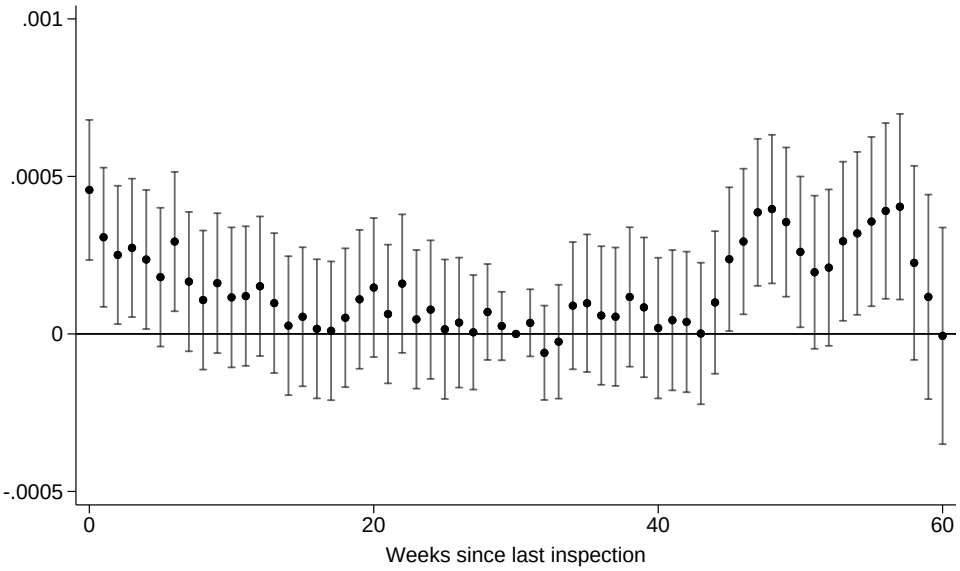
Figure 3 plots the survival coefficients against the effort coefficients for each point in

Figure 2: Inspection Hazard, Effort, and Patient Survival by Weeks Since Last Inspection

(a) Inspection Hazard and Coefficients on Hours Per Resident-Day (Std. Dev.)



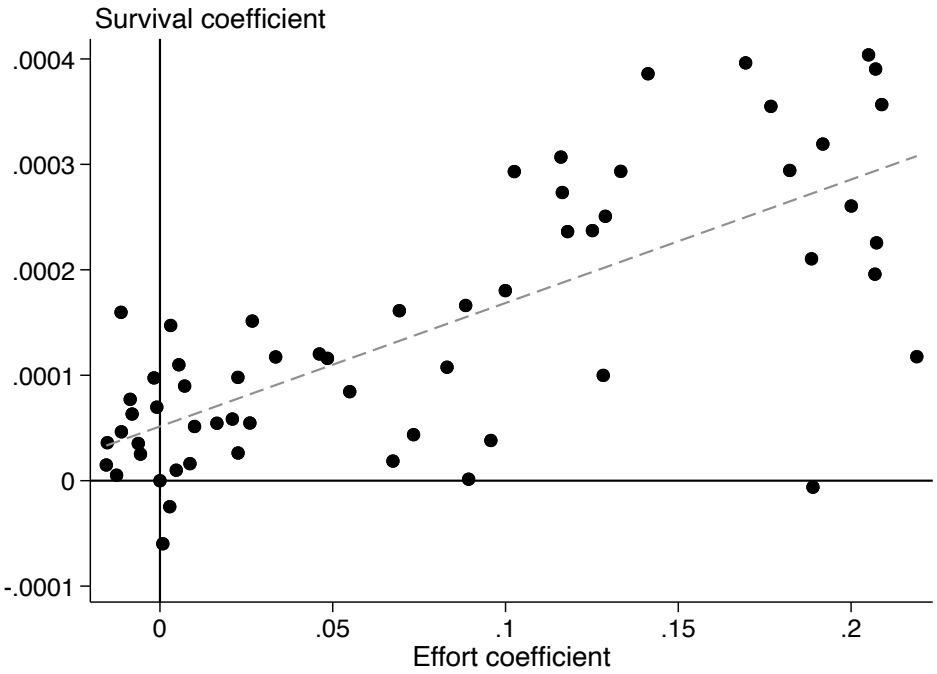
(b) Coefficients on 4-Week Survival Rate



Notes: Panel (a) plots the inspection hazards (right axis) and the coefficients on from estimating equation (1) on standardized staffing hours per resident-day (left axis). Panel (b) plots the coefficients from estimating equation (2) on 4-week patient survival rate, controlling for predicted patient mortality at admission.

the inspection cycle (i.e., each week from the last inspection), as well as the line of best fit between the two. There is a robust positive relationship between the two — survival is higher in weeks when effort is higher. The slope is 0.0012 (SE: 0.00014), indicating that a 1 standard deviation increase in staffing is associated with a 3.2 percent lower patient mortality rate. This relationship is robust to restricting only to coefficients corresponding to inspection cycle weeks greater than 30, i.e. when we exclude the revisit window (Appendix Figure C.5).

Figure 3: Comparison of Inspection Hazard to Effort vs. Survival Coefficients



Notes: This figure plots the effort coefficients and survival coefficients from estimating equations (1) and (2) in Figure 2. The slope of the line of best fit is 0.00117 (SE: 0.00014).

Due to the richness of the MDS data, we are also able to examine a wide variety of "intermediate" (i.e., sub-mortality) patient health outcomes. Figure 4 plots the corresponding estimates using several clinical measures derived from the assessments data. Panel (a) plots the activities of daily living (ADL) score, a numeric measure of a resident's independence in performing tasks related to personal care, including transfer, bed mobility, toileting and eating. Each activity is assessed from 0 to 4, with 0 meaning the resident completed each activity independently while 4 indicates total dependence on staff for each of the activities, and the activity scores are summed. Because a higher score indicates worse health, we negate it. The negated ADL score displays a similar pattern as effort and survival — it is lowest in the middle of the inspection cycle, when nursing homes are subject to the least oversight.

Panel (b) shows the share of patients not on antipsychotics and panel (c) plots the

share of patients without physical restraints, two common measures of nursing home quality that are used by CMS to construct public report cards, scaled so that larger values indicate better outcomes. Inappropriate antipsychotic prescribing (often used to sedate patients with Alzheimer's disease or a related dementia) and physical restraints are widely considered indicators of low-quality care, given the associated reductions in residents' quality-of-life, Curbing their use through public reporting has long been a policy goal for CMS. We find that both of these indicators also display a "U-shaped" pattern throughout the inspection cycle. These patterns likely explain the counterintuitive inverse "U-shape" we find when examining the share of patients without a fall or fracture in panel (d), as patients who are restrained are mechanically less likely to fall. Finally, we find no discernible effects on the share of patients without pressure ulcers (panel e), another common nursing home quality measure.

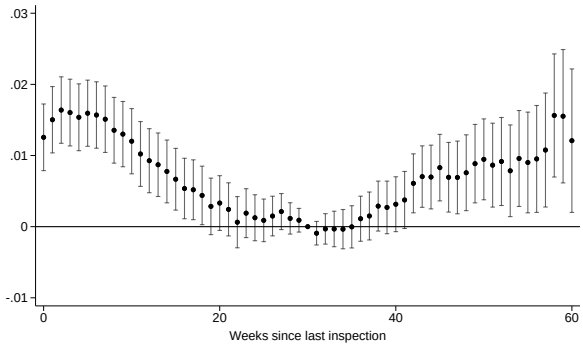
The positive relationship between the survival and effort coefficients in Figure 3 indicates that patient survival is higher at the points in the inspection cycle when nursing homes exert greater effort. It is possible that the improvements in observed patient health instead reflects another endogenous response of facilities to changes in inspection risk, in the form of selective admissions and discharge practices, such as in Gandhi (2023) and Hackmann et al. (2024), rather than resulting from increased facility effort. We engage in several exercises to assess the plausibility that changes in patient selection explain the patterns we observe.

First, we examine how a facility's overall occupancy, admission, and discharge decisions change over the inspection cycle. We find evidence of only very modest fluctuations in the number of patients in a facility (Appendix Figure C.9). From peak to trough, the typical facility faces differences of at most 0.016 standard deviations in the number of patients in residence at any time, corresponding to only 0.8 fewer patients relative to the baseline average of 82.3 residents a week. This change is driven by a very minor reduction in new admissions during inspection weeks and a corresponding reduction in new discharges in the weeks following an inspection (generally on the order of less than 0.004 standard deviations per week). In total, we find evidence that facility occupancy fluctuates only minimally across the inspection cycle.

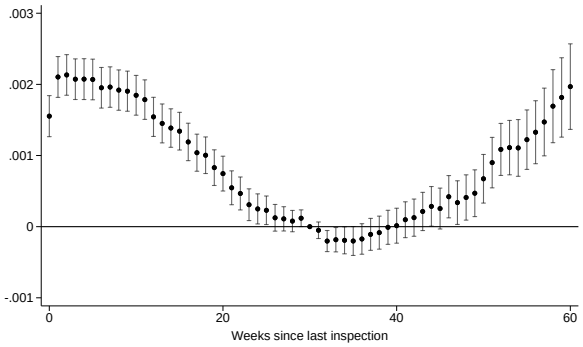
To understand whether these minor changes in admissions and discharge practices change patient composition in the facility (i.e., whether the marginal patient is healthier or sicker than the average patient), we rely on our univariate measure of observable patient health: predicted survival X_i from equation (2). Specifically, we examine how this index evolves across the inspection cycle, both overall as well as specifically for new admissions and discharges each week. Crucially, this index relies only on pre-admission information that is observable to the facility making admission or discharge decisions. We find no detectable

Figure 4: Additional Health Outcomes by Weeks Since Last Inspection

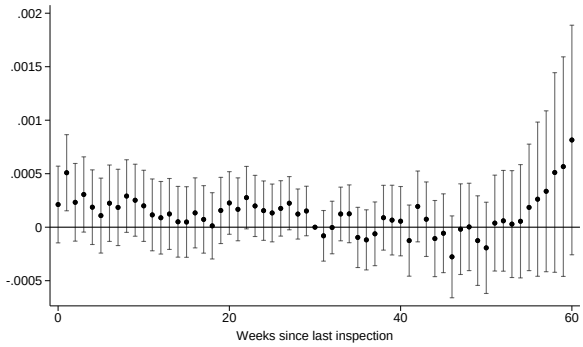
(a) (Negative) Activities of Daily Living Score



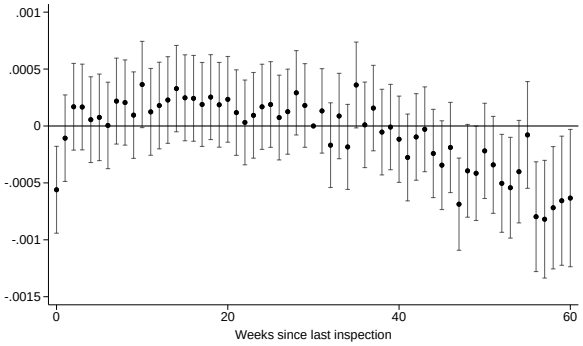
(b) Share of Patients Not on Anti-Psychotics



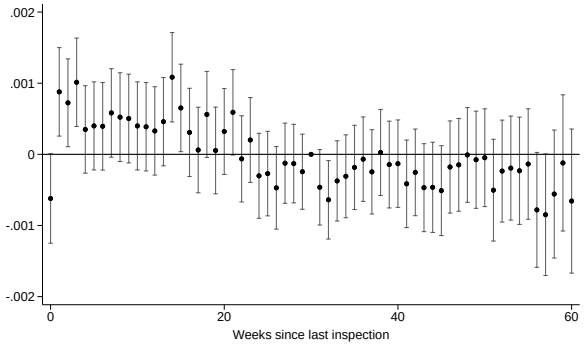
(c) Share of Patients Without Physical Restraints



(d) Share of Patients Without Fall or Fracture



(e) Share of Patients Without Pressure Ulcers



Notes: This figure plots the coefficients from estimating equation (2) on health outcomes measured in the MDS for patients in the facility in a given inspection week. All outcomes are normalized so that a higher value is associated with a better outcome. Panel (a) plots the negated activities of daily living score, panel (b) plots the share of patients not on anti-psychotics, panel (c) plots the share of patients without physical restraints, panel (d) plots the share of patients without a recent fall or fracture, and panel (e) plots the share of patients without pressure ulcers.

changes in predicted survival throughout the inspection cycle. We can rule out changes in predicted survival relative to baseline larger than 0.02 percent among all patients in the nursing home, 0.12 percent among admissions, and 0.12 percent among discharges (Appendix Figure C.10).

Moreover, consistent with these results, we find that excluding the predicted mortality control X_i generates estimates that are highly similar to our main specification (Appendix Figure C.11). Taken together, these results indicate that changing patient composition is unlikely to be driving the patient health effects observed in Figures 2 and 4. To further assess the robustness of our results to shifts in patient composition in Section 4.2 we develop an instrumental variables approach that is unaffected by endogenous patient selection.

Alternative Event Study Approach. In addition to changes in patient composition across the inspection cycle, a secondary concern could be that *facility* composition may change across the inspection cycle. That is, for a facility that is consistently inspected every 40 weeks, that facility's will observations never contribute to estimates of β_w for $w > 40$. The nursing fixed effects α_j should capture any time invariant cross-facility differences; nevertheless, in Appendix A, we explore how an alternative stacked event study research design would impact our results. Reassuringly, we recover estimates that are very similar to our primary inspection-cycle specification.

4.2 Instrumental Variables Approach

The estimates resulting from the prior specification illustrate how our outcomes of interest, nursing home staffing and patient outcomes, evolve across the inspection cycle. We can compare the estimates on effort and patient outcomes for the same week to characterize how the two co-vary. However, as noted, this positive relationship need not be causal, but could instead reflect efforts to assemble a healthier patient pool through selective admissions or discharge.

In addition to the predicted survival analyses discussed in Section 4.1, we further address this concern with an alternative instrumental variables approach. In particular, we construct a time-varying, distance-based instrument for the amount of effort a patient is *expected* to be exposed to, irrespective of the nursing home they are actually admitted to. Instrumented effort is instead based on geographic variation in where a patient lives as well as temporal variation in where nursing homes in the patient's vicinity are in their inspection cycle at the time of the patient's nursing home stay.

This approach expands on the standard distance-based instrument used to address selective admissions in healthcare (McClellan et al., 1994; Einav et al., 2022; Olenski and Sacher,

2024) by also incorporating temporal variation in the inspection cycle *within* a zip code. In other words, because our instrument varies over time depending on where nursing homes are in their inspection cycle, we can include patient home zip code fixed effects. Our comparisons are thus among patients in the same zip code who are exposed to different levels of nursing home effort based on the timing of their admission, relative to the inspection cycles of nearby nursing homes.

Approach. To identify which facilities patients are most likely to be admitted, we estimate a simple discrete choice logit model of patients' nursing home choices, considering only the distance between their home zip code centroid and each nursing home. We assume that the utility of patient i with home zip code h from choosing some nursing home j is given by:

$$u_{ij} = \mu \log d_{h(i)j} + \epsilon_{ij}, \tag{3}$$

where $\log d_{h(i)j}$ is the log distance between the centroid of patient i's zip and nursing home j. Assuming the error term ϵ_{ij} follows the type-I extreme value distribution, this specification implies that the probability patient i living in zip code z chooses nursing home j is:

$$p_{ij} = \frac{\exp(\mu \log d_{h(i)j})}{\sum_{k \in C_{h(i)}} \exp(\mu \log d_{h(i)k})}, \tag{4}$$

where $C_{h(i)}$ is the set of nursing homes available to someone with home zip code h.[6] We estimate μ using a high-dimensional Poisson model, given the equivalence of its likelihood function to a conditional logit (Guimarães et al., 2003; Correia et al., 2020; Brot-Goldberg et al., 2023).

For a given nursing home j and calendar week t, we observe the number of weeks w that have elapsed since j's last inspection. Using our regression estimates of $\widehat{\beta}_w$, $\widehat{\alpha}_j$, and $\widehat{\lambda}_t$ from equation (1), we can then compute the level of effort we expect facility j to expend w weeks after its last inspection. Specifically, we can compute each facility j's predicted level of effort for each possible w as:

$$\widehat{e}_{jw} \equiv \sum_{w \in [0,30)}^{(30,60]} \widehat{\beta}_w \; \mathbb{1}(\text{Week From Inspection}_{jt} = w) + \widehat{\alpha}_j + \frac{1}{N_w} \sum_t N_{wt} \widehat{\lambda}_t, \tag{5}$$

where N_w denotes the number of observations spent w weeks from the last inspection and N_{wt} denotes the number of observations in period t that are w weeks from last inspection.

[6] For computational tractability, we define patients' choice sets as those facilities within 25 kilometers of their zip code's centroid.

This last term therefore denotes the mean calendar week t fixed effect for facilities observed w weeks from their most recent inspection. Putting this together, $\widehat{e}_{jw}$ is the average predicted effort for a nursing home j that is w weeks from the last inspection.

Using these estimates of predicted effort, we can compute the instrumented effort we expect each patient i to receive in each period t, *irrespective* of which facility they are actually admitted to. Relying on patients' home zip codes (i.e., their pre-admission zip codes), we can construct instrumented efforts as:

$$\underbrace{\widehat{z}_{it}}_{\substack{\text{Instrumented}\\ \text{Effort}}} \equiv \sum_{j \in C_{h(i)}} \underbrace{\widehat{p}_{ij}}_{\substack{\text{Distance-Based}\\ \text{Demand}}} \underbrace{\widehat{e}_{j,w(j,t)}}_{\substack{\text{Predicted Effort from}\\ \text{Weeks from Inspection}}} \tag{6}$$

where $\widehat{e}_{j,w(j,t)}$ is the predicted effort of facility j in calendar week t, based on where in the inspection cycle that facility is at that point in time (i.e., $w(j,t)$). Instrumented effort $\widehat{z}_{it}$ can be interpreted as the effort patient i is expected to be exposed to in each calendar week based *only* on their home zip code.

In Figure 5 panel (a) we compare this "instrumented effort" to the "realized effort" which we denote (with a slight abuse of notation) as $\widehat{e}_{jt}$, which is the (predicted) effort at the nursing home they were actually admitted to. There is a strong positive relationship between the two, with a coefficient of 0.96 and an R^2 of 0.2. Panel (b) then plots the distributions of instrumented effort and its residual after absorbing zip fixed effects. The residualized value captures within-zip variation in effort exposure driven by the timing of when patients are admitted relative to the inspection cycle. Reassuringly, there is still substantial variation in our instrument even after accounting for zip fixed effects that is driven purely by the inspection cycle.

Finally, we use our instrumented effort measure in the following specification to estimate the causal relationship between nursing home effort and patient health outcome Y_{it}:

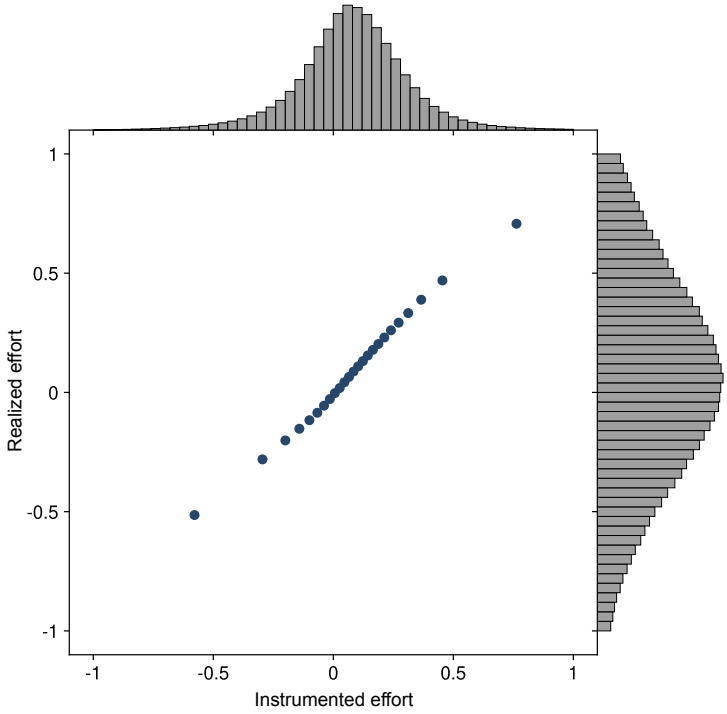
$$Y_{it} = \sum_{\tau=0}^{60} \delta_\tau \widehat{r}_{it} \times \mathbb{1}(\text{Week From Admission} = \tau) + \gamma X_i + \alpha_{h(i)} + \lambda_t + \eta_\tau + \varepsilon_{it}, \tag{7}$$

where τ denotes the week of patient i's stay relative to their admission week. As before, we control for predicted patient mortality, X_i. Because our instrument varies within zip code over time, we include home zip code fixed effects $\alpha_{h(i)}$. We stratify our results by week from admission to account for the fact that patient mortality varies dramatically depending on the week from admission, in addition to our prior results that short-stay patients are those most impacted by the inspection cycle.

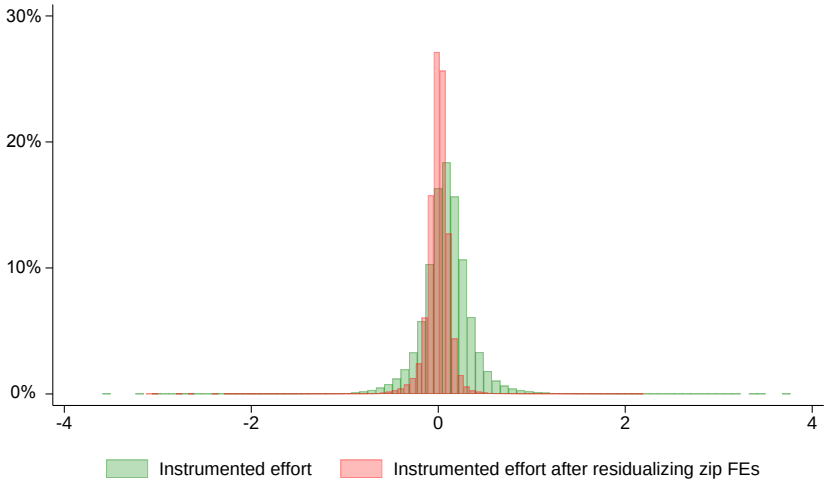
The δ_τ coefficients can be interpreted as the effect of a one-standard deviation increase

Figure 5: Distance Instrument

(a) Realized vs. Instrumented Effort



(b) Distribution of Instrumented Effort

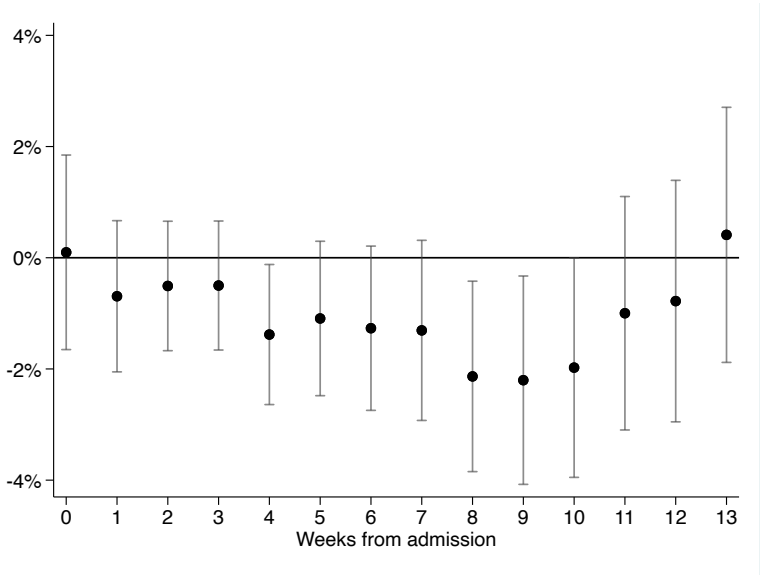


Notes: Panel (a) plots a scatter plot between instrumented effort $\widehat{z}_{it}$ and realized effort $\widehat{e}_{j(i)t}$. Instrumented effort is defined as equation (6): it is the average of realized effort at nursing homes within 25 kilometers of a patient's zip, weighted by the probability that the patient goes to that nursing home. Realized effort is the predicted effort defined in equation 5 at nursing home j, the nursing home the patient was actually admitted to. The predicted effort is calculated using the coefficients from estimating equation (1) on staffing. Panel (b) plots a histogram of raw instrumented effort as well as after residualizing for patient zip code fixed effects.

in inspection-induced nursing home effort on Y_{it} for patients who are τ weeks away from their admission to the nursing home. In order to interpret these effects as the causal effect of nursing home inspection risk, we require the following exclusion restriction assumption: after accounting for zip code and time fixed effects, there is no correlation between patient mortality and the effort exerted by geographically-proximate nursing homes other than through the timing of the nursing home inspection cycle.

Results. Figure 6 plots the δ_τ coefficients from estimating equation (7) on 4-week mortality for patients in the first 13 weeks of their nursing home stay. We restrict our focus to the first 13 weeks, which is equivalent to approximately the first 90 days of a patient's stay, since mortality is highest in these weeks (Appendix Figure C.8) and our results from the inspection-cycle analysis are concentrated among these "short-stay" patients (Appendix Figure C.7). The coefficients are normalized in percent terms relative to the average 4-week mortality rate in each week from admission; Appendix Figure C.13 shows the unscaled coefficients, which are in percentage point terms.

Figure 6: Distance IV Mortality Coefficients, Normalized by Mean



Notes: This figure plots the coefficients from the instrumented distance specification in equation (7) where the outcome variable is 4-week patient mortality, normalized in percent terms relative to average mortality for each week of admission. The regression is estimated on weeks 0-59 from admission and the plot shows the estimates on the first 13 weeks (equivalent to approximately the first 90 days) of the stay. All coefficients past week 13 are statistically insigificant at the 5 percent level.

The results show that exposure to inspection-induced nursing home effort reduces patient mortality. The protective effects on patient health are concentrated among patients who are

4 to 10 weeks out from their admission week. The lack of an effect in the first few weeks after admission could be because the effect of nursing home effort on health is cumulative, so patients may need several weeks of exposure before it has a demonstrable impact. Beyond 10 weeks from admission, mortality begins to wane, so the lack of an effect there could reflect the fact that there are fewer marginal deaths in this period.

5 A Dynamic Model of Nursing Home Effort

We next develop a simple dynamic model to formalize the main empirical patterns in our results and to quantify the influence of inspection timing on a nursing home's staffing decisions. Our aim is to capture a nursing home's incentive to raise effort — as proxied for by weekly staffing levels — in anticipation of an inspection, while acknowledging that staffing is costly and cannot be adjusted instantaneously without friction. The model will serve two primary purposes. First, it provides a theoretical framework that clarifies how a facility balances present costs against the potential future benefit of improved performance on inspections. Second, it gives us a vehicle for performing counterfactual policy experiments, such as altering the frequency or predictability of inspections, to assess how such changes might shift behavior and patient outcomes.

5.1 Model and Estimation

Setup and Timing. We treat the inspection cycle as our core unit of analysis and index time in weeks since the most recent inspection: $t = 1, 2, \ldots, T$, where T is the maximum plausible length of an inspection cycle. Federal statute dictates that inspections should not occur more than 15 months apart (i.e., $T \approx 65$); however, in practice, inspector shortages result in a non-trivial number cycles lasting longer than 65 weeks. To account for this, we allow for a greater maximum cycle length of $T = 100$. Each week t of the inspection cycle, the facility chooses how much costly effort e_t to exert. This effort encompasses the many costly actions a facility can take to improve its care processes. We observe one such action—staffing levels—at high frequency in our data; however, there are many more unobserved ways in which facilities can exert effort, such as providing on-the-job training, enforcing policies and standards of care more strictly, and performing practice inspections.

Unannounced inspections occur stochastically at the end of each period t with probability $h(t)$ that varies over the cycle. The facility's performance on the inspection depends on the effort is has been expending. Specifically, we model a facility's effort each week as contributing to a running "stock" s_t that depreciates over time. This effort stock then

determines the payoff $\pi(s_t)$ the facility receives when inspected. After inspection, the facility starts a new cycle: t returns to period 1 and the facility's effort stock is reset to a baseline level s_0.[7] We provide additional details on effort accumulation, effort costs, inspection payoffs, and the resulting Bellman equation below.

Accumulation of Effort. At the start of each period t, the facility holds a stock of effort $s_{t-1} \in S$, which summarizes its effort intensity through period $t-1$. The facility's chosen effort in period t updates their stock of effort in the following way:

$$\begin{aligned} s_t &= (1-\delta)e_t + \delta s_{t-1} \\ &= (1-\delta)e_t + (1-\delta)\delta e_{t-1} + \ldots + \delta^t s_0 \\ &= (1-\delta)\sum_{\tau=1}^{t} \delta^{t-\tau} e_\tau + \delta^t s_0, \end{aligned} \tag{8}$$

where $\delta \in [0,1]$ measures the extent to which effort stock—i.e., benefits from past effort—are persistent. A larger δ implies that effort in previous periods benefit current performance more. Two key interpretations stem from equation (8). The first is that s_t is the depreciation-discounted average of previous efforts (and the initial stock). The second is that since s_t is a convex combination of s_{t-1} and e_t, one can intuitively think of the choice of e_t relative to s_{t-1} as adjusting the stock of effort. We incorporate this intuition into our effort costs below.

Cost of Effort. In each period t, choosing a higher level of effort e_t incurs a cost that has two components. The first component captures the linear cost of scaling effort e_t. This could include, for example, employing additional hours from existing nursing and administrative staff at their current wage and benefits. The second is an "adjustment" cost if the facility chooses an effort that deviates too greatly from its recent efforts. For example, dramatically increasing staffing over a short period of time will likely require non-linear costs such as recruiting new staff, raising wages and benefits to improve retention, and paying overtime to increase hours from full-time workers. We capture this with a quadratic term in $e_t - s_{t-1}$, where stock s_{t-1} is a weighted average of prior efforts. Thus, our cost function is:

$$c(e_t, s_{t-1}) = \alpha_1 e_t + \alpha_2 (e_t - s_{t-1})^2, \tag{9}$$

where $\alpha_1 > 0$ and $\alpha_2 > 0$ respectively capture the linear and adjustment costs of effort.

[7] Our model is not constructed to capture behavior during the "revisit" window, and data on revisits are sparse. As such, we treat week 21, the first period after the 20 week revisit window, as period $t = 1$.

Inspection Payoffs. We let $h(t)$ denote the hazard rate of an inspection, or the probability that an inspection occurs at the end of period t. If an inspection occurs, the firm receives a payoff that depends on s_t, its accumulated effort at the end of the period:

$$\pi(s_t) = \gamma s_t, \tag{10}$$

where $\gamma > 0$ represents the benefit of better performance on an inspection.

Dynamic Optimization. The facility's problem is to choose a discrete effort level $e_t \in E$ each period to maximize the expected discounted payoffs over the entire cycle. Let $V_t(s_{t-1})$ be the facility's value of having stock s_{t-1} at the start of period t. After selecting e_t, the facility pays cost $c(e_t, s_{t-1})$, and its effort stock updates to s_t. Then, an inspection occurs with probability $h(t)$. If it does occur, the facility earns $\pi(s_t)$ for its performance, the period restarts at $t = 0$, and the facility's effort stock resets to s_0. If the inspection doesn't occur, the period increases to week $t + 1$. The corresponding Bellman equation is:

$$V_t(s_{t-1}) = \max_{e_t \in E} \left\{ -c(e_t, s_{t-1}) + h(t)\Big(\pi(s_t) + \beta V_0(s_0)\Big) + \big(1 - h(t)\big)\,\beta V_{t+1}(s_t) + \varepsilon_{e_t,t} \right\} \tag{11}$$

where $\beta < 1$ is the discount factor.[8] We also allow for an i.i.d. Type-I extreme value shock $\varepsilon_{e_t,t}$ on the facility's decision in each period to account for random factors that affect the firm's decision-making not captured by the deterministic part of the model.

Denote by $\tilde{V}_t(e_t, s_{t-1})$ the deterministic component of the period t return from choosing effort e_t.[9] Then the choice of effort e_t is multinomial logit with the usual inclusive value:[10]

$$P(e_t \mid s_{t-1}) = \frac{\exp\Big(\tilde{V}_t(e_t, s_{t-1})\Big)}{\sum_{e \in E} \exp\Big(\tilde{V}_t(e, s_{t-1})\Big)} \tag{12}$$

$$V_t(s_{t-1}) = \log\left[\sum_{e \in E} \exp\Big(\tilde{V}_t(e, s_{t-1})\Big)\right]. \tag{13}$$

Given these, it is straightforward to solve equation (11) via backwards induction. Terminal V_T is computed from equation (13) with $h(T) = 1$. V_t for $t < T$ are then computed

[8] In practice, because our periods are weekly, we impose a high discount factor of $\beta = 0.99$.

[9] Formally:

$$\tilde{V}_t(e_t, s_{t-1}) = -c(e_t, s_{t-1}) + h(t)\Big[\pi(s_t) + \beta V_0(s_0)\Big] + \big[1 - h(t)\big]\,\beta V_{t+1}(s_t).$$

[10] For ease of notation, we exclude the Euler-Mascheroni Constant from the inclusive value.

iteratively from $t = T - 1$ back to $t = 1$.

Estimation. We estimate parameters $\theta = (\alpha_1, \alpha_2, \gamma, \delta, s_0)$ via maximum likelihood on the observed effort choices. The empirical log likelihood follows directly from equation (12):

$$\ln \mathcal{L}(\theta) = \sum_{j,c,t>20} \ln\Big(P(e_{j,c,t}|s_{j,c,t-1})\Big),$$

where j indexes facility, c indexes cycle, and t indexes week. In practice, discretization of the state space results in the objective function $\ln \mathcal{L}(\theta)$ being non-smooth in δ and s_0. Therefore, we employ a nested approach to optimization, performing a grid search over δ and s_0 in an outer-loop and using gradient-based methods for $\alpha_1, \alpha_2, \gamma$ in an inner loop.

5.2 Estimates and Counterfactuals

5.2.1 Estimates

Figure 7 panel (a) summarizes our estimated parameters. We find sensible estimates: the reward to a higher effort stock when inspected is approximately 8 times the per-period level cost of effort. The implied adjustment cost are quite sizable, which reflects a relatively high level of inertia apparent in the data. These adjustment costs also rationalize the fact that the inspection hazard often sharply increases, facilities ramp up their effort more slowly and starting before the increase in inspection hazard actually occurs. That we estimate a relatively small δ suggests that effort tends to depreciate quickly over time, and more recent effort choices carry substantially greater weight in the inspection payoffs. This strengthens facilities' incentive to time their efforts to be as close as possible to the inspection period.

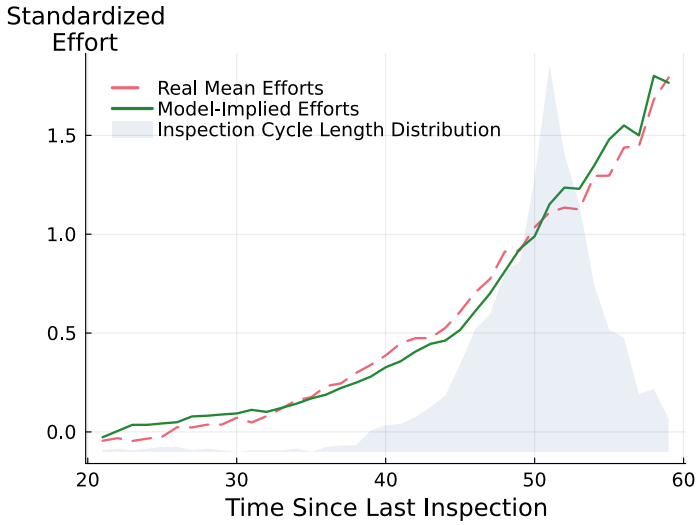
To examine the fit of our model, we compare the model-implied effort choices from equation (12) against the empirical choices. Figure 7 panel (b) compares the predicted and empirical average efforts for different periods in the inspection cycle. We find that at each period t, the average model-implied efforts are close to the empirical mean actually chosen by firms in the data. Even though our maximum likelihood procedure doesn't target these moments specifically, the quality of fit is quite high. This suggests that our model does a reasonably good job capturing the dynamics of optimal effort choice.

Figure 7: Model Fit

(a) Estimated Model Parameters

Parameter	Estimate	Description
α_1	0.076	Level effort cost
α_2	0.303	Adjustment cost
γ	0.641	Inspection reward
δ	0.201	Depreciation rate of effort stock
s_0	-0.159	Initial effort stock

(b) Model-Implied Efforts Over Time



Notes: Top panel presents estimates of the model parameters for California nursing homes. Bottom panel presents a measure of model fit, plotting both the mean California nursing home efforts over time as well as the model-implied optimal effort taking the empirical effort stock as fixed. The figure also includes the empirical distribution of inspection cycle lengths in California.

5.2.2 Counterfactuals

We use the estimated model to perform two sets of counterfactual exercises exploring alternative inspection regimes. For each, we use the model estimates to simulate counterfactual paths of facility effort—and, by extension, patient mortality as implied by our results in Section 4—over the counterfactual inspection cycles. First, to understand the importance of predictability, we examine how nursing home efforts and patient mortality would change under a fully "unpredictable" inspection regime. That is, we consider a regime with the same average frequency of inspections but in which inspections arrive at a constant hazard rate, making the elapsed time since the previous inspection completely uninformative. Second, in light of recent policy efforts to increase the pace of nursing home inspections (U.S. Senate Committee on Aging, 2023), we simulate the impact of increased inspection frequency, holding fixed the current degree of predictability. Note that as in our estimation, we omit 20 week revisit window following prior inspections to avoid taking a stand on how revisits might change under counterfactual regimes. Appendix B provides details on implementation for each of our counterfactual exercises.

Unpredictable Inspections. We begin by examining what would happen if the inspection hazard were constant across the weeks following a visit, rather than increasing gradually over the cycle. We set this constant hazard equal to the inverse of the factual average cycle length so that the average inspection frequency is the same in the factual and counterfactual. Importantly, since the counterfactual only changes the predictability of inspections and not their frequency, it represents a budget-neutral alternative to the status quo inspection regime.

Figure 8 presents the results of our counterfactual simulation. Panel (a) contrasts the counterfactual and factual hazard rates. The counterfactual constant hazard is greater than the empirical hazard rate for many periods, since inspections during the first ten months are empirically very rare. However, the factual inspection hazard grows to eclipses the counterfactual uniform hazard during the period when most inspections typically occur. Panel (b) presents the optimally chosen effort under each of the inspection regimes. The principal finding is that placing facilities under a constant probability of inspection each week encourages a more even approach to effort. Rather than exerting minimal effort in parts of the cycle with low inspection risk and then scrambling to add staff later, facilities maintain a relatively high level of staffing throughout the inspection cycle since they might be evaluated in any period. The effort path under a constant inspection hazard lies well above the factual baseline during the periods when the factual hazard is near zero but is substantially below the high level of effort facilities exert in the factual during the period

when inspections are highly likely.[11]

Panel (b) allows us to distinguish in which effort is greater under each regime. However, simply comparing the curves in panel (b) can be misleading since the choice of inspection regime also affects the distribution of cycle lengths. Panel (c) illustrates this point by plotting the share of weeks spent at each time since last inspection under both the factual and constant hazard regimes. While efforts under the factual hazard rate are higher later in the cycle (e.g., for $t > 45$), panel (c) shows that only a small fraction of periods are actually *spent* later in the cycle. Instead, the majority of the facility's time is spent early in the cycle, which are t for which facilities' efforts under factual hazards are relatively low. The overwhelming importance of effort in these early periods stems from the fact that any cycle that includes period t must also include periods $1, 2, \ldots, t-1$. This is particularly strong for factual hazards, where inspections early in the cycle are extremely rare, so almost all cycles include all $t \leq 45$.

Panel (d) synthesizes the information above by plotting a distribution efforts from panel (b) drawn according to the corresponding distribution of periods implied by panel (c). This provides the distribution of weekly efforts that facilities exert under each regime. This figure captures the main takeaway from this exercise: while there is a long tail of effort under the empirical hazards coming from the late-cycle periods just prior to inspection, the weekly average effort is considerably higher in the constant hazard counterfactual.

The link between nursing home effort and care outcomes in Section 4 suggests that this increase in average weekly effort could meaningfully benefit patients. To quantify this improvement, we employ our estimate of the relationship between effort and patient survival suggested by Figure 3. Applying this estimate to our counterfactual efforts from Figure 8 implies that moving to the flat hazard rate would generate 116.2 additional annual lives saved in California alone, or an additional 1,742.6 lives nationally.

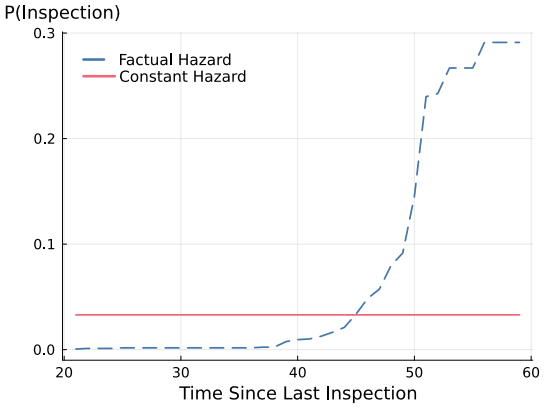
These findings underscore the importance of timing in inspection design. Even without expending resources to increase the frequency of inspections, simply making inspections more unpredictable reduces slacking in the early weeks of an inspection cycle. Doing so leads to a steadier level of effort from facilities that, on average, substantially benefits patients and even saves lives.

Increasing Inspection Frequency. Our second counterfactual increases the frequency of inspections sufficiently to reduce the average cycle length by 25%, but otherwise keeps

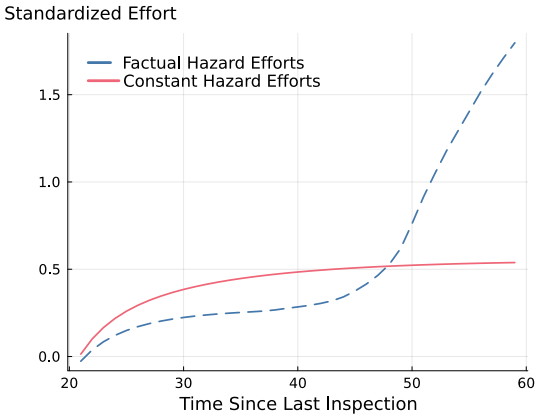
[11] That both the counterfactual and empirical effort curves begin at the same effort level is an artifact of our simulations holding the initial stock s_0 to be the same under both inspection regimes. A plausible alternative that is less conservative would be to allow s_0 in the constant hazard counterfactual to be determined in equilibrium as the long-run distribution or average of s_t for $t > 1$.

Figure 8: Effect of Making Inspections Unpredictable

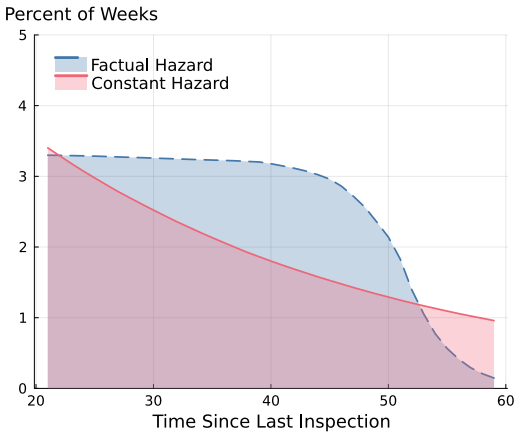
(a) Inspection Hazards



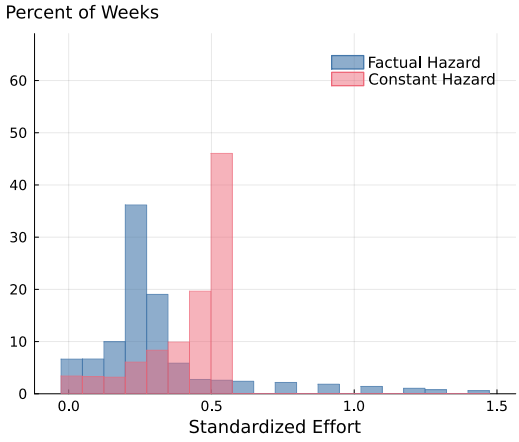
(b) Effort Curves



(c) Period Distributions



(d) Effort Distributions



Notes: Figure presents the results of the first counterfactual exercise. All results correspond to the California nursing home market, over the years 2017-2019. In this counterfactual, we impose a flat hazard function that yields a constant probability of inspection. Panel (a) presents both the factual as well as the counterfactual (flat) hazard rate. Panel (b) presents the optimal chosen effort across the inspection cycle, for both the factual and counterfactual hazard functions. Panel (c) plots the distribution of times since inspection under each of the hazard functions considered, reflecting that changes in the hazard function mechanically change where in an inspection cycle time is spent. Panel (d) combines the information from the prior two panels, plotting histograms of the optimal effort, weighted by the share of time spent allocating that level of effort.

the shape of the hazard function unchanged. Such a scenario is particularly relevant given renewed interest by policymakers in increasing the pace of nursing home inspections (U.S. Senate Committee on Aging, 2023).

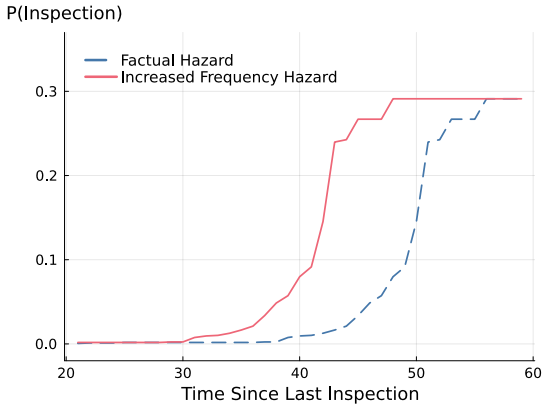
Figure 9 presents the results of our second counterfactual simulation. As before, we begin our analysis after the 20 week revisit window. Panel (a) contrasts the factual inspection hazard with the counterfactual inspection hazard that is shifted to the left. Panel (b) examines what this accelerated inspection hazard implies for optimal effort over the course of the inspection cycle. In the earliest periods, optimal effort looks similar under the two policies. However, because inspection risk increases earlier under the counterfactual, optimal efforts increase earlier too.

As with the other counterfactual, the increasing the frequency of inspections changes the distribution of how time is spent over the inspection cycle. Panel (c) presents the distributions of time since last inspection under both the factual and counterfactual. The shorter counterfactual cycles mean that although the effort curves increase earlier and reach higher levels, facilities spend even more time in the early parts of a cycle, and considerably less time in the later parts of the cycle.

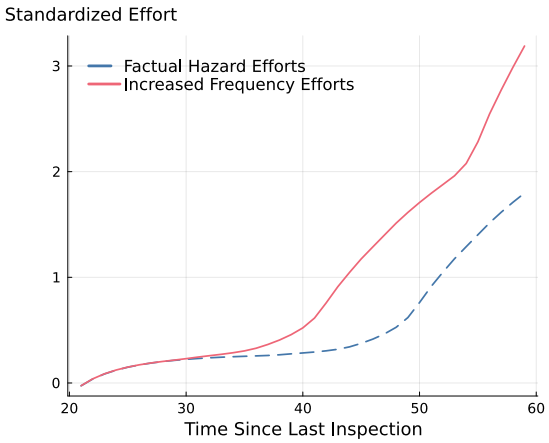
That the earliest periods have muted efforts even under the counterfactual limits the ability for the counterfactual to dramatically improve facilities' efforts. Panel (d) shows this, plotting the distribution efforts from panel (b) drawn according to the corresponding distribution of periods implied by panel (c). While the increased frequency of inspections elicits higher levels of effort throughout the cycle, the shift in the distribution of periods offsets these gains. On average, there is approximately no gain in mean efforts across facility-weeks from increasing the frequency of inspections.

Figure 9: Effect of Increasing Inspection Frequency

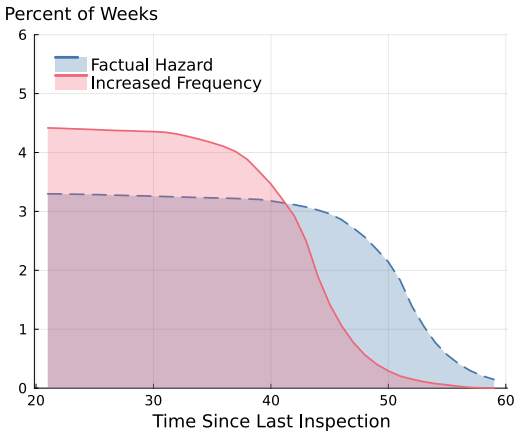
(a) Inspection Hazards



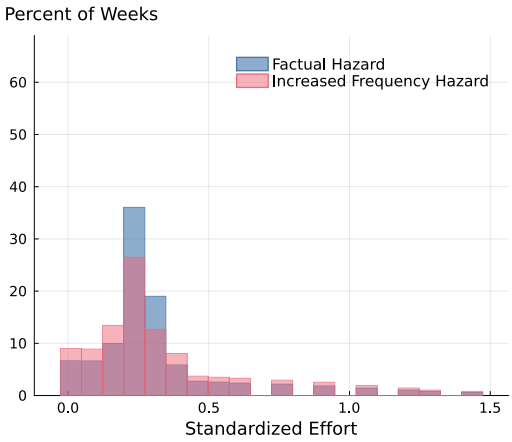
(b) Effort Curves



(c) Period Distributions



(d) Effort Distributions



Notes: Figure presents the results of the second counterfactual exercise. All results correspond to the California nursing home market, over the years 2017-2019. In this counterfactual, we impose an increased frequency hazard function that shortens inspection cycle lengths. Panel (a) presents both the factual as well as the counterfactual (increased frequency) hazard rate. Panel (b) presents the optimal chosen effort across the inspection cycle, for both the factual and counterfactual hazard functions. Panel (c) plots the distribution of times since inspection under each of the hazard functions considered, reflecting that changes in the hazard function mechanically change where in an inspection cycle time is spent. Panel (d) combines the information from the prior two panels, plotting histograms of the optimal effort, weighted by the share of time spent allocating that level of effort.

6 Conclusion

This paper explores how the timing and predictability of inspections affect staffing decisions and patient outcomes in nursing homes. Our empirical analysis reveals significant temporal

variation in staffing levels, which rise systematically as an inspection becomes more likely and fall sharply once the immediate threat subsides. We link these patterns to resident outcomes, documenting improvements in a variety of patient health measures, including mortality rates, in periods of heightened inspection risk.

To interpret these findings, we develop a simple dynamic model in which nursing homes balance staffing costs against the benefits of being well-prepared when an inspection arrives. Estimating the model with high-frequency staffing data shows that it can replicate the observed ramp-up in staffing. We then use the estimated framework to examine alternative inspection regimes. Although increasing inspection frequency can prompt additional short-term effort, we find that making inspections more random generates more persistent gains in average staffing and therefore patient outcomes. This result underscores the importance of optimally designing *when* inspections occur, not just how often, in order to sustain compliance and minimize shirking.

Taken together, our findings highlight both the promise and the limitations of inspection-based regulation. While existing policy clearly deters some level of shirking, it may inadvertently create "valleys of low effort" early in the cycle. A more randomized approach to inspections could help alleviate these lulls, potentially enhancing resident well-being without substantially raising regulatory costs. Future work might investigate how alternative policy tools, such as public disclosure of interim performance metrics or stronger enforcement penalties, could further strengthen the incentives formed by inspections.

References

Ball, Ian and Jan Knoepfle, "Should the timing of inspections be predictable?," 2023. arXiv preprint arXiv:2304.01385.

Barnett, Michael L., Andrew R. Olenski, and Anupam B. Jena, "Patient Mortality During Unannounced Accreditation Surveys at US Hospitals," *JAMA Internal Medicine*, 2017, *177* (5), 693–700.

Blundell, Wesley, Gautam Gowrisankaran, and Ashley Langer, "Escalation of Scrutiny: The Gains from Dynamic Enforcement of Environmental Regulations," *American Economic Review*, 2020, *110* (8), 2558–2585.

Brot-Goldberg, Zarek C., Samantha Burn, Timothy Layton, and Boris Vabson, "Rationing Medicine Through Bureaucracy: Authorization Restrictions in Medicare," 2023. NBER Working Paper No. 30878.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Center for Medicare Advocacy, "Report: Too Much Secrecy in the Nursing Home Enforcement System," 2023. Available at https://medicareadvocacy.org/report-too-much-secrecy-in-the-nursing-home-enforcement-system/.

Cheng, Alden, "Demand for Quality in the Presence of Information Frictions: Evidence from the Nursing Home Market," 2023. Working Paper.

Ching, Andrew T., Fumiko Hayashi, and Hui Wang, "Quantifying the Impacts of Limited Supply: The Case of Nursing Homes," *International Economic Review*, 2015, *56* (4), 1291–1322.

CMS, "SFF Scoring Methodology," 2008. Available at https://www.cms.gov/medicare/provider-enrollment-and-certification/surveycertificationgeninfo/policy-and-memos-to-states-and-regions-items/cms1215978.

_ , "Payroll-Based Journal Public Use Files: Technical Specifications," April 2022. Available at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Staffing-Data-Submission-PBJ.html.

_ , "Quality, Safety & Oversight - General Information," December 2024. Available at https://www.cms.gov/medicare/health-safety-standards/quality-safety-oversight-general-information.

_ , "Nursing Home Care Compare and Provider Data Catalog - Health Deficiencies," February 2025. Available at https://data.cms.gov/provider-data/dataset/r5ix-sfxw.

_ , "Nursing Home Care Compare and Provider Data Catalog - Inspection Dates," February 2025. Available at https://data.cms.gov/provider-data/dataset/svdt-c123.

Correia, Sergio, Paulo Guimarães, and Tom Zylkin, "Fast Poisson estimation with high-dimensional fixed effects," *The Stata Journal*, 2020, *20* (1), 95–115.

Cronin, Christopher J. and William N. Evans, "Nursing home quality, COVID-19 deaths, and excess mortality," *Journal of Health Economics*, 2022, *82*, 102592.

Deshpande, Manasi and Yue Li, "Who is screened out? Application costs and the targeting of disability programs," *American Economic Journal: Economic Policy*, 2019, *11* (4), 213–248.

Dillender, Marcus, "What happens when the insurer can say no? Assessing prior authorization as a tool to prevent high-risk prescriptions and to lower costs," *Journal of Public Economics*, 2018, *165*, 170–200.

Dranove, David and Ginger Zhe Jin, "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 2010, *48* (4), 935–963.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India," *Econometrica*, 2018, *86* (6), 2123–2160.

Dunn, Abe, Joshua D Gottlieb, Adam Hale Shapiro, Daniel J Sonnenstuhl, and Pietro Tebaldi, "A Denial a Day Keeps the Doctor Away," *Quarterly Journal of Economics*, 2024, *139* (1), 187–233.

Einav, Liran, Amy Finkelstein, and Neale Mahoney, "Producing Health: Measuring Value Added of Nursing Homes," 2022. NBER Working Paper No. 30228.

Eliason, Paul, Riley League, Jetson Leder-Luis, Ryan McDevitt, and James Roberts, "Ambulance Taxis: The Impact of Regulation and Litigation on Health Care Fraud," *Journal of Political Economy*, 2024.

Friedrich, Benjamin U and Martin B Hackmann, "The returns to nursing: Evidence from a parental-leave program," *Review of Economic Studies*, 2021, *88* (5), 2308–2343.

Furtado, Delia and Francesc Ortega, "Does Immigration Improve Quality of Care in Nursing Homes?," *Journal of Human Resources*, 2023.

Gandhi, Ashvin, "Picking Your Patients: Selective Admissions in the Nursing Home Industry," 2023. Working Paper.

_ **and Andrew Olenski**, "Tunneling and Hidden Profits in Health Care," 2024. NBER Working Paper No. 32258.

_ , _ , **Krista Ruffini, and Karen Shen**, "Alleviating Worker Shortages Through Targeted Subsidies: Evidence from Incentive Payments in Healthcare," *Review of Economics and Statistics*, 2024, pp. 1–31.

_ , **Huizi Yu, and David C. Grabowski**, "High Nursing Staff Turnover In Nursing Homes Offers Important Quality Information," *Health Affairs*, 2021, *40* (3), 384–391.

_ , **Youngjun Song, and Prabhava Upadrashta**, "Have Private Equity Owned Nursing Homes Fared Worse Under COVID-19?," 2020. Working Paper.

Geng, Fangli, David G. Stevenson, and David C. Grabowski, "Daily Nursing Home Staffing Levels Highly Variable, Often Below CMS Expectations," *Health Affairs*, 2019, *38* (7), 1095–1100.

Grabowski, David C, "Medicaid reimbursement and the quality of nursing home care," *Journal of Health Economics*, 2001, *20* (4), 549–569.

_ **and Robert J Town**, "Does information matter? Competition, quality, and the impact of nursing home report cards," *Health Services Research*, 2011, *46* (6pt1), 1698–1719.

_ , **Jonathan Gruber, and Joseph J Angelelli**, "Nursing home quality as a common good," *Review of Economics and Statistics*, 2008, *90* (4), 754–764.

_ , **Zhanlian Feng, Richard Hirth, Momotazur Rahman, and Vincent Mor**, "Effect of nursing home ownership on the quality of post-acute care: an instrumental variables approach," *Journal of health economics*, 2013, *32* (1), 12–21.

Grabowski, David, Jonathan Gruber, and Brian McGarry, "Immigration, The Long-Term Care Workforce, and Elder Outcomes in the US," 2023. NBER Working Paper No. 30960.

Guimarães, Paulo, Octávio Figueirdo, and Douglas Woodward, "A Tractable Approach to the Firm Location Decision Problem," *Review of Economics and Statistics*, 2003, *85* (1), 201–204.

Gupta, Atul, Sabrina T Howell, Constantine Yannelis, and Abhinav Gupta, "Owner incentives and performance in healthcare: Private equity investment in nursing homes," *The Review of Financial Studies*, 2024, *37* (4), 1029–1077.

Hackmann, Martin B, "Incentivizing better quality of care: The role of Medicaid and competition in the nursing home industry," *American Economic Review*, 2019, *109* (5), 1684–1716.

Hackmann, Martin B., R. Vincent Pohl, and Nicolas R. Ziebarth, "Patient versus Provider Incentives in Long-Term Care," *American Economic Journal: Applied Economics*, July 2024, *16* (3), 178–218.

Jin, Ginger Zhe and Phillip Leslie, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *The Quarterly Journal of Economics*, 2003, *118* (2), 409–451.

Johnson, Matthew S., David I. Levine, and Michael W. Toffel, "Improving Regulatory Effectiveness through Better Targeting: Evidence from OSHA," *American Economic Journal: Applied Economics*, 2023, *15* (4), 30–67.

Konetzka, R Tamara and Marcelo Coca Perraillon, "Use of nursing home compare website appears limited by lack of awareness and initial mistrust of the data," *Health Affairs*, 2016, *35* (4), 706–713.

Konetzka, R. Tamara, Sally C. Stearns, and Jeongyoung Park, "The staffing-outcomes relationship in nursing homes," *Health Services Research*, 2008, *43* (3), 1025–1042.

Konetzka, Tamara, Kevin Yan, and Rachel M Werner, "Two decades of nursing home compare: what have we learned?," *Medical Care Research and Review*, 2021, *78* (4), 295–310.

Lazear, Edward P., "Speeding, Terrorism, and Teaching to the Test," *The Quarterly Journal of Economics*, 2006, *121* (3), 1029–1061.

League, Riley, "Administrative Burden and Consolidation in Health Care: Evidence from Medicare Contractor Transitions," 2022. Working Paper.

Leder-Luis, Jetson, "Can Whistleblowers Root Out Public Expenditure Fraud? Evidence from Medicare," *Review of Economics and Statistics*, 2023, pp. 1–49.

Levine, David I., Michael W. Toffel, and Matthew S. Johnson, "Randomized government safety inspections reduce worker injuries with no detectable job loss," *Science*, 2012, *336* (6083), 907–911.

Lin, Haizhen, "Revisiting the relationship between nurse staffing and quality of care in nursing homes: An instrumental variables approach," *Journal of Health Economics*, 2014, *37*, 13–24.

Macambira, Danil, Michael Geruso, Anthony Lollo, Chima D. Ndumele, and Jacob Wallace, "The Private Provision of Public Services: Evidence from Random Assignment in Medicaid," 2022. NBER Working Paper No. 30390.

Matsudaira, Jordan D, "Government Regulation and the Quality of Healthcare Evidence from Minimum Staffing Legislation for Nursing Homes," *Journal of Human Resources*, 2014, *49* (1), 32–72.

McClellan, Mark, Barbara J McNeil, and Joseph P Newhouse, "Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality?: Analysis using instrumental variables," *JAMA*, 1994, *272* (11), 859–866.

National Academies of Sciences Engineering and Medicine, *The National Imperative to Improve Nursing Home Quality: Honoring Our Commitment to Residents, Families, and Staff*, Washington, DC: National Academies Press, 2022.

Olenski, Andrew, "Reallocation and the (In)efficiency of Exit in the U.S. Nursing Home Industry," 2023. Working Paper.

_ **and Szymon Sacher**, "Estimating nursing home quality with selection," *Review of Economics and Statistics*, 2024, pp. 1–31.

Olken, Benjamin A., "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, 2007, *115* (2), 200–249.

Rau, Jordan, ""Like A Ghost Town": Erratic Nursing Home Staffing Revealed Through New Records," *Kaiser Health News* July 2018.

Ruffini, Krista, "Worker Earnings, Service Quality, and Firm Profitability: Evidence from Nursing Homes and Minimum Wage Reforms," *Review of Economics and Statistics*, 2022, pp. 1–46.

Shi, Maggie, "Monitoring for Waste: Evidence from Medicare Audits*," *The Quarterly Journal of Economics*, 2024, *139* (2), 993–1049.

Thomas, Katie, "Medicare Star Ratings Allow Nursing Homes to Game the System," *The New York Times* August 2014.

U.S. Congress, "Omnibus Budget Reconciliation Act of 1987," H.R. 3545, Public Law 100–203, 101 Stat. 1330 (1987) 1987. Enacted by the 100th U.S. Congress. Signed into law on December 22, 1987.

U.S. Senate Committee on Aging, "Uninspected and Neglected: A Report by the Majority Staff of the U.S. Senate Special Committee on Aging," Technical Report, U.S. Senate May 2023.

Varas, Felipe, Iván Marinovic, and Andrzej Skrzypacz, "Random Inspections and Periodic Reviews: Optimal Dynamic Monitoring," *Review of Economic Studies*, 2020, *87* (6), 2893–2937.

Werner, Rachel M, Edward C Norton, R Tamara Konetzka, and Daniel Polsky, "Do consumers respond to publicly reported quality information? Evidence from nursing homes," *Journal of Health Economics*, 2012, *31* (1), 50–61.

Zou, Eric Yongchen, "Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality," *American Economic Review*, July 2021, *111* (7), 2101–2126.

Online Appendix

A Alternative Event Study Approach

Our primary research approach – the inspection-cycle specification – restricts the sample to weeks $w \leq 60$ from last inspection to ensure sufficient sample size to estimate β_w. Yet even with this restriction, there could still be differences in sample composition in the later weeks of a cycle, as not all inspection cycles last at least 60 weeks. Accordingly, the estimates for weeks further away from the last inspection may be biased by differences in the types of facilities with observations for these weeks. This concern motivates an alternative event study specification. While any static differences across nursing homes should be addressed by the inclusion of time-invariant facility fixed effects α_j, the aim of this secondary specification is to ensure that our primary results are robust to differences in sample composition.

Design. In this specification, we use a balanced panel of weeks around each inspection. When constructing the panel of inspection events for this sample, we include the 30 weeks before and after each inspection for each facility, regardless whether another inspection occurs in these weeks. The resulting facility-week panel is balanced in that each inspection event is associated with a 61 week window, including the week of the inspection. Consequently, the corresponding estimates are unaffected by differences in sample composition arising from varying inspection cycle lengths. We then "stack" each inspection event together in the spirit of Cengiz et al. (2019) and Deshpande and Li (2019), again clustering our standard errors at the facility and calendar week level to account for duplicate facility-week observations.

Regression equation (14) summarizes this stacked event study design for facility j in calendar week t and index inspection event c:

$$Y_{jct} = \sum_{r \in [-30,-1)}^{[0,30]} \beta_r \ \mathbb{1}(\text{Week Rel. Inspection}_{jct} = r) + \alpha_j + \lambda_t + \delta_c + \varepsilon_{jct}, \tag{14}$$

where $\mathbb{1}(\text{Week Rel. Inspection}_{jct} = r)$ is an indicator for whether facility j is r weeks relative to the inspection that defines event c in calendar week t. We augment the fixed effects structure from equation (1) with an additional inspection event fixed effect δ_c. Because the omitted term is relative week $r = -1$, the parameters of interest (the β_r coefficients) can be interpreted as the difference in the outcome relative to the week prior to the index inspection event.

For completeness, as with equation (2), we also consider the corresponding specification

for patient i:

$$Y_{ict} = \beta_0 + \sum_{r \in [0,30)}^{(30,60]} \beta_r \, \mathbb{1}(\text{Week Rel. Inspection}_{j(i),c(i),t} = r) + \gamma X_i + \alpha_{j(i)} + \lambda_t + \delta_{c(i)} + \varepsilon_{ict} \quad (15)$$

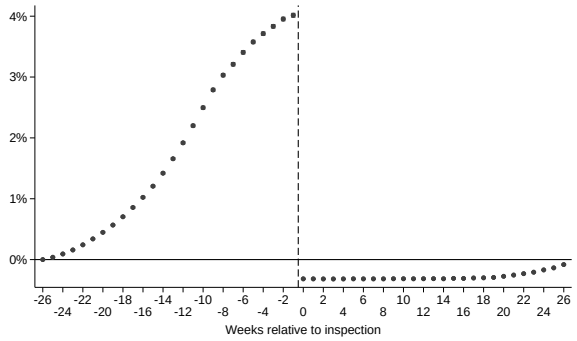
where, as before, X_i controls for patient i's predicted mortality at admission.

Results. Our findings are consistent with the results from Section 4.1. Appendix Figure A.1 plots the event studies results for the inspection hazard, staffing hours per resident-day, and the 4-week patient survival rate. Panel (a) shows that inspection hazard is rising rapidly in the weeks ahead of an inspection, sharply falls at the week of the inspection, and remains relatively very low in the 6 months following the inspection. Staffing largely follows an "inverse-U" shape, with a spike in the weeks -1 and 0. The smaller spike in staffing in week -1 reflects the fact that inspections are 4 days long on average (panel b). Since we measure inspection date as when the inspection is completed, in many instances the inspection has already begun in week -1.[12] The patterns of the patient survival results mirror the staffing results – patient survival increases in the weeks leading up to the inspection when inspection risk is highest and then ramps once the inspection has passed (panel c).

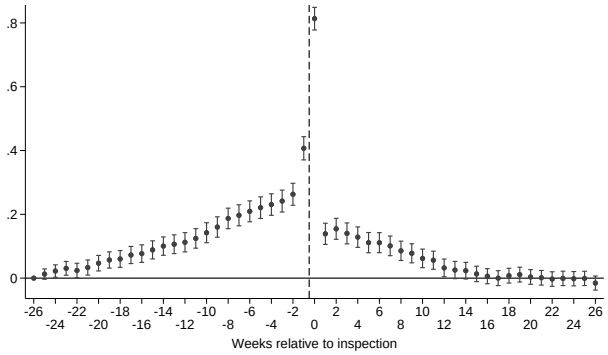
[12]Appendix Figure C.12 plots daily event studies in the 60 days before and after the inspection, centered around either the start date (panel a), which is available for a subset of years through FOIA request, or the end date (panel b), which is available for our entire sample. These indicate that the staffing spike begins once the inspection has begun.

Figure A.1: Event Studies: Inspection Hazard, Effort, and Patient Survival

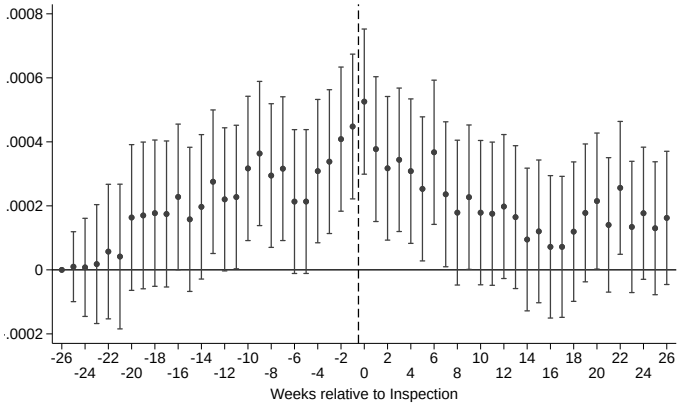
(a) Inspection Hazard



(b) Staffing Hours per Resident-Day (Std. Dev.)



(c) 4-Week Survival Rate



Notes: This figure plots the estimates from equations (14) and (15) on (a) the inspection hazard rate, (b) standardized hours per resident-day, and (c) 4-week patient survival rate.

B Further Details on Counterfactuals

Unpredictable Inspections: Our first counterfactual scenario – which imposes a constant inspection hazard rate – asks what the equilibrium effort levels would be under a memoryless inspection arrival rate. Because our model is in discrete time, where the unit of time is the number of weeks between inspections, this counterfactual is equivalent to imposing a geometric distribution on the time between inspections.[13]

Recall that our baseline model uses a monotone transformation of the empirical hazard rate, denoted h, to ensure that inspection risk is non-decreasing over time. We start by computing the mean number of periods between arrivals implied by this transformed hazard rate using the following formula:

$$E[L] = \sum_{w=1}^{T} \left(\prod_{\tau=1}^{w-1} (1 - h_\tau) \right) \tag{16}$$

where L is a random variable denoting the number of periods between inspections and h_τ is the monotone empirical hazard for τ periods since the last inspection. The mean number of periods between inspections is the sum of the survival rates for each period. The data contain a maximum observed number of periods, T, which we use for determining the sum.

$E[L]$ is therefore the mean number of periods between inspections implied by the monotone-transformed empirical hazard rate. Our first counterfactual imposes $L \sim \text{Geom}(1/E[L])$. Because the geometric distribution carries the memoryless property, this implies a counterfactual constant hazard $\tilde{h}(t) := \bar{h} = 1/E[L]$. We then re-solve the Bellman equation using this constant hazard rate $\bar{h}$ holding our estimates of the parameters fixed, and simulate counterfactual efforts implied by this alternative policy.

Increasing Inspection Frequency: Our second counterfactual scenario – increasing inspection frequency without changing the shape of the hazard rate – asks what the equilibrium effort levels would be if we maintained the current inspection regime (i.e., did not switch to a memoryless process) but simply increased the pace of nursing home inspections. Doing so requires an alternative hazard function $\tilde{h}(t)$ such that the implied mean time between inspections is 25% lower than the one implied by the monotone empirical rate.

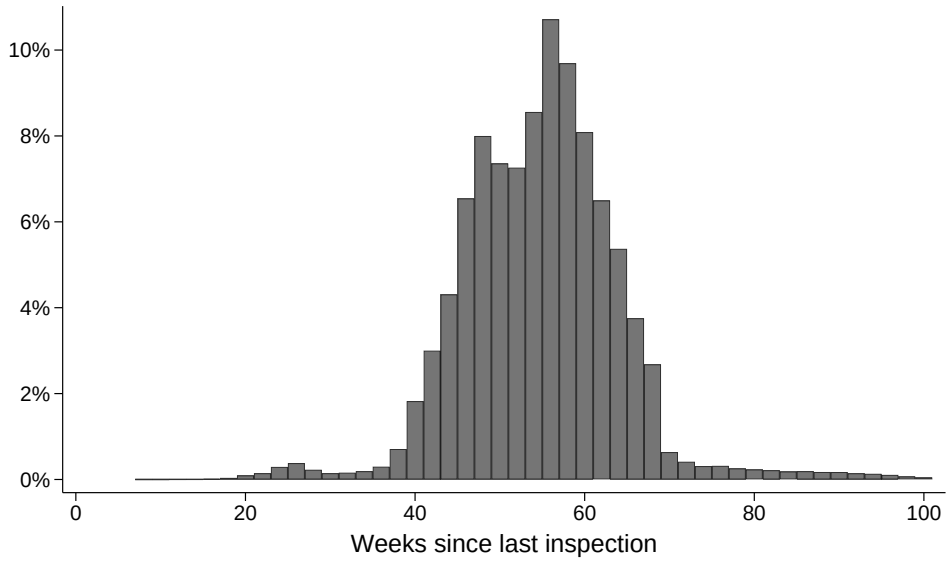
To find such an increased frequency hazard function $\tilde{h}(t)$ without changing the original function's 'shape', we employ the following iterative procedure. Starting from the original empirical hazard $h(t)$ (a vector), we remove one period from the start and add one period of $\max h(t)$ to the end. We then compute the new $\widetilde{E[L]}$ implied by this modified hazard

[13] A continuous-time alternative would instead use an exponential distribution for this exercise.

vector using equation (16). We iterate this procedure until the modified mean time between inspections $\widetilde{E[L]} \approx \alpha E[L]$, where for example $\alpha = 0.75$ implies a 25% decrease in the mean inspection frequency. Doing so 'shifts' the hazard function to the left, while maintaining the function's shape beyond the initial excised region.

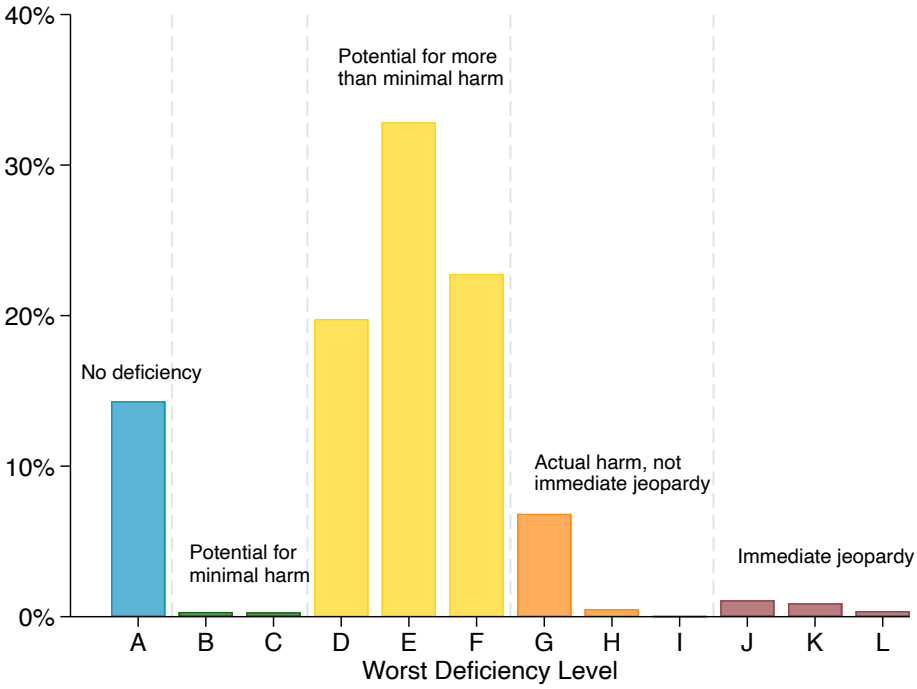
C Appendix Figures

Figure C.1: Histogram of Weeks Since Last Inspection



Notes: This figure plots a histogram of weeks since last inspection for standard health surveys in 2013-2019. The mean is 54.4 and the standard deviation is 9.9.

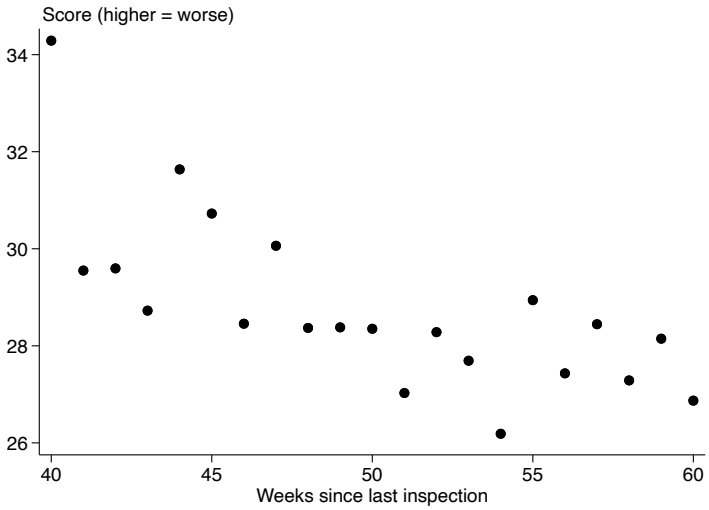
Figure C.2: Inspection Outcomes, by Highest Deficiency Category



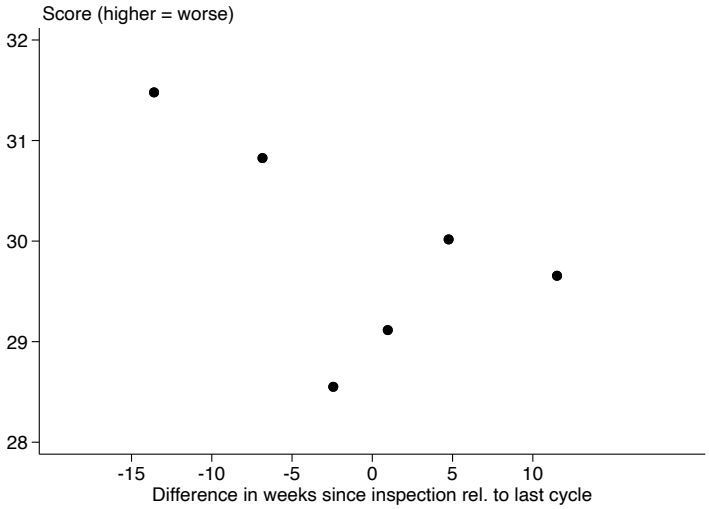
Notes: This figure plots a histogram of inspection outcomes in 2013-2019 by the highest (worst) deficiency level. Level A corresponds to no deficiency or an isolated deficiency with no actual harm with potential for minimal harm was found, while levels B-C correspond to a pattern of or widespread deficiencies (respectively) that pose no actual harm with potential for minimal harm. Levels D-F correspond to isolated, pattern of, and widespread deficiencies (respectively) with no actual harm with potential for more than minimal harm that is not immediate jeopardy. Levels G-I correspond to isolated, pattern of, and widespread deficiencies (respectively) with actual harm that is not immediate jeopardy. Levels J-K correspond to isolated, pattern of, and widespread deficiencies (respectively) with immediate jeopardy to resident health or safety (CMS, 2008).

Figure C.3: Relationship Between Time Since Last Survey and Inspection Outcome

(a) Deficiency Score vs. Time Since Last Inspection

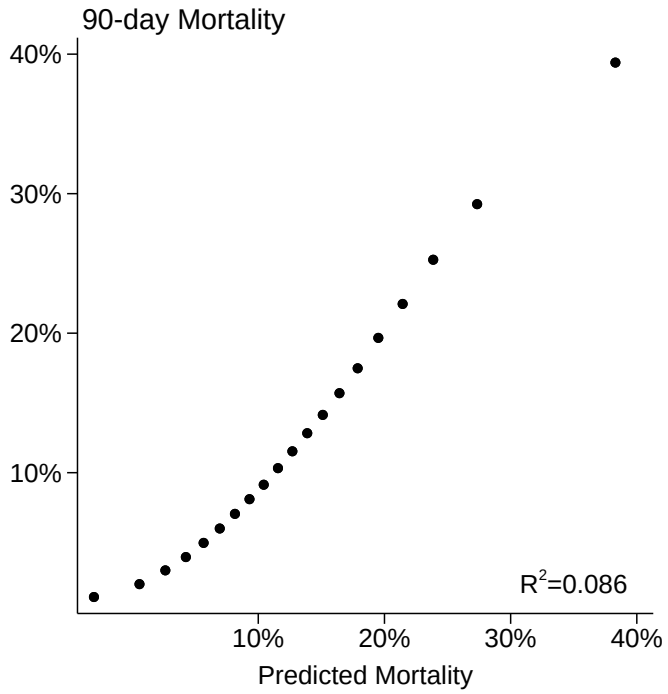


(b) Deficiency Score vs. Difference in Time Since Last Inspection Relative to Previous Cycle



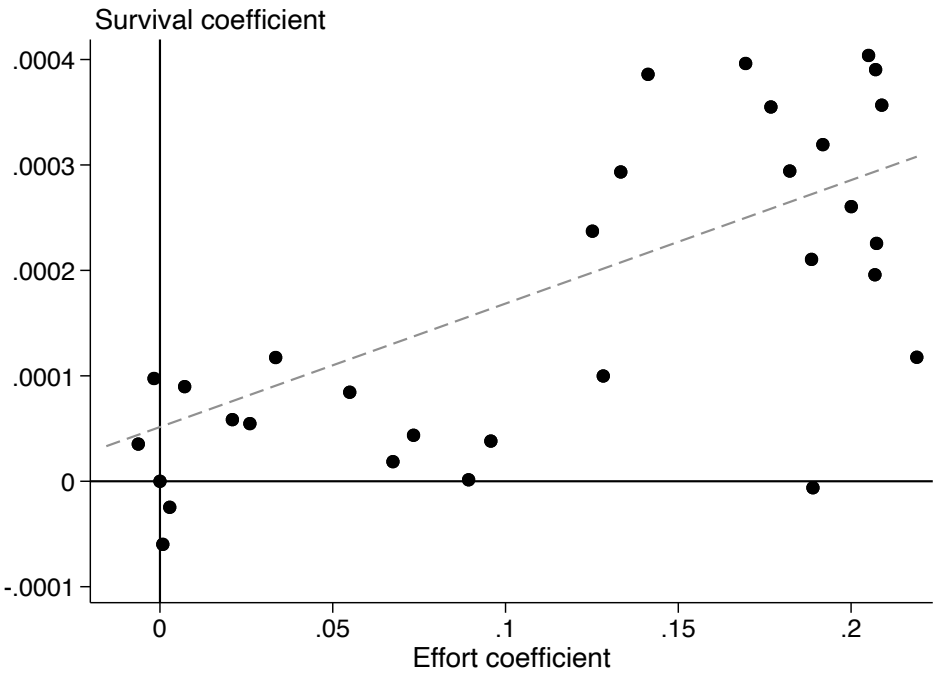
Notes: This figure plots binscatters of the relationship between the inspection deficiency score and (a) the number of weeks since a facility's last inspection and (b) the difference in the number of weeks since the last inspection, relative to the previous inspection cycle. The deficiency score is the sum of the total number of deficiencies found during the inspection, weighted by their severity. It is used to assign facilities to the Special Focus Facility (SFF) program, which subjects the worst-scoring facilities to more frequent inspections. We drop observations during which a facility is participating in SFF in the analysis. State-year, facility, and calendar week fixed effects have been residualized out and the sample is restricted to cycles with 40-60 weeks between inspections. The mean and median deficiency scores are 28.13 and 16, respectively.

Figure C.4: Predicted vs. Actual Mortality



Notes: This figure plots a binscatter of predicted vs. actual 90-day mortality after admission to the facility. We use a high-dimensional fixed effect model that includes age, race, sex, chronic conditions noted in the initial admission assessment, pre-admission hospital stay diagnosis group (DRG) and length of stay, initial payer, and whether the patient qualifies for end stage renal dialysis coverage.

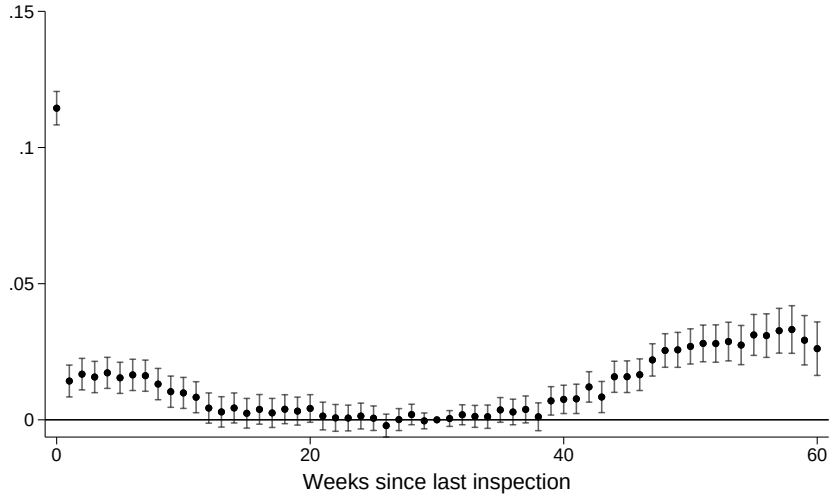
Figure C.5: Comparison of Inspection Hazard to Effort vs. Survival Coefficients, Restricted to $\geq$ 30 Weeks from Inspection



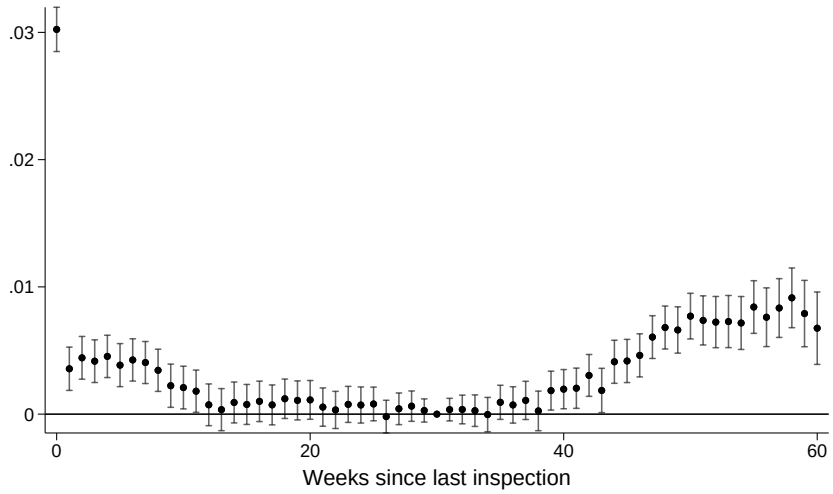
Notes: This figure plots the effort coefficients and survival coefficients from estimating equations (14) and (15) in Figure 2, restricted to weeks from inspection $w > 30$. The slope of the line of best fit is 0.0012973 (SE: 0.00023).

Figure C.6: Alternative Staffing Measures by Weeks Since Last Inspection

(a) Hours Per Resident-Day



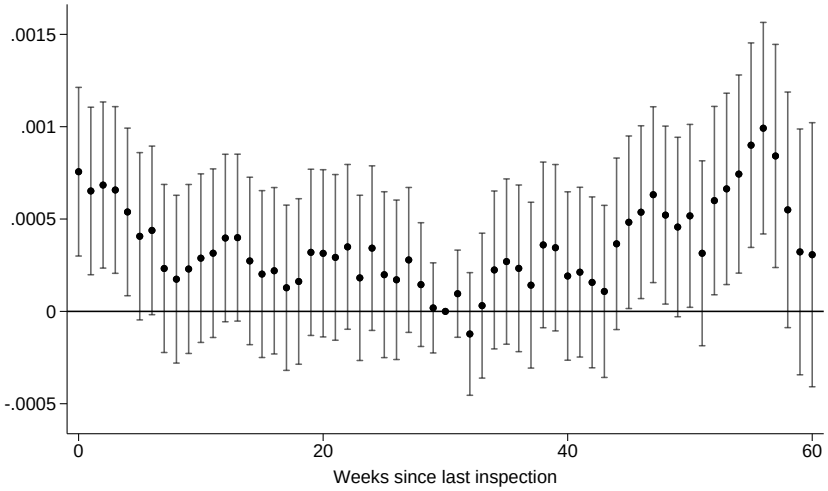
(b) Log Total Hours



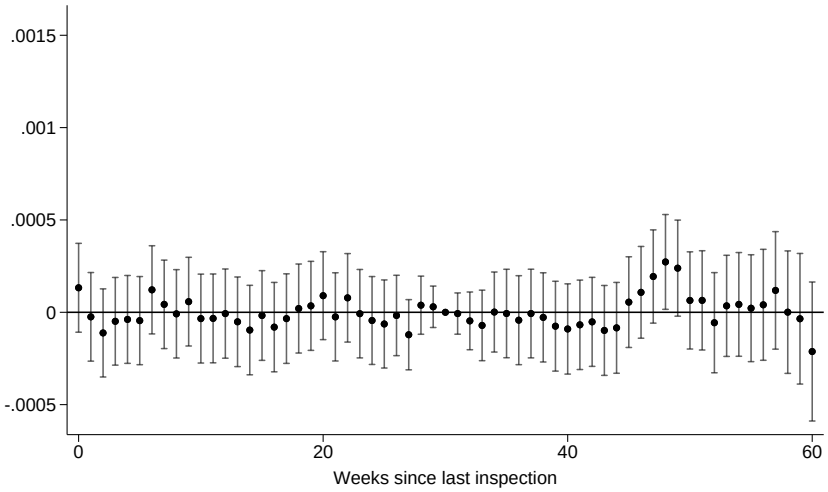
Notes: This figure plots the estimates from equation (1) on alternative nursing home staffing measures: (a) hours per resident-day and (b) log total hours.

Figure C.7: Patient Survival by Weeks Since Last Inspection, By Week of Stay

(a) Coefficients on 4-Week Survival Rate, $\leq$ 12 Weeks From Admission

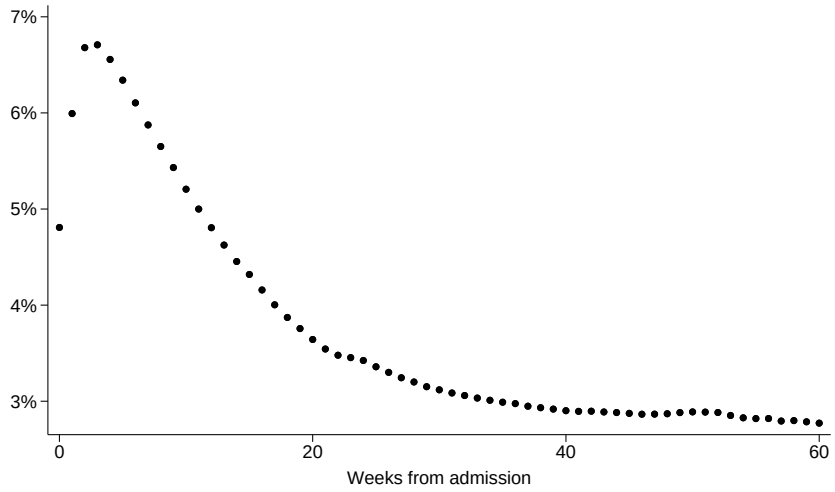


(b) Coefficients on 4-Week Survival Rate, $>$ 12 Weeks From Admission



Notes: This figure plots the estimates from equation (2), split by whether a patient is (a) less than 13 weeks from admission (i.e., in their first 90 days of the stay) or (b) greater than 13 weeks from admission.

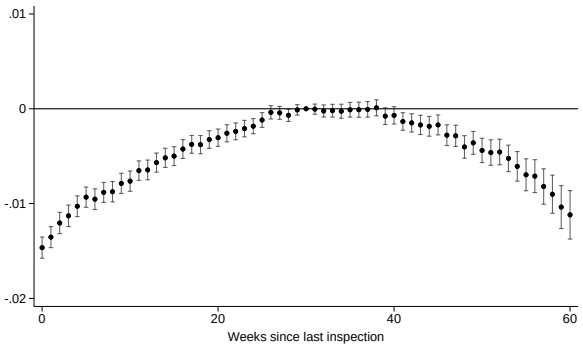
Figure C.8: Patient Mortality, by Weeks from Admission



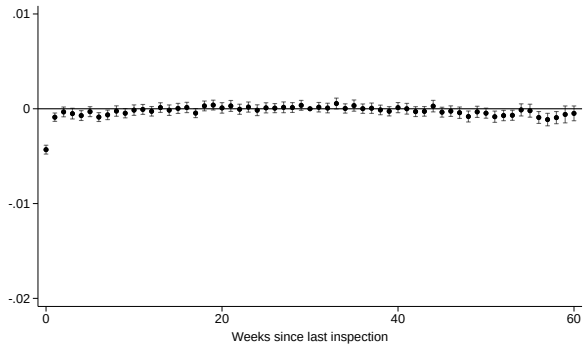
Notes: This figure plots the 4-week patient mortality rate by weeks from admission.

Figure C.9: Patient Census, Admissions, and Discharges by Weeks Since Last Inspection

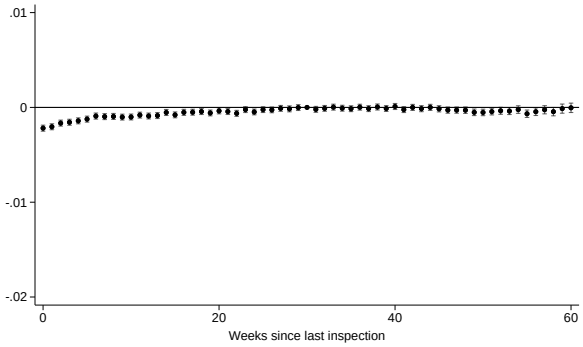
(a) Patient Census



(b) Admissions



(c) Discharges

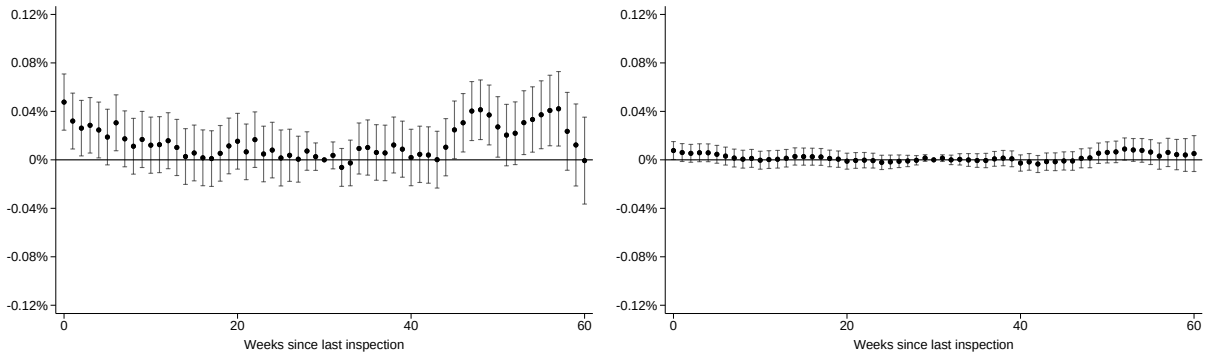


Notes: This figure plots the coefficients from estimating equation (1) on (a) patient census, or the number of patients in a facility in a given week, (b) admissions, and (c) discharges. All values are normalized in percent terms relative to the average patient census in the omitted week. We construct patient stay spells using MDS admission and discharge assessments. We identify a stay as starting when a patient is observed receiving an admission assessment, and ending with a discharge assessment where return is not anticipated, when a patient dies, or if another facility reports an admission assessment for the patient. If a stay does not have a discharge assessment, death, or is interrupted by another admission, then we set the end date of the stay to be 150 days after the last observed assessment.

Figure C.10: Actual and Predicted Survival Rates by Weeks Since Last Inspection

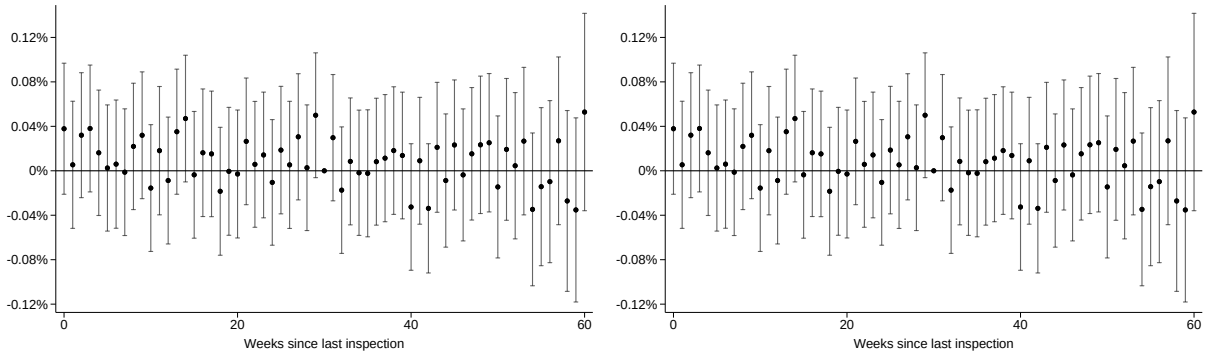
(a) Actual Survival, All Patients

(b) Predicted Survival, All Patients



(c) Predicted Survival, Admissions

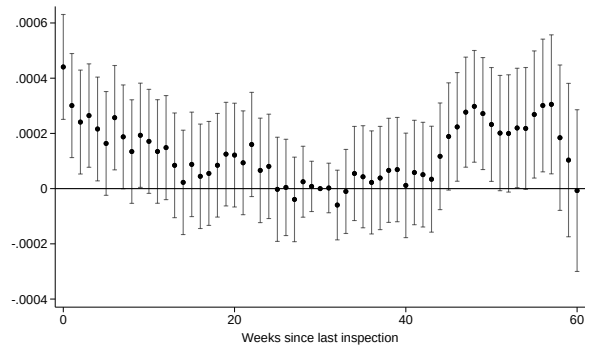
(d) Predicted Survival, Discharges



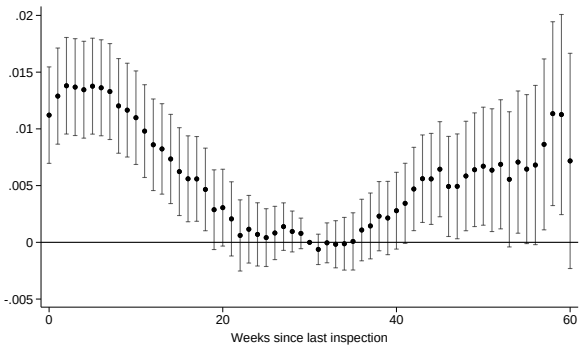
Notes: This figure plots the coefficients from estimating equation (2) on actual and predicted survival rates. Panel (a) recreates the plot from Figure A.1 and depicts the 4-week survival rate for all patients in the facility, scaled to percent terms. Panel (b) plots the 90-day predicted survival rate from admission for all patients in the facility, panel (c) plots the predicted survival rate for all admissions in a facility week, and panel (d) plots the predicted survival rate for all discharges. All estimates are normalized in percent terms relative to the average of the outcome for all patients in the omitted week.

Figure C.11: Survival Rate and Additional Health Outcomes by Weeks Since Last Inspection, without Predicted Survival Controls

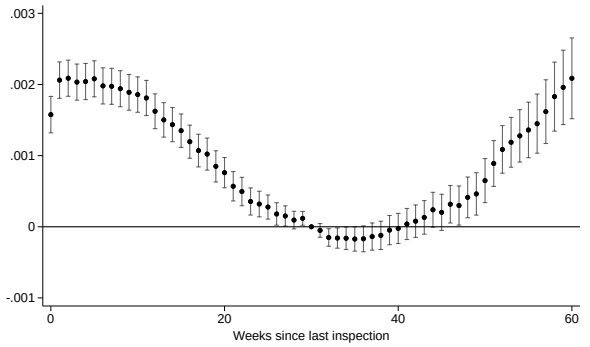
(a) 4-Week Survival Rate



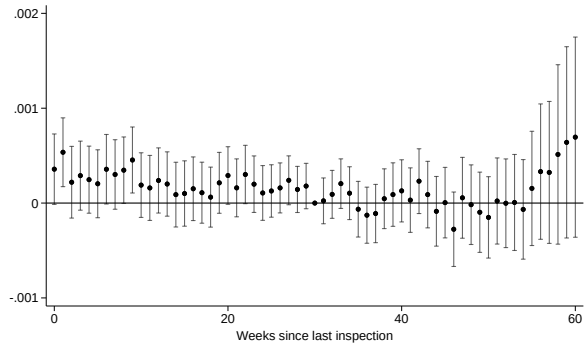
(b) (Negative) Activities of Daily Living Score



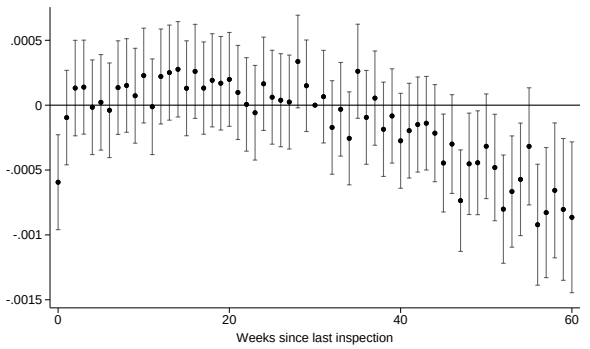
(c) Share of Patients Not on Anti-Psychotics



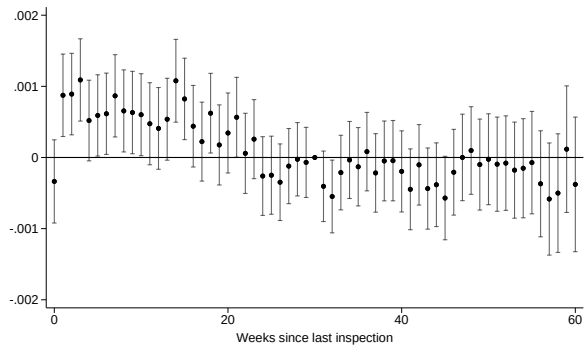
(d) Share of Patients Without Restraints



(e) Share of Patients Without Fall or Fracture



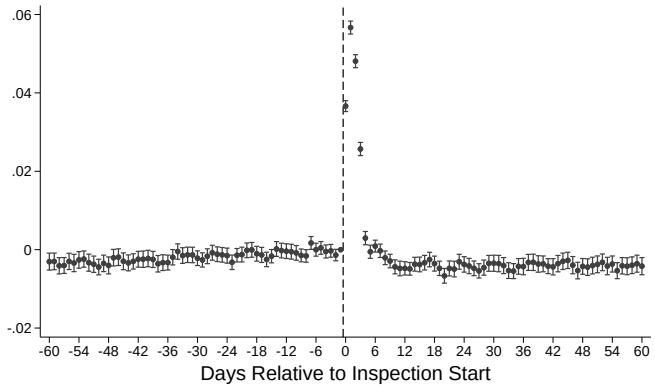
(f) Share of Patients Without Pressure Ulcers



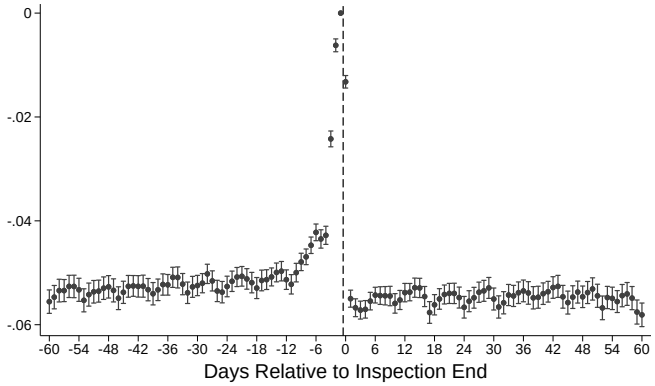
Notes: This figure plots the estimates from equation (2), excluding the control for predicted patient survival.

Figure C.12: Daily Event Studies, by Inspection Start and End Dates

(a) Staffing Hours per Resident-Day, relative to Inspection Start Date

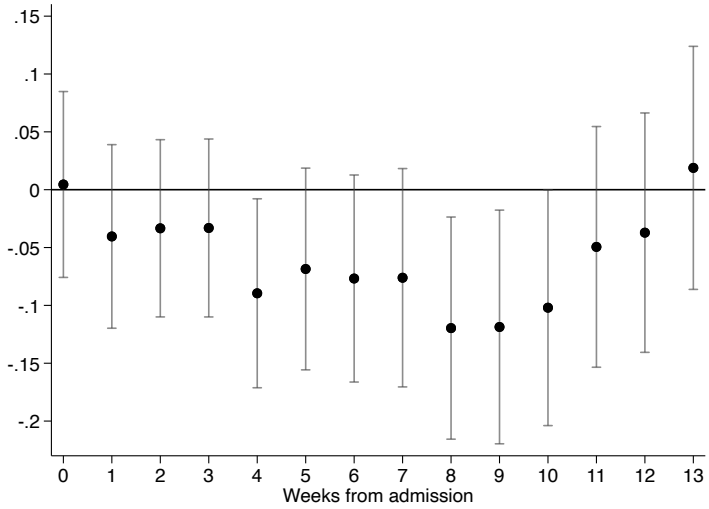


(b) Staffing Hours per Resident-Day, relative to Inspection End Date



Notes: This figure plots the estimates from a daily analogue to the event study specification in equation (14), centered around the inspection start date (panel a) and inspection end date (panel b). Survey start and end dates are acquired from a Freedom of Information Act request for inspections between 2015 and 2023.

Figure C.13: Distance IV Mortality Coefficients, Unscaled



Notes: This figure plots the coefficients from the instrumented distance specification in equation (7) where the outcome variable is 4-week patient mortality. The regression is estimated on weeks 0-59 from admission and the plot shows the estimates on the first 13 weeks (equivalent to approximately the first 90 days) of the stay.