NBER WORKING PAPER SERIES

PREDICTABLY UNPREDICTABLE INSPECTIONS

Ashvin Gandhi
Andrew Olenski
Maggie Shi

Working Paper 34491
http://www.nber.org/papers/w34491

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2025

We are very grateful to Nicolas R. Ziebarth for an excellent discussion of this paper. We also thank Zarek Brot, Yiqun Chen, Liran Einav, Amy Finkelstein, Joshua Gottlieb, Neale Mahoney, Kelli Marquardt, Ariel Pakes and Yu Fu Wong for helpful conversations and discussions, as well as seminar audiences at the Whistler Junior Health Economics Summit, Chicago Health Economics Workshop, RAND, the Johns Hopkins Health Care Economics Workshop, Stanford University, ASHEcon, the NBER Summer Institute, and the California Health Economics Conference. Aman Aggarwal, Michelle Gard, Sangwook Suh, and Fuman Xie provided excellent research assistance. This project received financial support from Arnold Ventures and the Becker Friedman Institute. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Predictably Unpredictable Inspections
Ashvin Gandhi, Andrew Olenski, and Maggie Shi
NBER Working Paper No. 34491
November 2025
JEL No. I18, L51

ABSTRACT

Inspections are a common tool for acquiring information and incentivizing compliance. Although inspections are typically unannounced, their timing often follows a predictable schedule. We study how this predictability shapes firm effort and patient outcomes in U.S. nursing homes, leveraging detailed administrative data on staffing, care, and health outcomes. Nursing homes "slack" in the low-risk period following an inspection and ramp up effort as their next inspection approaches. Patient survival mirrors this pattern, suggesting that these fluctuations in effort have meaningful consequences for the quality of patient care. We embed these estimates in a dynamic model capturing how inspection regimes incentivize effort and provide information about quality. Our estimates indicate that moving to unpredictable inspections could induce as much additional effort as increasing the frequency of inspections by 12%, while only minimally reducing their informational value.

Ashvin Gandhi
University of California, Los Angeles
Anderson School of Management
and NBER
ashvin.gandhi@anderson.ucla.edu

Maggie Shi
University of Chicago
Harris School of Public Policy
and NBER
m.shi@uchicago.edu

Andrew Olenski
Lehigh University
ano223@lehigh.edu

1 Introduction

Inspections are a key tool used by principals to acquire information and to incentivize compliance from agents. Government regulators use inspections widely to ensure compliance with various standards, including for product quality, environmental protection, safety, and public health. Private firms likewise employ inspections to monitor suppliers, contractors, franchisees, and even their own establishments.

Inspections can directly improve compliance by identifying, penalizing, and correcting violations. However, even in periods when an inspection does not occur, the mere threat of one can incentivize compliance. Inspection regimes commonly leverage this, specifying only approximate frequencies—such as annual or biannual—rather than exact dates. While this avoids immediate anticipatory "window-dressing," the requirement that inspections happen at regular intervals necessarily means that their timing is often partially *predictable*. Figure 1 illustrates this across a variety of regulatory contexts: in each case, the time elapsed since the last inspection is highly informative about whether the next inspection is imminent.

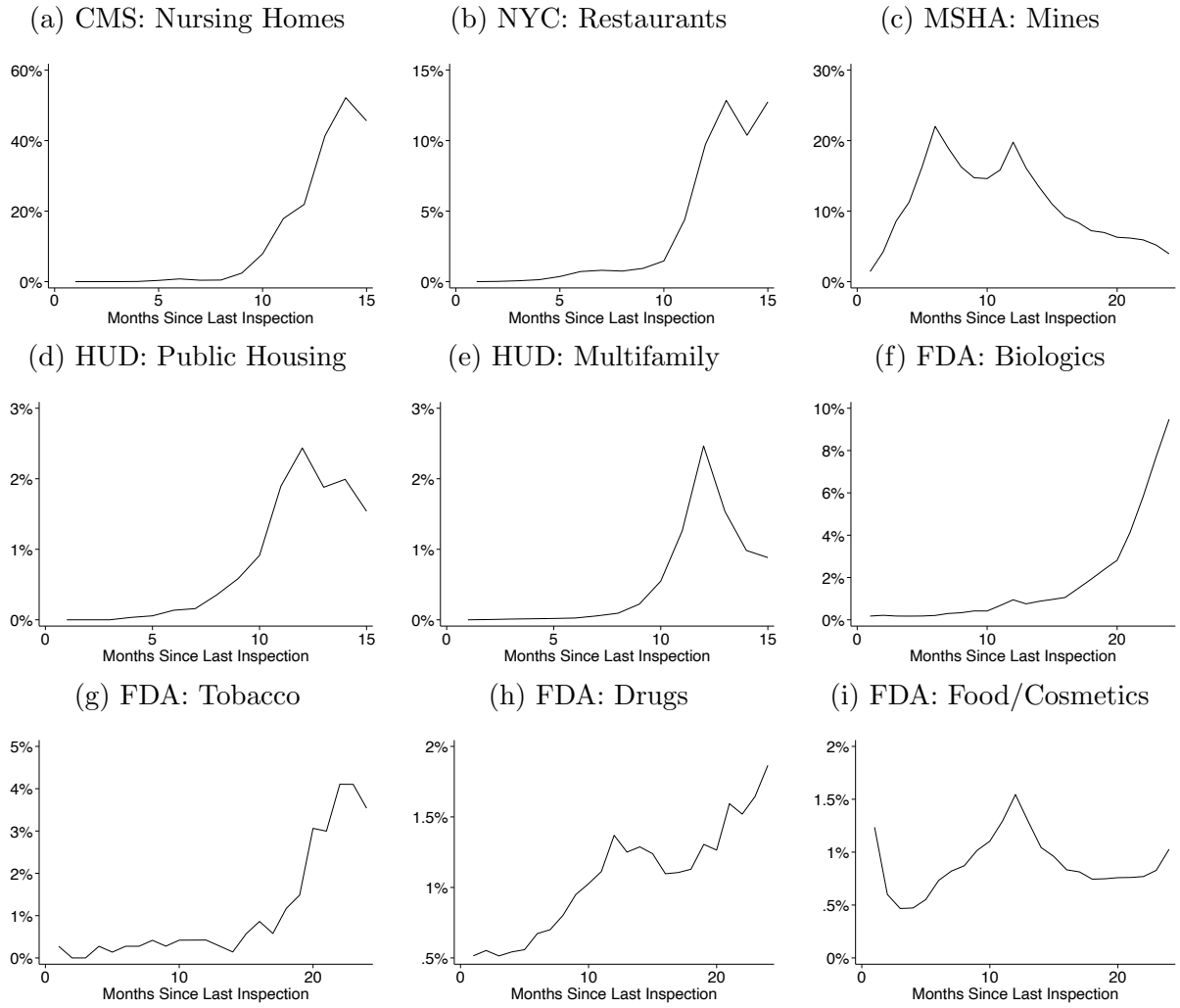
Whether inspections should be predictable is theoretically ambiguous (Varas et al., 2020; Ball and Knoepfle, 2023). The benefits of unpredictable inspections depend on the extent to which compliance responds to inspection risk, while the costs depend on how much information is lost due to irregular inspection timing. The question of how predictable inspections should be is therefore empirical and depends on the magnitudes of these costs and benefits.

This paper evaluates this tradeoff empirically in a high-stakes regulatory setting: the U.S. nursing home industry, where facilities undergo unannounced inspections approximately annually. Using high-frequency administrative data, we first examine how facilities' effort and patient health dynamically respond to inspection risk. These empirical patterns then inform a dynamic structural model that captures both the incentive and informational roles of inspections within a unified framework. We use this model to assess how changing features of the inspection regime—such as the frequency or predictability of inspections—would affect facilities' effort and patients' health outcomes.

Our main empirical analysis centers on the dynamic relationship between inspection hazard, facility effort, and patient health in the period *between* routine health inspections. We first confirm that inspection timing is distinctly cyclical: following an inspection, the hazard rate for the next inspection is initially near-zero but begins to rise sharply starting after approximately 35 weeks until the next inspection occurs.

To quantify effort, we construct a composite index from daily payroll-based staffing measures and time-varying measures of care inputs derived from patient assessments. We find that facilities exert low effort when inspection risk is minimal and ramp up effort later in the

Figure 1: Inspection Hazard Rates Across Regulatory Settings



Notes: This figure plots inspection hazard rates across different regulatory settings. They correspond to (a) nursing home inspections conducted on behalf of CMS (2010-2023), (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 (d) public housing properties and (e) multifamily assisted properties by the Department of Housing and Urban Development (2018-2021), and inspections of manufacturers of (f) biologics, (g) tobacco goods, (h) drugs, and (i) food and cosmetics by the Food and Drug Administration (2009-2024).

cycle as the inspection hazard increases. Importantly, the empirical relationship between inspection risk and effort is concave, suggesting that smoothing inspection risk over time may prompt higher average effort.

Our primary health outcome of interest is patient mortality. We find that patients are more likely to die during the periods in which inspections are unlikely and effort is low. However, survival rates improve as effort increases later in the inspection cycle. The covariation between effort and mortality over the inspection cycle implies that a one standard-deviation increase in facility effort lowers patient mortality by 5 percent. Balance tests indicate these effects are not driven by changes in patient composition over the inspection cycle. Furthermore, we find that changes in effort over the inspection cycle affect mortality even when instrumenting for choice of facility using a distance-based demand model.

Guided by our empirical evidence, we develop a dynamic model of nursing homes' strategic effort provision. This model allows us to simulate how effort and health outcomes would change under alternative inspection regimes. In each period, the facility balances the costs of exerting effort today against the intertemporal benefits of additional effort should an unannounced inspection occur in the near future. Under this model, facilities develop an effort strategy based on the anticipated arrival process of inspections.

But inspections do not exist solely to incentivize effort: they also inform the principal about agents' time-varying latent quality. The Centers for Medicare & Medicaid Services (CMS) uses inspections to monitor facilities and publishes recent inspection results to inform consumers about facility quality. The timing of inspections dictates how regularly such information is updated. We therefore model each inspection as generating signals about facility quality, the informativeness of which gradually decays over time. This characterizes the information gained from each inspection and how regulators' uncertainty about a facility's quality increases over the inspection cycle.

We use our model to simulate the impact of inspections on effort, health outcomes, and uncertainty about quality. We first benchmark the current inspection regime against the extreme counterfactual of eliminating nursing home inspections entirely. We find that the effort currently induced by the threat of inspections saves 850.6 lives per year (54.9 lives per 1,000 inspections) and reduces the regulator's average uncertainty about latent quality by 56.9%. Increasing the frequency of inspections by 25% saves 25.3% more lives and reduces uncertainty by an additional 2.8 percentage points. However, this may be costly or infeasible given reports that health departments are struggling to perform enough inspections to meet even the current standard for inspection frequency (U.S. Senate, 2023).

In contrast, simply making the timing of inspections unpredictable is budget-neutral. We find that doing so would increase the number of lives saved by 12.2%, which is an equiv-

alent benefit to increasing the frequency of inspections by 12%. However, unpredictability comes at an informational cost: in the average week, regulators would have 2.6% less information about facilities' quality than under the current regime.[1] Furthermore, we find that unpredictability and frequency are complementary: increasing the frequency of inspections is 57.8% more effective at saving lives when inspections are unpredictable than under the current level of predictability. These findings highlight the predictability of inspections as an important but underutilized lever for policymakers. They suggest that optimizing *when* inspections occur may be as consequential as determining *how many* inspections to conduct.

While an extensive empirical literature has demonstrated that increasing the frequency (Olken, 2007; Levine et al., 2012; Blundell et al., 2020; Shi, 2024; Lin et al., 2025), targeting (Duflo et al., 2018; Johnson et al., 2023), and salience (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 another key dimension of the design of inspection policies: their predictability. Recent theoretical work has highlighted the important tradeoff that unpredictability creates between incentives and information (Varas et al., 2020; Ball and Knoepfle, 2023), but no empirical work has quantified this. The main challenge in empirically studying inspection predictability is that compliance is typically only observed during inspections.[2] We are able to overcome this in the nursing home context, where high-frequency administrative data allow us to measure firms' effort and compliance between inspections.

This paper also contributes to our understanding of inspections as a widely-used tool in health care, including for hospitals, home health agencies, dialysis facilities, clinical laboratories, and hospices (CMS, 2024b). Existing work on inspections has largely focused on their overall effects (Alsan and Yang, 2025) or their frequency (Lin et al., 2025). Papers considering strategic responses have focused primarily on window-dressing in the period during or immediately surrounding the inspection (Barnett et al., 2017; Chen and Dillender, 2025). In contrast, our paper examines strategic responses over the long window between inspections and considers how these might change under alternative inspection regimes.

In the healthcare setting, inspections are just one of many monitoring tools—including insurer scrutiny, denials, and prior authorization—that are used to either reduce costs or improve quality of care (Dillender, 2018; League, 2022; Macambira et al., 2022; Brot-Goldberg et al., 2023; Leder-Luis, 2023; Eliason et al., 2024; Gandhi and Shi, 2025). Our findings suggest that reducing the predictability of these tools may improve their efficacy without incurring the costs of employing them more frequently (U.S. Senate, 2023; Dunn et al., 2024).

[1] In contrast, perfectly predictable inspections occurring at regularly scheduled intervals would decrease the number of lives saved by 16.5% but would increase the information regulators have by <1%.

[2] The notable exception is Zou (2021), which uses satellite-based measures of air quality to identify noncompliance on unmonitored days under the Clean Air Act.

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; Gupta et al., 2024; Gandhi et al., 2025; Gandhi and Olenski, 2025), and reimbursement design (Grabowski, 2001; Hackmann, 2019; Gandhi et al., 2024). While the public report card system used to provide information about nursing home quality has received considerable attention (Grabowski and Town, 2011; Werner et al., 2012; Cheng, 2023; Konetzka et al., 2021), the inspection process that generates this key input remains relatively under-studied, with contemporaneous work just beginning to fill this gap (Chen and Dillender, 2025; Lin et al., 2025). We contribute to this literature by examining the inspection process, an important yet under-studied component of the regulatory environment, with a particular focus on the timing of when inspections happen.

2 Background

2.1 Nursing Home Industry

This paper considers staffing levels, patient outcomes, and inspections in skilled nursing facilities, colloquially referred to as nursing homes. The nursing home industry is large, accounting for approximately one percent of U.S. gross domestic product, and is composed of a mix of for-profit and non-profit operators. Nursing homes provide a continuum of care that spans two broad categories. First, they offer post-acute, rehabilitative services for individuals transitioning from a hospital stay. Second, they deliver long-term custodial care for residents who require daily assistance with activities of daily living (ADLs)—e.g., bathing, eating, and toileting—over an extended period.

The hallmark issue facing the nursing home industry is the pervasiveness of low-quality care. Patients routinely face harm indicative of abuse and neglect. Avoidable falls, pressure ulcers ("bed sores"), and inappropriate chemical or physical restraints are such common occurrences that the government reports the relative frequency of these incidents as part of facilities' public report cards. Such harm is commonly attributed to inappropriately low levels of staffing: most facilities are under-staffed relative to clinical benchmarks and even statutory minimums (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).

Information on nursing homes' inputs and outputs is uniquely rich, even within the

healthcare sector. 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. On the output side, all patients, regardless of payer, must receive detailed clinical assessments at admission, at discharge, and at regular intervals in between. Additionally, since almost all nursing home patients are Medicare enrollees, their mortality is easily observed in Medicare data, even when death occurs outside of the facility.

2.2 Nursing Home Inspections

Inspection Process. Nearly all facilities are certified by CMS to provide care and receive public reimbursement through Medicare and Medicaid. As part of this certification, CMS mandates approximately annual inspections for facilities. These inspections serve two primary purposes: they monitor and report on nursing home quality, and they enforce compliance by identifying and correcting quality deficiencies.

These inspections—formally referred to as "surveys"—are conducted by state survey agencies. While states have some discretion in scheduling and logistics, they must adhere to federal guidelines governing what to evaluate and how often. State survey agencies are required to conduct an unannounced inspection of each facility at least every 15 months and roughly once per year on average (U.S. Congress, 1987).[3] Empirically, the gap between inspections of the same facility is typically around a year: Figure H.1 shows that 74 percent of inspections occur between 40 and 60 weeks after the previous inspection.

During an inspection, the state survey agency sends a team of inspectors on-site, who spend on average 4.9 days reviewing the facility's practices, interviewing residents and staff, and observing conditions to identify any deficiencies in care or safety (Figure H.2). The teams are typically comprised of up to five trained inspectors, including at least one registered nurse. The surveys begin with an initial "entrance conference" with facility personnel, which is followed by a general tour of the facility, selection of residents for in-depth interviews, and investigations into particular areas. Inspectors are trained to follow a complex set of guidelines—the operations manual for conducting surveys is almost 900 pages—on how to identify violations in areas ranging from resident quality of life, food and nutrition services, infection control measures, and safety of the physical environment (CMS, 2024a). Inspectors take detailed notes of any observations they make, interviews they conduct, and conclusions they draw. Once the inspection has concluded, the survey team submits their findings to CMS.

[3] In addition to "standard" unannounced inspections (and revisits related to them), state survey agencies also conduct *ad hoc* complaint investigations to respond to individual allegations of quality deficiencies, which are submitted by residents or their families.

Inspection Outcomes and Implications. The key outcome of the inspection is the "Statement of Deficiencies," which is a list of all violations identified during the inspection. Inspectors can cite facilities for over 200 possible violations. Each deficiency cited is rated by scope and severity, which can range from minor isolated issues to widespread harm (CMS, 2023, 2025b). Between 2013 and 2019, 84.9 percent of inspections found at least one deficiency with potential for minimal harm or more (Figure H.3), and the average number of deficiencies identified per inspection is 5.4.

Regulators use the inspection findings as an enforcement tool to ensure facilities are in compliance with the stated guidelines. Facilities cited for deficiencies must submit a "Plan of Correction" outlining the specific corrective actions they will take to address each deficiency, and the date by which the deficiency will be resolved. Inspectors then conduct up to three "revisits" to verify compliance and that all deficiencies have been corrected. The first two revisits typically occur within 12 weeks of the initial inspection, and the third revisit must occur within 24 weeks. In practice, 99.4% of inspection cycles have completed their last revisit within 20 weeks of the initial inspection (Figure H.4).

Facilities that fail to address all deficiencies within 6 months of their inspection risk termination of their CMS contracts. Additionally, depending on the severity and urgency of the deficiencies, initial inspection findings may also trigger more serious enforcement actions at the discretion of the state or CMS. For example, facilities can be added to the Special Focus Facility (SFF) program, during which they are subject to additional scrutiny and threats to their license (CMS, 2008).[4]

Inspections also serve an informational purpose, as their findings are made publicly available for patients and families evaluating nursing homes. CMS publishes the Statement of Deficiencies, as well as a redacted version of the full notes, from each facility's most recent inspection on their online Care Compare tool (CMS, 2025a). The Care Compare website also reports a facility's "Five-Star Rating," which aggregates information from inspection results, staffing levels, and other quality measures (CMS, 2025c,e). When constructing this composite rating, inspection results receive the greatest weight, owing to CMS's belief that they provide the strongest signal of facility quality. Previous literature has shown that these ratings steer at least some patient demand (Konetzka et al., 2015; Cheng, 2023).

For these reasons, nursing home operators are keenly aware of the inspection process and are widely believed to adjust their behavior in anticipation of new inspections. 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 Advo-

[4]Other enforcement actions include civil monetary fines, denial of Medicare or Medicaid payments for new admissions, and the transfer of existing residents to other facilities.

cacy, 2023), a phenomenon also reported in popular news media (Thomas, 2014; Rau, 2018). While it is unlikely that operators are obtaining information about the exact date of the next inspection months in advance, the fact that inspections occur on relatively regular intervals makes their timing partially predictable. See Section B.1 for evidence that operators are responding to general inspection risk rather than receiving tip-offs about exact dates.

3 Data and Sample Construction

3.1 Data

Nursing Home Inspections. Our primary source for data on inspections is the Survey Summary files from the CMS online data repository. From these, we obtain inspection end date, inspection type (standard, revisit, or complaint), and a list of all cited deficiencies for all 2006-2019 inspections.[5] Table G.1 describes the 20 most common deficiencies (out of 205 in total), along with their relative frequencies. Common deficiencies include issues pertaining to food sanitation, infection control, resident safety, and low-quality care.

Staffing. We measure staffing using the Payroll-Based Journal (PBJ) from 2017 through 2019. The PBJ is an administrative database containing records for all direct care staff in all CMS-certified nursing homes. These data detail the daily shifts of each direct care worker, including hours worked, level of nursing certification, employee or contractor status, and administrator status. Because PBJ data derive from payroll records, they are considered more reliable than self-reported staffing measures (Geng et al., 2019; Gandhi et al., 2021).

Patient Characteristics, Care, and Health Outcomes. Our analysis of patient health outcomes relies primarily on mortality reported in from the 2013-2019 Medicare Master Beneficiary Summary File (MBSF) enrollment files, which include patient demographics, chronic conditions, home zip codes, and death dates for all Medicare beneficiaries enrolled in traditional Medicare or Medicare Advantage.

To identify nursing home use, we leverage the 2013-2019 Minimum Data Set (MDS), which captures regular assessments from all CMS-certified nursing homes. The MDS provides 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 classify the MDS measures into patient care *inputs*—such as the share of residents receiving antipsychotics—which capture nursing

[5]For completeness, we supplement with data from the Quality, Certification, and Oversight Reports (QCOR) website, which are used by state survey agencies and include a small number of surveys not found in the Survey Summary files. However, QCOR data are only available starting 2010. We also obtained inspection *start* dates and data on inspection revisits in 2017-2019 from Freedom of Information Act requests.

home treatment decisions, and patient health *outcomes*—such as measures of activities of daily living (ADLs)—which reflect the downstream consequences of those inputs.

Finally, we use the 2013-2019 Medicare Provider Analysis and Review (MedPAR) files, which contain information on all Medicare-covered hospital stays, to incorporate measures from recent hospitalizations in our analyses.

Samples. We construct analysis samples for three time periods: (i) the "Facility Effort" sample spans 2017-2019, as the PBJ staffing data are only available in these years; (ii) the "Patient Care" sample spans 2013-2019 to match our sample of patient data; and (iii) the "Inspection Outcomes" sample spans 2006-2019, which includes all years for which we have inspection data. Table G.2 reports summary statistics for each sample. Section A.1 details additional sample restrictions and how we construct a panel of nursing home stays.

3.2 Composite Effort Index

A key input into our analysis is the responsiveness of nursing home "effort" to inspection incentives. We conceptualize the effort level as a univariate input choice which varies from week to week and is costly for the nursing home to increase. In practice, there are many operational routines that change across the inspection cycle which may impact patient health outcomes. Facilities may retain more qualified or experienced staff, those staff may adhere more or less strictly to care protocols, management practices may adjust, and so on. Though not all of these dimensions of effort are observable, changes in these routines are likely to be well-approximated by both the labor and non-labor inputs we do observe.

We combine these multiple nursing home inputs into a single composite effort index using factor analysis (Bartholomew et al., 2011). Let facility j's choice of input choice k (such as the overall staff/resident ratio) in week t be x_{jt}^k. Each observed input k is a noisy proxy of latent nursing home effort e_{jt}:

$$x_{jt}^k = \lambda_k e_{jt} + \varepsilon_{jt}^k, \tag{1}$$

where λ_k is a factor loading representing how strongly the observed measure reflects latent effort. The ε_{jt}^k terms represent idiosyncratic errors that are assumed to be uncorrelated across k, and thus the only covariance across observed measures is driven by latent effort e_{jt}. Conceptually, the objective is to compute e_{jt} by estimating the factor loadings λ_k and inverting equation (1). Section A.2 provides an overview of factor analysis and the estimation procedure in further detail.

We construct the effort index using several observed input measures, normalized so that higher values indicate greater effort. Staffing variables — all measured in hours of direct care per resident-day — include total, registered nurse (RN), administrator, full-time, overtime,

and weekend hours. We also include an indicator for the use of contract workers and two non-labor inputs: the share of patients not being physically restrained and the share not receiving antipsychotics (a common form of "chemical restraint"). Table A.1 reports the estimated factor loadings and correlations for each measure. Total, weekend, full-time, and RN staffing are most strongly related to latent effort, as is less-frequent use of antipsychotics. Physical restraints exhibit little correlation with effort.

We standardize the resulting effort measure using the within-facility standard deviation calculated during the middle of their inspection cycles — specifically, weeks 21 to 30 following the prior inspection. This "flat" period corresponds to the interval between the end of the revisit window and the beginning of the next inspection window. We allow the means to vary across cycles. Accordingly, our effort estimates can be interpreted as standard deviations away from each facility's average mid-cycle effort level, when inspection risk is minimal.

4 Inspection Cycle Analysis

Our empirical analyses consist of several specifications, each relying on variation in inspection risk over time. In this section, we consider our primary specification, which studies how outcomes evolve throughout the inspection cycle, i.e., the period between inspections.

4.1 Research Design

Our primary empirical specification for outcomes measured at the facility-level is summarized by equation (2), which models an outcome of interest for facility j in calendar week t:

$$Y_{jt} = \sum_{w \in [0,30)}^{(30,60]} \beta_w \, \mathbb{1}(\text{Week Since Inspection}_{jt} = w) + \alpha_j + \lambda_t + \varepsilon_{jt}, \tag{2}$$

where $\mathbb{1}(\text{Week Since Inspection}_{jt} = w)$ is an indicator for week w of an inspection cycle, i.e., whether calendar week t is w weeks after facility j's most recent inspection. Y_{jt} is a facility-level outcome, such as the effort index or one of its inputs. We include facility fixed effects α_j and calendar-week fixed effects λ_t to account for time-invariant differences in outcomes across providers and aggregate fluctuations across time, respectively. Because the omitted week from inspection is $w = 30$, the parameters of interest—the β_w coefficients—can be interpreted as the difference in the outcome relative to the middle of an inspection cycle.

We also consider a patient-level analogue, given by equation (3):

$$Y_{it} = \sum_{w \in [0,30)}^{(30,60]} \beta_w \, \mathbb{1}(\text{Week Since Inspection}_{j(i),t} = w) + \gamma X_i + \alpha_{j(i)} + \lambda_t + \varepsilon_{it}, \tag{3}$$

where Y_{it} is a patient-week level measure such as 4-week survival, and $j(i)$ denotes the facility of patient i. The only substantive modification from equation (2) is the inclusion of a patient-level control X_i for predicted survival based on baseline patient characteristics, which helps account for any changes in sample composition over the inspection cycle. Section B.2 describes the construction of the predicted survival measure in more detail. In all analyses, we compute robust standard errors clustered at the facility-level.

4.2 Effort and Health Over the Inspection Cycle

Main Results. Figure 2 documents the findings of our inspection cycle analysis on our two primary outcomes: nursing home effort and 4-week patient survival. Panel (a) plots the coefficients on effort by weeks since last inspection along with the inspection hazard — the probability of receiving the next inspection that week, conditional on having not received one up to that point in the cycle. The inspection hazard is relatively flat up until week 35 and then subsequently rises rapidly, reflecting the predictability of the inspection cycle.

Panel (a) shows that, after a short-lived spike in effort during the inspection, effort follows a "U-shaped" pattern throughout the rest of the cycle.[6] For about 20 weeks after the inspection, it remains elevated but is declining. This aligns with the period of the inspection cycle in which nursing homes are subject to revisits following up on any deficiencies identified in the initial inspection.[7] We denote the first 20 weeks after the inspection as the "potential revisit window," as almost all revisits occur in this period (Figure H.4). Effort then remains flat in the middle of the cycle, where there is minimal inspection risk. Eventually, as the inspection hazard increases and the next inspection becomes imminent, effort gradually rises again, until a new inspection occurs and the cycle repeats itself.

Figure A.1 decomposes this pattern and plots each component of the effort index separately. All components with non-trivial factor loadings exhibit a similar "U-shape," indicating that the cyclical pattern we observe is common to the factors contributing to our effort measure. We also confirm that our results are robust to using *across*-facility (rather than within-facility) standard deviations (Figure H.5) and using a simpler effort measure which considers only variation in total staffing hours, unadjusted by patient census (Figure H.6).

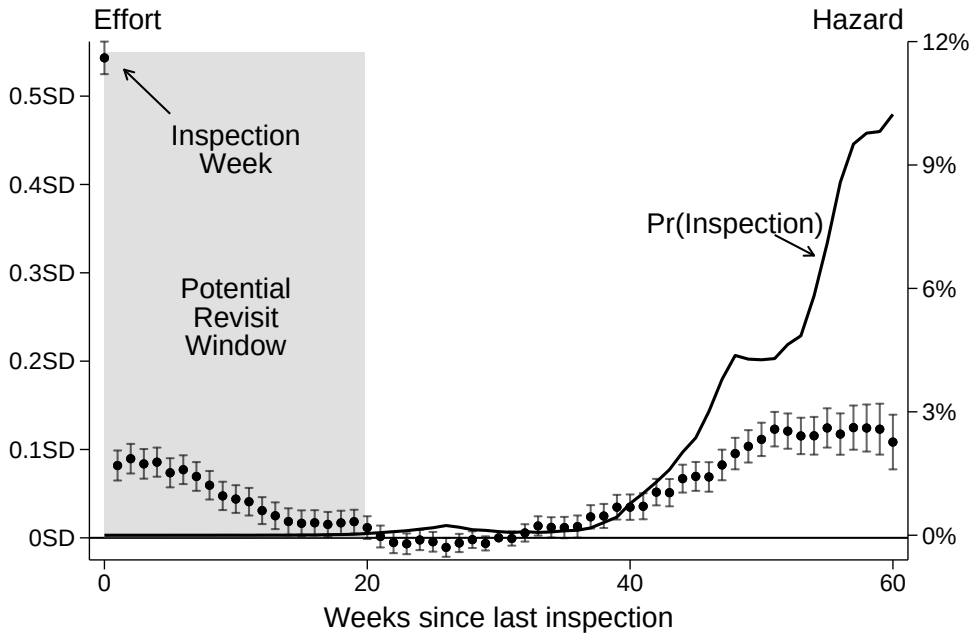
Panel (b) plots the estimates from our patient-level specification using 4-week survival as the dependent variable. The pattern of the coefficients also follows a "U-shape" similar to the effort coefficients: survival is elevated in the weeks immediately following an inspection, dips in the middle of the cycle, and then rises again as the next inspection approaches.

[6] Given that the large but transient disruption during the week of the inspection is not the focus of our paper, we estimate but omit the coefficient for week 0 in all analyses after Figure 2.

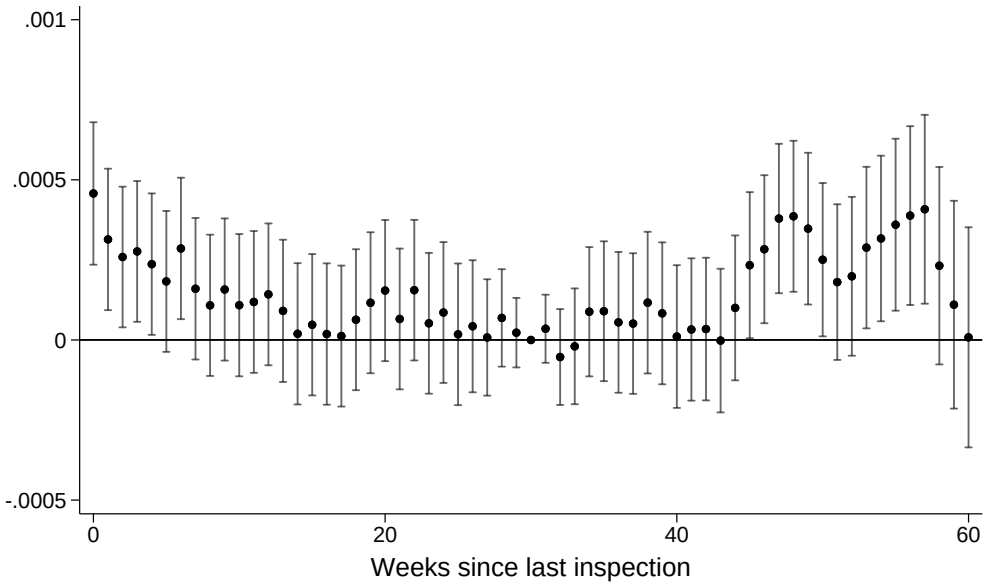
[7] For more detailed discussions of effort during the inspection and the revisit period, see Section B.3.

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

(a) Inspection Hazard and Facility Effort (Std. Dev.)



(b) Patient 4-Week Survival



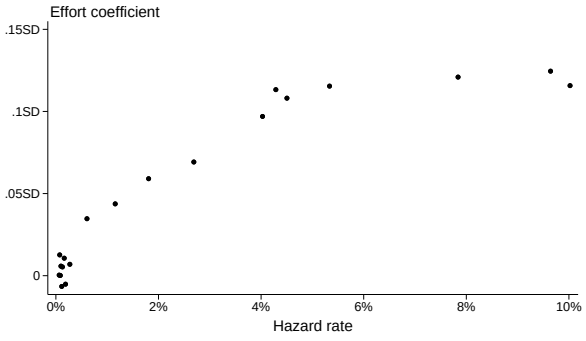
Notes: Panel (a) plots the inspection hazards (right axis) and the coefficients on from estimating equation (2) on within-facility standard deviations of effort (left axis). Panel (b) plots the coefficients from estimating equation (3) on 4-week patient survival rate, controlling for predicted patient mortality at admission. The effort index is a composite measure combining labor and non-labor inputs, and is described in further detail in Section 3.1 and Appendix Section A.2.

Heterogeneity. Given the differences in care needs between post-acute and long-term care patients detailed in Section 2, we examine heterogeneity in survival effects by "short-stay" (those within 90 days of admission, whose care often consists primarily of post-acute rehabilitation) and "long-stay" (the remainder, whose care is primarily custodial). Short-stay patients—who account for 77.9% of stays—drive the overall survival effects, and we find only minimal changes in survival across the inspection cycle for long-stay patients (Figure B.5). 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 (Figure H.7).

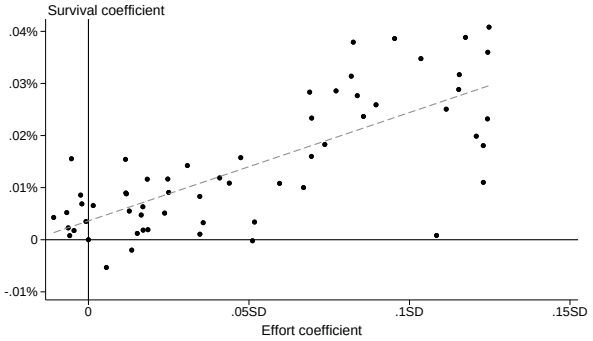
Section B.4 presents our remaining heterogeneity analyses, which include heterogeneity in effort and survival by facility type and patient characteristics. Most importantly, facilities in states with shorter inspection cycles increase effort more quickly than those in states that have longer cycles. This demonstrates that it is indeed the cyclicality of inspection risk driving the patterns in effort that we observe. For-profit facilities also exhibit greater responsiveness to inspection risk than non-profit facilities, and correspondingly show more pronounced cyclicality in patient survival. Survival is most responsive to inspection-induced effort for patients with low predicted survival rates and for patients with a recent hospital stay, consistent with survival effects being driven by patients with high baseline mortality. Differences across remaining facility and patient characteristics are relatively modest.

Figure 3: Comparison of Inspection Hazard, Effort, and Survival Coefficients

(a) Inspection Hazard vs. Effort Coefficients



(b) Effort vs. Survival Rate Coefficients



Notes: Panel (a) plots the relationship between inspection hazard and the effort coefficients in Figure 2. Panel (b) plots the relationship between the effort and survival coefficients in Figure 2. Coefficients are derived from estimating equations (2) and (3). The slope of the line of best fit in panel (b) is 0.002 (SE 0.0002). Coefficients for panel (a) are restricted to weeks from inspection $w \geq 21$. The effort index is a composite measure combining labor and non-labor inputs, and is described in further detail in Section 3.1 and Appendix Section A.2.

Relationship Between Effort, Inspection Risk, and Survival. Figure 3 plots the relationship between the inspection hazard rate and the effort coefficients for each week in

the inspection cycle, as well as the relationship between the effort and survival coefficients. Panel (a) shows a concave relationship between inspection hazard and effort, consistent with convex costs or concave returns to increasing effort. Panel (b) then shows a linear relationship between survival and effort with a slope of 0.002 (SE: 0.0002), indicating that a 1 standard deviation increase in effort is associated with 5 percent lower patient mortality rate.[8] Taken together, these two relationships imply a concave relationship between inspection hazard and survival. This suggests that smoothing inspection risk over time—i.e., making inspections less predictable, holding fixed their frequency—should improve overall survival rates.

Additional Health Outcomes. Given the richness of the MDS data, in Appendix B.5 we examine a variety of intermediate, sub-mortality patient health outcomes. For example, we test for cyclicality in the activities of daily living (ADLs), which provide measures of patient independence with activities such as eating or toileting, and serve as key quality-of-life measures. As with patient survival and facility effort, we find "U-shaped" patterns across the inspection cycle when examining ADLs, either in composite form or separately by activity. Periods of high effort correspond to improvement in each of these measures (Figure B.9). We find similar, though more muted, patterns when assessing the rates of pressure ulcers (Figure B.10), a common marker of inadequate nursing home care.

Alternative Event Study Approach One potential concern is that facility composition may change across the inspection cycle. For example, a facility that was always inspected within 50 weeks can never contribute to estimates of β_w for $w > 50$. While the facility fixed effects α_j should capture any time-invariant, cross-facility differences, we nevertheless explore how an alternative stacked event study research design would impact our results in Section B.6. Reassuringly, we recover estimates that are very similar to our primary specification.

4.3 Assessing Selective Admissions and Discharges

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. While this is consistent with a causal relationship between inspection-induced effort and improved patient health, it is theoretically possible that our findings could be driven by selective admissions (Gandhi, 2023) or discharges (Hackmann et al., 2024) along dimensions that are uncorrelated with our control for predicted survival. We conduct several exercises to assess the plausibility and magnitude of this potential bias.

First, we examine how a facility's overall admission and discharge decisions change over the inspection cycle in Figure H.9. We find evidence of only very modest fluctuations—there

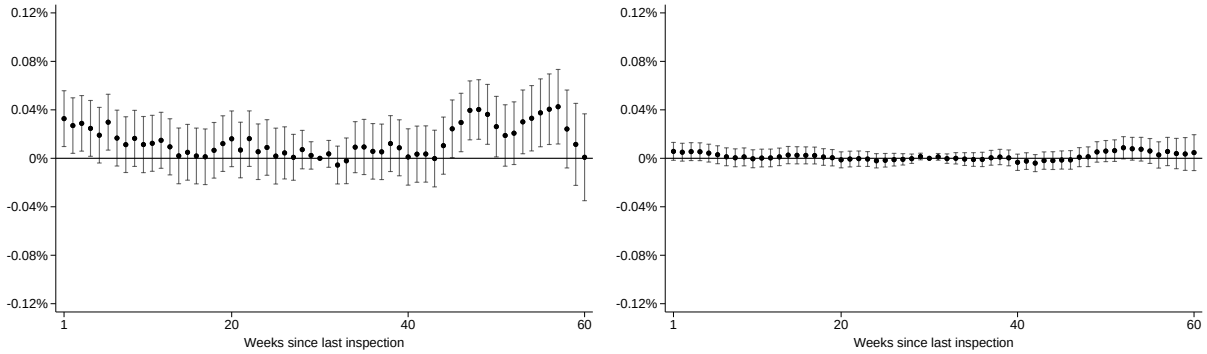
[8] We find very similar results when excluding the revisit window in Figure H.8.

are minor reductions in new admissions and discharges without anticipated return in the revisit period and leading up to the next inspection. For discharges with an anticipated return (i.e., patient is discharged to the hospital but anticipated to return to the nursing home), we find no evidence of systematic changes across the inspection cycle.

Figure 4: Actual and Predicted Survival Rates by Weeks Since Last Inspection

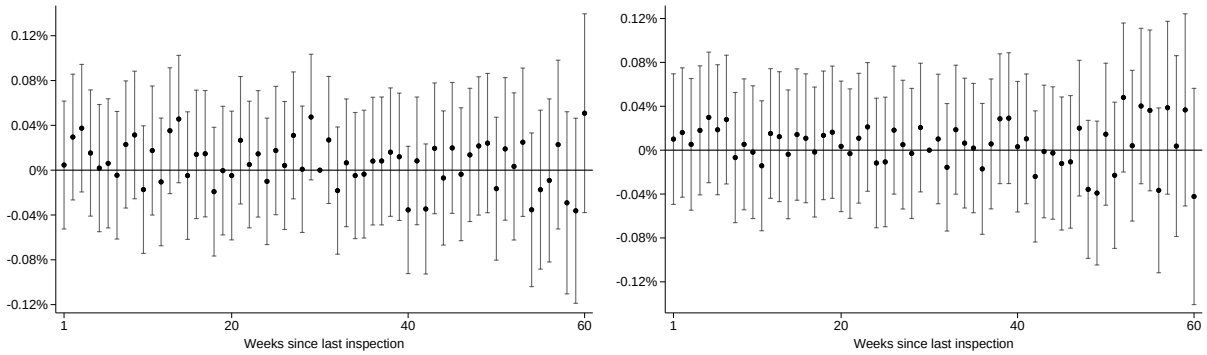
(a) Actual Survival, All Patients

(b) Predicted Survival, All Patients



(c) Predicted Survival, Admissions

(d) Predicted Survival, Discharges



Notes: This figure plots the coefficients from estimating equation (3) on actual and predicted survival rates. Panel (a) recreates the plot from Figure 2 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 with no return anticipated. All estimates are normalized in percent terms relative to the average of the outcome for all patients in the omitted week. The coefficients for week 0 are estimated but not reported.

Second, to understand whether these modest fluctuations in admissions and discharges 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 from equation (3): predicted survival X_i. Specifically, we examine how this measure evolves across the inspection cycle, both overall as well as specifically for new admissions

and discharges each week. Crucially, this measure relies only on *pre-admission* information observable to the facility and is not influenced by any direct effect of the stay itself on their health. We find no detectable changes in predicted survival throughout the inspection cycle. We can rule out changes in predicted survival relative to baseline larger than 0.02% among all patients in the nursing home, 0.14% among admissions, and 0.14% among discharges (Figure 4). Consistent with these results, we find that excluding the predicted survival control X_i generates estimates that are highly similar to our main specification (Figure H.10). While these results do not directly rule out systematic selection on unobservables over the inspection cycle, it seems implausible that such selection on unobservables would be particularly strong when selection on observables appears to be quite weak.

5 Zip-Code Level Approach

The exercises in Section 4.3 demonstrate that observable patient risk does not vary across the inspection cycle. To further assess the robustness of our estimates to endogenous *unobserved* patient selection—that is, selection that is uncorrelated with observable characteristics, but nevertheless impacts survival—we next develop a zip-code level approach that is agnostic to the exact facility where patients received care. Specifically, our approach leverages patients' *expected* exposure to cycle-induced variation in survival, based only on where they live and when their stay occurs.

This approach builds upon standard distance-based instruments used to address selective admissions in healthcare (McClellan et al., 1994; Einav et al., 2022; Olenski and Sacher, 2024), which leverage patients' well-documented preference for nearby providers. One key difference is that our measure varies *within* zip code over time, which allows us to control for baseline geographic differences using zip code fixed effects. Our comparisons are thus among patients in the same zip code whose exposure to cycle-induced variation in survival is determined by how recently nearby facilities were inspected when the patient needs nursing home care.

Distance-Based Model of Facility Choice. We specify a distance-based facility discrete choice model similar to Einav et al. (2022). Patient i with home zip code $h(i)$ receives utility from choosing nursing home j given by:

$$u_{ij} = \mu \log d_{h(i),j} + \epsilon_{ij}, \tag{4}$$

where $\log d_{h(i),j}$ is the log-distance between the centroid of patient i's home zip code and nursing home j, and the error term ϵ_{ij} follows the type-I extreme value distribution. This yields the familiar logit choice probability:

$$p_{ij} = \frac{\exp(\mu \log d_{h(i),j})}{\sum_{k \in C_{h(i)}} \exp(\mu \log d_{h(i),k})}, \tag{5}$$

where $C_{h(i)}$ is the choice set of nursing homes available to someone with home zip code h.[9]

Expected Survival Effect. Our analysis requires a measure of the expected change in i's survival in calendar week t due to the current cycle-induced survival effects of nearby facilities. To construct this, we combine our estimates from panel (b) of Figure 2 with the distance-based choice model from equation (5):

$$\underbrace{\widehat{z}_{it}}_{\substack{\text{Expected} \\ \text{Survival Effect}}} \equiv \sum_j \underbrace{\widehat{p}_{ij}}_{\substack{\text{Distance-Based} \\ \text{Choice Probability}}} \times \underbrace{\widehat{\beta}_{w(j,t),}}_{\substack{\text{Estimated} \\ \text{Survival Effect}}} \tag{6}$$

where $\widehat{\beta}_{w(j,t)}$ gives the survival effect based on cycle-week w that facility j is in at time t, and $\widehat{p}_{ij}$ is the distance-based probability that i receives care at j.

Note that $\widehat{z}_{it}$ does not condition on the actual facility where i receives care. However, we still find a strong positive relationship between $\widehat{z}_{it}$ and $\widehat{\beta}_{w(j(i),t)}$, indicating a strong predictive power of our distance-based choice model. Panel (a) of Figure 5 documents this: a simple regression of $\widehat{\beta}_{w(j(i),t)}$ on $\widehat{z}_{it}$ yields a coefficient of 0.86 and an R^2 of 0.15.

Regression Specification. We then estimate the following relationship between the expected survival effect and actual 4-week survival Y_{it}:

$$Y_{it} = \sum_{\tau=0}^{59} \delta_\tau \, \widehat{z}_{it} \times \mathbb{1}(\text{Week-of-Stay} = \tau) + \gamma X_i + \eta_\tau + \alpha_{h(i)} + \lambda_t + \varepsilon_{it}. \tag{7}$$

As before, we control for predicted patient mortality, X_i, and include calendar time fixed effects λ_t. In addition, we now include home zip code fixed effects $\alpha_{h(i)}$.[10] Notably, we do not include facility fixed effects to avoid conditioning on the facility where i received care.

Our coefficients of interest, δ_τ, are stratified by week-of-stay τ relative to the patient's admission. This approach reflects our earlier finding that patient mortality is most sensitive to the inspection cycle in the earliest weeks of a stay (Figure B.5). Accordingly, the relative magnitudes of δ_τ for different τ reflect when in a stay the health effects are largest. We include week-of-stay fixed effects η_τ to flexibly account for baseline differences in survival

[9] We define patients' choice sets as those facilities within 25 kilometers of their zip code's centroid.

[10] Figure H.12 plots the distribution of the expected survival effect with and without residualizing out home zip fixed effects. The residualized value captures within-zip variation driven by the timing of patient admissions and the inspection cycle. Reassuringly, there is still substantial variation in our instrument even after accounting for home zip fixed effects.

rates as residents progress through their stay. Importantly, $\widehat{z}_{it}$ varies over time for the same patient, as each week changes where local facilities are in their inspection cycles.

Identifying Assumption. The key identifying assumption underlying our distance-based instrument is that *within-zip-code, over-time* variation in unobserved patient risk is uncorrelated with the inspection cycle. Equivalently, changes in the composition of nursing home patients from a given zip code must be orthogonal to nearby facilities' inspection cycles. A potential violation would arise if, for example, all facilities in an area face the same inspection schedule and collectively adjust their admissions practices in response, such that the pool of nursing home patients from that zip code systematically covaries with the expected survival effect measure $\widehat{z}_{it}$. In this "aggregate cream-skimming" case, the zip code-level approach would be vulnerable to similar selection concerns as our main approach.

Such systematic variation seems implausible. Nevertheless, to assess this concern, we examine whether admissions at the zip code-level vary systematically with the inspection cycle. Figure H.11 plots the relationships between a zip code's expected survival effect and (a) the number of new admissions from that zip code that week, and (b) the predicted 90-day survival rate for patients from that zip code in a nursing home that week, after residualizing out zip code and calendar week fixed effects. Reassuringly, we find no evidence of a systematic relationship between either of these outcomes and the instrument.

Results. Figure 5 panel (b) plots the δ_τ coefficients from estimating equation (7) on 4-week survival. We restrict our focus to the coefficients for the first 13 weeks, which is the "short stay" period where we previously found the effect concentrated. The estimates indicate a positive relationship between a patient's survival and their expected survival effect during the first five weeks of a stay. These initial weeks have particularly high baseline mortality and represent a period during which most patients receive high-intensity post-acute care.

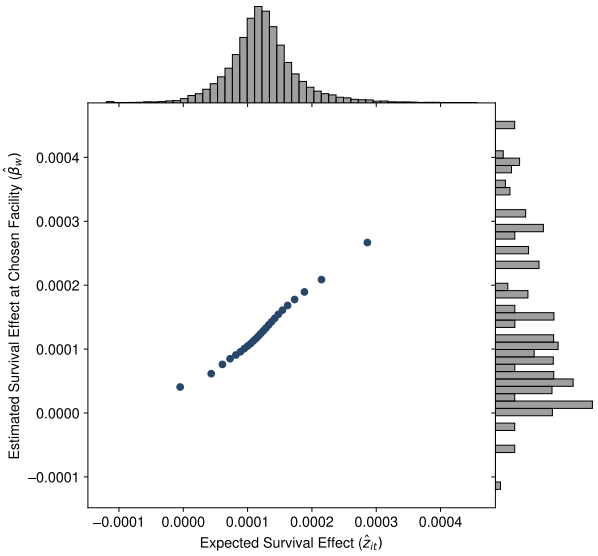
While the individual δ_τ are noisily estimated, in aggregate they provide a direct test of the assumption that unobserved patient selection does not vary systematically with the inspection cycle. The estimates capture the relationship between actual and expected survival, so under the null hypothesis of no unobserved selection, they should move one-for-one, on average. In contrast, an aggregate estimate close to zero—indicating no relationship between actual and expected survival—would suggest that the patterns in Figure 2 panel (b) reflect systematic patient selection on unobservable rather than inspection-induced improvements in care.

We formalize this by computing the weighted average $\bar{\delta} = 1$ and testing whether $\bar{\delta} = 1$.[11] Absent unobserved selection, a one percentage point increase in expected survival should

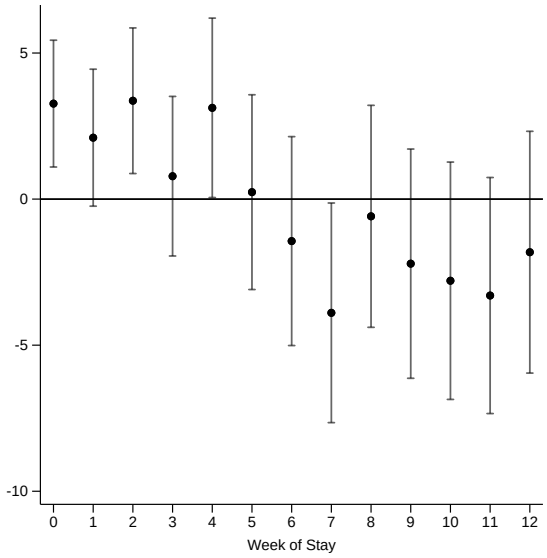
[11] Each δ_τ is weighted by the share of patient-week observations at week-of-stay τ. Early weeks receive greater weight because most stays are relatively short.

Figure 5: Zip-Code Level Approach

(a) Estimated vs. Expected Survival Effects



(b) Zip-Code-Level Survival Results



Notes: Panel (a) plots the relationship between the expected survival effect $\widehat{z}_{it}$ and the estimated survival effect of the facility-week the patient was actually admitted to, $\hat{\beta}_{w(j(i),t}$. The expected survival effect is defined in equation (6). The estimated survival effect for the chosen facility is assigned using the $\widehat{\beta}_w$ coefficients from equation (3). Panel (b) plots the coefficients from the specification in equation (7). The regression is estimated on weeks 0-59 from admission and the plot shows the estimates on the first 13 weeks of the stay. All coefficients past week 13 are statistically insignificant at the 5 percent level, with the exception of week 17, which is positive and statistically significant (coefficient: 5.32, SE: 2.22). We find that the weighted average $\bar{\delta} = 1.014$, which is indistinguishable from one ($p = 0.979$).

correspond to a unitary increase in actual survival. Indeed, we find $\bar{\delta} = 1.014$, virtually indistinguishable from 1 ($p = 0.979$), consistent with no role for unobserved selection in explaining our findings.

6 Dynamic Model of Nursing Home Inspections

In this section, we develop a dynamic model of nursing home inspections. First, in Section 6.1, we incorporate the role of inspections in incentivizing compliance by modeling nursing homes as strategically providing costly effort in anticipation of an inspection. Second, in Section 6.2, we capture the information-gathering role of inspections by modeling them as providing signals about facilities' time-varying latent quality. As time passes, the signal from the most recent inspection decays and becomes less informative.

6.1 Model of Nursing Home Effort

Inspection Cycles. An inspection cycle is the period after one inspection has occurred lasting until the next inspection arrives. Within a cycle k, we index time as weeks $w \in \mathbb{N}$ since the last inspection. At the end of each week, an unannounced inspection occurs with probability $h(w)$. This hazard rate weakly increases until it reaches a maximum on week $\overline{w}$, after which it is constant. When an inspection occurs, cycle k ends and the facility begins a new cycle.

A unique feature of our setting is that the initial weeks of an inspection cycle are spent in a "revisit window" during which the inspector may return to follow up on citations from the previous inspection. We have little data on or insights into the revisits process, and our model is not designed to characterize behavior during this period. We therefore treat behavior during the revisit window as fixed and model only behavior that occurs after the final revisit occurs on week τ_k. In other words, when modeling the strategic provision of effort, we essentially treat the "effective start" of the cycle as $w = \tau_k + 1$. See Sections 2.2 and B.3 for a more detailed discussion of revisits.

Target vs. Realized Effort. In each week w of inspection cycle k, facility j chooses a target level of effort $e^*_{j,k,w}$ that is not observed by the econometrician. Realized effort $e_{j,k,w}$ is drawn from a logistic distribution centered at $e^*_{j,k,w}$ with standard deviation σ_e. This stochastic wedge between target and realized effort reflects difficulties that facilities may face in precisely tuning effort at the weekly level, such as staff turnover, holidays, patient surges, and other operational frictions. From a practical perspective, this noise also helps rationalize empirical effort paths that are not perfectly smooth or monotone over the cycle.

Accumulated Effort and Inspection Performance. A facility's performance during an unannounced inspection depends on the cumulative effort expended over the current cycle. Specifically, the effort realized in each week contributes to an "effort stock" $s_{j,k,w}$ that determines facility j's expected performance $\pi(s_{j,k,w})$ should an inspection occur that week. This stock starts from zero in period τ_k and both accumulates and depreciates over the cycle:

$$s_{j,k,w} = (1-\delta)e_{j,k,w} + \delta s_{j,k,w-1} \tag{8}$$

where $\delta \in [0,1]$ measures persistence in the benefits of past effort. Note that since $s_{j,k,w}$ is a convex combination of $s_{j,k,w-1}$ and $e_{j,k,w}$, one can think of $e_{j,k,w}$ as an adjustment to the prior stock of effort $s_{j,k,w-1}$. We embed this intuition into our model of effort costs below.

Cost of Effort. Higher levels of effort contribute to costs in two ways. $e_{j,k,w}$ incurs a cost that has two components. The first component captures the potentially convex costs of

scaling effort $e_{j,k,w}$, which we model with a quadratic in $e_{j,k,w}$. This would encompass, for example, the amount spent on typical wages, salaries, and benefits for staff. The second is a quadratic adjustment cost that captures whether current effort deviates too greatly from recent efforts. For example, dramatically increasing staffing over a short period of time will likely entail costly recruitment, training, overtime hours, contract workers, and raising wages and benefits for retention. Likewise, rapid reductions in effort may also be costly and involve layoffs, cut shifts, and other difficult reductions. Thus, our cost function is as follows.

$$c(e_{j,k,w}, s_{j,k,w-1}) = \alpha_1 e_{j,k,w} + \alpha_2 e_{j,k,w}^2 + \alpha_3 (e_{j,k,w} - s_{j,k,w-1})^2, \tag{9}$$

where (α_1, α_2) and α_3 capture the level and adjustment costs of effort, respectively.

Dynamic Effort Optimization. Each week, facility j chooses its target effort aiming to maximize its expected discounted payoffs. Let $V_w(s_{j,k,w-1})$ be the facility's value of having stock $s_{j,k,w-1}$ at the start of week w in cycle k. After selecting target $e^*_{j,k,w}$, effort $e_{j,k,w}$ is realized, the facility pays cost $c(e_{j,k,w}, s_{j,k,w-1})$, and effort stock updates to $s_{j,k,w}$. Then, an inspection occurs with probability $h(w)$. If it doesn't occur, the period increments to week $w+1$. If it does occur, the facility earns $\pi(s_{j,k,w})$ for its performance and starts a new cycle k'. The facility does not know the length of the revisit window for k' and assumes that the effective start of the next cycle is drawn from the empirical distribution. The corresponding Bellman equation is:

$$\begin{aligned} V_w(s_{j,k,w-1}) = \max_{e^*_{j,k,w}} \int \Big[& - c(e_{j,k,w}, s_{j,k,w-1}) \\ & + h(w)\Big(\pi(s_{j,k,w}) + \beta \int V_{\tau_{k'}+1}(0) dP(\tau_{k'})\Big) \\ & + \big(1 - h(w)\big)\, \beta V_{w+1}(s_{j,k,w})\Big] dP(e_{j,k,w} \mid e^*_{j,k,w}) \end{aligned} \tag{10}$$

where $\beta < 1$ is the discount factor.[12] Importantly, when $w \geq \overline{w}$, the Bellman equation becomes stationary. Thus, we first solve the stationary Bellman at $w = \overline{w}$ and trivially apply backwards induction for $w < \overline{w}$.

6.2 Latent Quality, Effort, and Inspection Performance

We next specify how a facility's quality and effort determine its inspection performance. This serves two purposes. First, it defines the payoff function $\pi(s_{j,k,w})$ that enters the facility's dynamic problem. Second, it characterizes the information that a regulator gathers about quality from intermittent inspections.

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

In this section, we consider the evolution of a facility's latent quality over time, including across cycles. For this reason, it is convenient to work in calendar time t, rather than in terms of cycles k and week-within-cycle w. Of course, these are equivalent: for a given facility j, calendar time uniquely corresponds to a specific cycle k and week-within-cycle w. Towards this end, we slightly abuse notation by denoting $s_{j,t}$ to signify $s_{j,k(j,t),w(j,t)}$.

Quality. A facility's quality is a persistent latent state with a facility-specific intercept:

$$q_{j,t} = \varphi_j + \rho q_{j,t-1} + \epsilon_{jt} \tag{11}$$

where $\varphi_j \sim \mathcal{N}(0, \sigma_\varphi^2)$ captures static heterogeneity across facilities, $\rho \in (0,1)$ characterizes the persistence of quality, and $\epsilon_{jt} \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is an idiosyncratic weekly innovation.

Inspection Performance. Neither the regulator nor the econometrician observe quality $q_{j,t}$ directly. Instead, the regulator conducts periodic inspections in which she evaluates $D = 160$ potential deficiencies, indexed by $d = 1, \ldots, D$.[13] Let $x_{j,t}^d$ be an indicator for deficiency d being cited in an inspection of facility j at time t. We model each $x_{j,t}^d$ as a probit with probabilities determined by latent quality $q_{j,t}$ and accumulated effort $s_{j,t}$:

$$p_{j,t}^d = \Pr(x_{j,t}^d = 1 | q_{j,t}, s_{j,t}) = \Phi\big(\kappa_d + \gamma_d(q_{j,t} + \lambda s_{j,t})\big), \tag{12}$$

where κ_d governs the baseline prevalence of deficiency d and γ_d captures its sensitivity to latent quality and accumulated effort. We stack these across D as vectors $\boldsymbol{\kappa}$ and $\boldsymbol{\gamma}$, respectively. λ governs the substitutability between quality and accumulated effort.

We assume that facilities' effort choices in Section 6.1 are made without precise knowledge of current latent quality t. This means that their anticipated payoff from effort stock $s_{j,t}$ is determined by their expected number of deficiencies without conditioning on $q_{j,t}$:

$$\pi(s_{j,t}) = -\int \sum_{d=1}^{D} \Phi\big(\kappa_d + \gamma_d(q_{j,t} + \lambda s_{j,t})\big) dP(q). \tag{13}$$

Regulator's Uncertainty. Because inspections occur only intermittently, the regulator's information about quality is not always up to date. On a given week t, the regulator's belief about facility j's quality is based on the most recent inspection outcome—i.e., the vector of deficiencies $\mathbf{x}_{j,t-w(j,t)}$—and their knowledge of facility-specific intercept φ_j. Since the regulator learns φ_j in the long-run, we focus our analysis on uncertainty due to variation in quality over time. Specifically, we examine $\mathrm{Var}(q_{j,t} \mid \mathbf{x}_{j,t-w(j,t)}, \varphi_j)$, which increases as

[13] We exclude some rare deficiencies (1.1% of citations) for issues with sample size and convergence.

more time has elapsed since the most recent inspection.

7 Estimation Procedure

Estimation proceeds in the reverse order. In Section 7.1, we estimate the model parameters related to inspection performance and latent quality, as described in 6.2. Then, in Section 7.2, we estimate the parameters of the effort choice model described in 6.1.

7.1 Inspection Performance

In this section, we provide an overview of estimation that highlights key details. The full procedure is detailed in Appendix C.

Data. Estimation relies on the Statement of Deficiencies produced in each inspection, as described in Sections 2.2 and 3.1. These data record $\mathbf{x}_{j,t}$, the vector of D deficiency indicators summarizing the results of an inspection of facility j occurring in calendar week t. Because inspections occur only intermittently, $\mathbf{x}_{j,t}$ is defined only for the subset of weeks during which j is inspected. Since we are often interested in comparing consecutive inspections, we define $\Delta(j,t)$ to be the length of the cycle that ended with an inspection of facility j at time t.[14] In particular, then $t - \Delta(j,t)$ yields the date of the prior inspection. For convenience, we slightly abuse notation in simply writing Δ as a shorthand for $\Delta(j,t)$.

Moments. We estimate the inspection performance parameters $\{\boldsymbol{\kappa}, \boldsymbol{\gamma}, \lambda, \sigma_\varphi, \sigma_\epsilon, \rho\}$, as well as the effort accumulation rate δ, via the simulated method of moments. We include three sets of moments. First, we match the share of inspections with citations for deficiency d separately by cycle length Δ: $\mathbb{E}[x^d_{j,t} \mid \Delta] = \bar{x}^d_\Delta$, where $\bar{x}^d_\Delta$ is the empirical mean of $x^d_{j,t}$ for cycles of length Δ. Matching the empirical rates helps to pin down the baseline prevalence parameter κ_d for each deficiency. Doing so separately for different cycle lengths allows us to also identify λ, which relates accumulated effort to performance, since effort stock tends to be higher at the end of longer (i.e., larger Δ) inspection cycles. These moments are particularly important for ensuring that the model generates realistic deficiency rates. As such, we strictly enforce the minimization of these particular moments.[15] Additionally, as we show below, this provides further benefit in dramatically simplifying the computational burden of the estimation problem.

Second, we match the covariance between each individual deficiency and the sum of all other deficiencies: $\text{Cov}(x^d_{j,t}, \sum_{d' \neq d} x^{d'}_{j,t})$. This helps identify the γ_d governing the sen-

[14] Equivalently, $\Delta(j,t) = w(j,t-1)+1$, where $w(j,t-1)$ denotes the week-within-cycle immediately prior to the inspection at time t. Recall that w is indexed such that an inspection occurring at time t corresponds to $w(j,t) = 0$.

[15] Formally, we minimize the other moments subject to the constraint that this first set of moments are minimized. This is equivalent to placing arbitrarily larger weights on the first set of moments.

sitivity of each deficiency to common latent quality $q_{j,t}$. Third, we match a measure of the autocovariance between consecutive inspections, separately for each cycle length Δ: $\text{Cov}\big(x^d_{j,t}, \sum_{d'} x^{d'}_{j,t-\Delta} | \Delta\big)$. These moments help identify the extent to which latent quality $q_{j,t}$ is time-varying. The remainder of this section describes two key computational tricks that aid in computing and minimizing these moments.

Concentrating Out $\boldsymbol{\kappa}$ and λ. As discussed above, to ensure that the model maximally matches the empirical deficiency rates and their relationship with cycle length, we enforce the minimization of our first set of moments—$\mathbb{E}[x^d_{j,t} \mid \Delta] = \bar{x}^d_\Delta$ for each d and Δ—as a constraint in estimation. Conveniently, as we show in Theorem 1, minimizing these moments entails that given parameters $(\boldsymbol{\gamma}, \rho, \sigma_\varphi, \sigma_\epsilon, \delta)$, the optimal $\boldsymbol{\kappa}$ and λ are straightforward to compute via linear regression.

Theorem 1 (Proof in Appendix C) *For every admissible parameter vector $(\boldsymbol{\gamma}, \rho, \sigma_\varphi, \sigma_\epsilon, \delta)$, the unique $(\boldsymbol{\kappa}, \lambda)$ minimizing $\mathbb{E}[x^d_{j,t} \mid \Delta] = \bar{x}^d_\Delta$ for each d and Δ are the regression coefficients given by the linear relationship:*

$$\Phi^{-1}(\bar{x}^d_\Delta) \times \sqrt{1 + \gamma_d^2 \sigma_q^2} = \kappa_d + \lambda(\gamma_d \bar{s}_\Delta), \tag{14}$$

where $\sigma_q^2 \equiv Var(q_{j,t})$ and $\bar{s}_\Delta$ is the average effort stock at the end of cycles of length Δ.

Stacking equation (14) across all combinations of d and Δ yields a linear system that is easily solved given a candidate set of parameters $(\boldsymbol{\gamma}, \rho, \sigma_\varphi, \sigma_\epsilon, \delta)$. This provides a straightforward approach to "concentrate out" the nuisance parameters $\boldsymbol{\kappa}$ and λ, considerably simplifying the parameter search.

Long-run Distributions. Simulating the second set of moments entails drawing from the long-run distribution of $q_{j,t}$, and simulating the third set of moments entails drawing from the joint distribution of $q_{j,t}$ and the quality during the previous inspection $q_{jt-\Delta}$. Theorem 2 precisely characterizes the joint distribution for $q_{j,t}$ and $q_{jt-\Delta}$, making it fast and straightforward to simulate our second and third sets of moments.

Theorem 2 (Proof in Appendix C) *The vector of latent qualities follows a bivariate normal distribution with mean vector and covariance matrix given by:*

$$\begin{bmatrix} q_{j,t} \\ q_{j,t-\Delta} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_q^2 & V_\varphi + \rho^\Delta V_\epsilon \\ V_\varphi + \rho^\Delta V_\epsilon & \sigma_q^2 \end{bmatrix} \right) \tag{15}$$

where:

$$V_\varphi \equiv \frac{\sigma_\varphi^2}{(1-\rho)^2}, \qquad V_\epsilon \equiv \frac{\sigma_\epsilon^2}{1-\rho^2}, \quad \sigma_q^2 \equiv \text{Var}(q_{j,t}) = V_\varphi + V_\epsilon.$$

With each simulation draw from equation (15), one can compute the model-implied deficiency probabilities—i.e., $p^d_{j,t}$ and $p^d_{j,t-\Delta}$ from equation (12) for each d—necessary to compute the covariance moments.

7.2 Estimating Nursing Home Effort Incentives

Section 6.1 characterizes nursing home effort choices as the solution to a dynamic programming problem. The optimal policies implied by equation (10) are determined by a number of factors that must be estimated.

Inspection Hazard $h(w)$. We estimate the empirical hazard function non-parametrically for each $w < \overline{w}$ individually (and for all $w \geq \overline{w}$ together). For modeling and expositional clarity, we impose that $h(w)$ be weakly monotone. See Appendix E for details.

Noise in Realized Effort σ_e. It is computationally convenient to estimate σ_e prior to the main estimation of effort costs. To do this, notice in Figure 2 panel (a) that average effort levels are empirically flat between periods 21 and 30. Assuming the target effort levels are also flat over this period (i.e., $e^*_{j,k,w} \approx e^*_{j,k,w-1}$), the variance in observed effort reflects only noise: $\text{Var}(\Delta_w e_{j,k,w}) \approx \text{Var}(\Delta_w \epsilon^e_{j,k,w}) = 2\sigma_e^2$, where each $\epsilon^e_{j,k,w} \equiv e_{j,k,w} - e^*_{j,k,w}$ is an independent logistic with standard deviation σ_e. Therefore, we can infer σ_e from the observed variance in $\Delta_w e_{j,k,w}$ over this period:

$$\hat{\sigma}_e = \sqrt{\frac{\widehat{\text{Var}}(\Delta_w e_{j,k,w})}{2}}. \tag{16}$$

Inspection Benefit $\pi(s_{j,k,w})$. Equation (13) parameterizes the relationship between effort stock s and performance on inspections as a function of parameters $\{\kappa_d, \gamma_d, \lambda, \delta\}$ that we estimate in Section 7.1. Therefore, we simply utilize $\hat{\pi}(s_{j,k,w})$ characterized by $\{\hat{\kappa}_d, \hat{\gamma}_d, \hat{\lambda}, \hat{\delta}\}$.

Cost of Effort $c(e_{j,k,w}, s_{j,k,w-1})$. Our estimation principally focuses on estimating the effort costs that rationalize the observed pattern of effort over inspection cycles. This includes (α_1, α_2), which characterize the level cost of effort, and α_3, which characterizes the adjustment cost of effort. We estimate these via simulated method of moments, in which we match the average effort levels $(\overline{e}_w)$ and effort stocks $(\overline{s}_w)$ for $w \in \{21, \ldots, \overline{w}-1, \overline{w}\}$.[16]

[16] Note that we pool all $w \geq \overline{w}$ when constructing moments.

8 Estimates

8.1 Inspection Performance Estimates

Figure 6 presents our estimates of the process underlying inspection performance. Panel (a) summarizes the estimated parameters. These results are sensible: quality is highly persistent ($\rho = 0.95$), effort depreciates relatively quickly ($\delta = 0.30$), and approximately 46.7% of the overall variation in $q_{j,t}$ can be explained by persistent cross-facility variance (i.e., $\frac{V_\varphi}{\sigma_q^2} \approx 0.467$). Importantly, the relatively small estimated δ suggests that more recent effort choices carry substantially greater weight in the inspection payoffs. This creates a strong incentive for facilities to time their effort provision to be as close as possible to the inspection period.

We also estimate $D = 160$ γ_d and κ_d terms. The γ_d terms are systematically negative, indicating that higher quality is associated with lower likelihood of citation.[17] Figure H.13 presents the joint distribution of γ_d and κ_d, which exhibits little systematic relationship between baseline prevalence and structural sensitivity to quality. The semi-elasticities implied by our estimates are sensible: at mean quality, moving from initial effort stock to the average end-of-cycle effort stock produces a 30.2% average relative reduction in the probability of a deficiency citation, with a 17.1% standard deviation across deficiencies (Figure H.15). Panel (b) evaluates model fit and reassuringly shows that the distribution of deficiencies per inspection are similar for the model and the data.

Inspection Payoffs Our estimates parameterize the payoff function $\pi(s)$ given by equation (13). We find that this function is increasing and concave over its typical domain, suggesting diminishing returns to exerting higher levels of effort (Figure H.16).

Regulator's Uncertainty About Quality. Our estimates allow us to quantify the regulator's uncertainty about latent quality. On week w of the inspection cycle, the regulator's uncertainty is characterized by the conditional variance of quality $q_{j,t}$ given the results of the most recent inspection $\mathbf{x}_{j,t-w}$ and the facility's known static component of quality.[18] Taking the expectation over the joint distribution of $\mathbf{x}_{j,t-w}$ and φ_j yields our uncertainty measure:

$$U_w = \mathbb{E}\big[\mathrm{Var}(q_{j,t} \mid \mathbf{x}_{j,t-w}, \varphi_j)\big]/\sigma_q^2, \tag{17}$$

where we normalize by the unconditional variance in quality, σ_q^2, as it is the maximum possible uncertainty and represents the case in which the regulator has no information. Appendix C.5 provides derivations and additional details about this measure.

[17] Consistent with these patterns, we find reduced form evidence that higher effort and effort stock correspond to better inspection performance (Figure H.14).

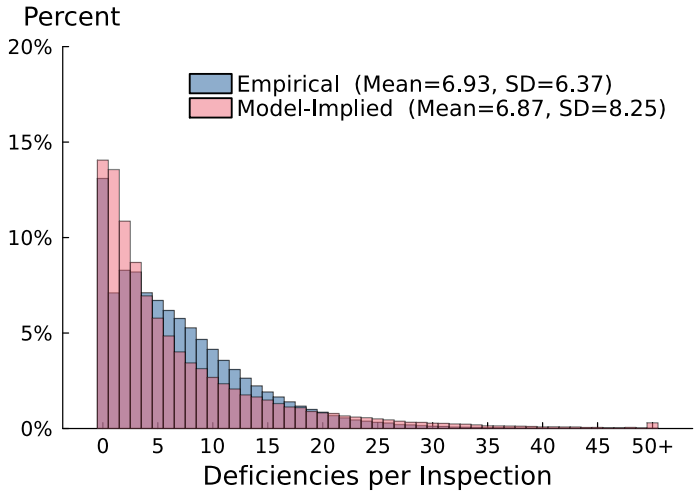
[18] In the long-run, a regulator will eventually learn φ_j with certainty, and so we condition on this as known.

Figure 6: Inspection Performance Estimation Results

(a) Estimated Model Parameters

Parameter	Estimate	SE	Description
σ_φ	0.038	0.002	Standard deviation of facility-specific intercepts
σ_ϵ	0.246	0.004	Standard deviation of quality shocks
ρ	0.947	0.002	Persistence of quality
λ	1.454	0.245	Effect of effort on performance
δ	0.300	0.025	Depreciation rate of effort
V_φ/σ_q^2	0.467	0.004	Share of quality variance explained by φ
	Mean	SD	
κ_d	-0.659	0.385	Parameters governing baseline prevalence
γ_d	-2.479	0.961	Sensitivity to quality and effort

(b) Model Fit

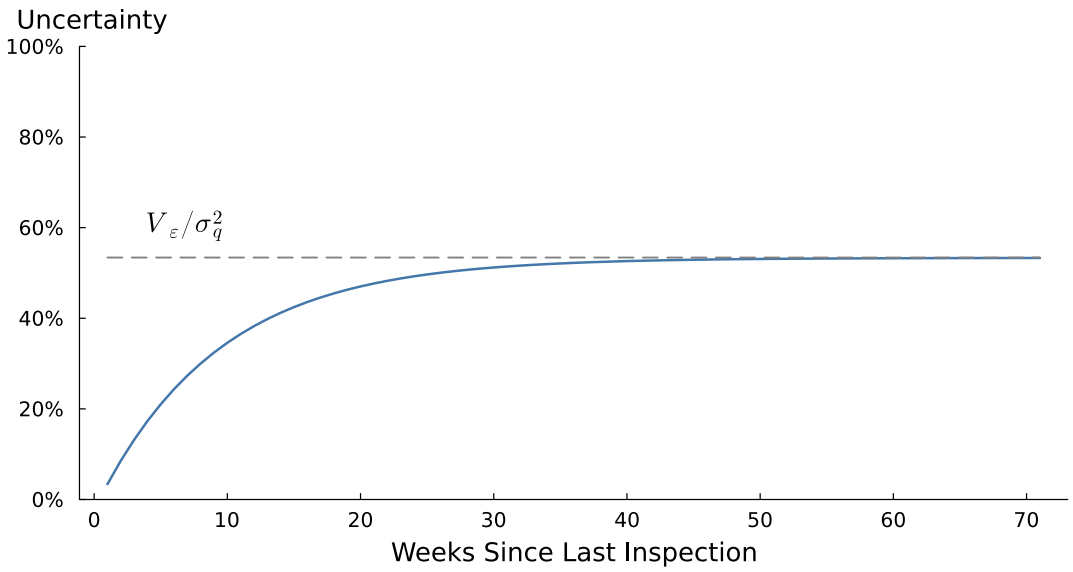


Notes: Panel (a) presents estimates of the inspection performance model parameters. Panel (b) presents a measure of model fit, plotting both the empirical and model-implied distributions of inspection deficiencies.

This measure characterizes the regulator's average uncertainty on week w of an inspection cycle and has several intuitive properties. First, it is increasing in w: the longer a facility goes without inspection, the more uncertain the regulator becomes. Second, the level of uncertainty at $w = 0$ (i.e., immediately after an inspection occurs) reflects the informativeness of the inspection signal. Third, the measure is bounded between 0% (exact knowledge of $q_{j,t}$) and 100% (no information about $q_{j,t}$). Finally, the asymptotic level of uncertainty, $\lim_{w\to\infty} U_w = \frac{V_\epsilon}{\sigma_q^2}$, captures the relative contribution of dynamic versus static components of quality variation. If most of the variation in quality is static, then inspection timing has little effect on long-run uncertainty. Conversely, if latent quality fluctuates meaningfully over

time, then inspection timing plays a much larger role in shaping the regulator's uncertainty.

Figure 7: Uncertainty Over Quality



Notes: This figure characterizes the regulator's uncertainty over quality, given by U_w. Derivation of this uncertainty measure provided in Appendix C.5.

Figure 7 illustrates these properties for the U_w implied by our estimates. Immediately after an inspection, average uncertainty over quality is low—just 3.5% of the maximum—indicating that inspections are highly informative. However, this signal decays over time, and average uncertainty rises each week towards an asymptote of $V_\epsilon/\sigma_q^2 = 53.3\%$, the share of variance in quality that is dynamic.[19] This implies that inspection timing can impact only about one half of the total potential uncertainty about quality. Note that in Section 9, when we characterize the benefits and information provided by inspection regimes, we work with $1 - U_w$ (Figure H.17), which we describe as the regulator's "information."

8.2 Effort Model Estimates

Figure 8 presents our estimates of the parameters governing nursing homes' effort choices. Panel (a) summarizes the estimated parameters. The results are economically sensible. First, they imply that the cost of effort is convex, which discourages facilities from concentrating effort in few high-intensity periods and explains why average effort remains moderate even very late in the cycle. Second, we estimate substantial adjustment costs, consistent with facilities' slow ramping of effort even when inspection hazard increases rapidly. Panel (b)

[19] Equivalently, the share of variance that is static is $V_\varphi/\sigma_q^2 = 46.7\%$.

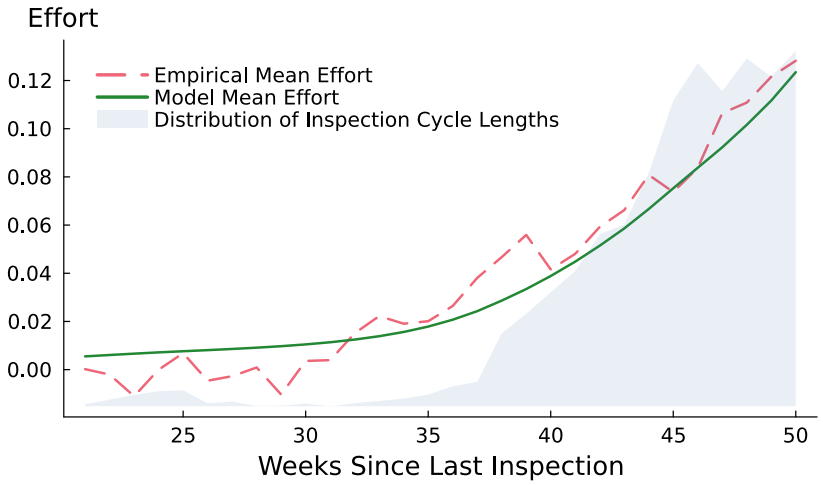
evaluates model fit and reassuringly shows that the model closely replicates the evolution of effort over the inspection cycle, which is the key empirical pattern motivating our exercise.

Figure 8: Effort Choice Estimation Results

(a) Estimated Model Parameters

Parameter	Estimate	SE	Description
α_1	-0.016	0.007	Linear effort cost
α_2	2.768	0.060	Quadratic effort cost
α_3	23.652	4.223	Adjustment cost
σ_e	0.619	0.001	SD of effort noise

(b) Model Fit



Notes: Panel (a) presents estimates of the effort parameters. Panel (b) presents the corresponding model fit, plotting both the mean nursing home effort levels over time as well as the model-implied optimal effort. The figure also includes the empirical distribution of inspection cycle length.

Note that for our primary results, we estimate the model jointly across all states. We assess robustness in Appendix D by estimating our model separately across different groups of states. Reassuringly, we find similar results when subsetting to states with similar inspection frequencies or when examining a single large state (California).

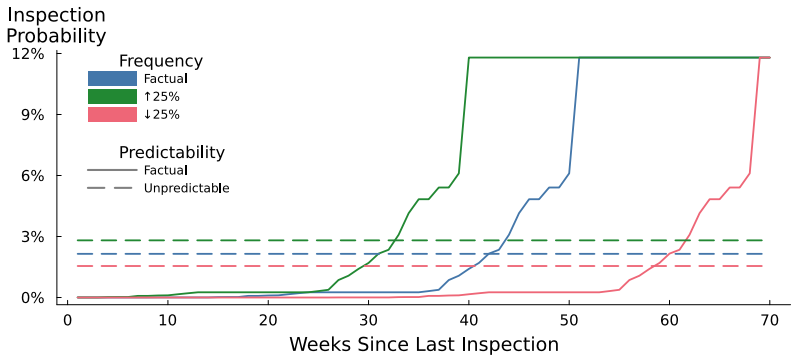
9 Quantifying the Value of Inspections

We next use the estimated model to simulate different inspection regimes. Each regime is fully characterized by its inspection hazard rate $h(w)$, which captures both the frequency and the predictability of inspections. Throughout, we benchmark regimes against a counterfactual with no inspections, i.e., a regime of $h(w) = 0$ for all w. Correspondingly, the results we present are the benefits of each regime relative to having no inspections at all. We then

contrast these benefits across different regimes..

Figure 9: Overview of Inspection Regimes

(a) Inspection Hazard Rates



(b) Summary of Results

	Frequency (Per Facility-Year)	Predictability	Lives Saved Annually	Lives Saved Per 1,000 Inspections	Information (%)	Information (%) Per 1,000 Inspections
No Inspections (Benchmark)	0	-	0	-	0	-
Current Regime	0.99	Factual	850.6	54.9	56.9	3.7
Increase Frequency (↑ 25%)	1.25	Factual	1,066.2	54.4	59.7	3.0
Decrease Frequency (↓ 25%)	0.74	Factual	635.1	54.9	54.1	4.7
Unpredictable	0.99	Unpredictable	954.5	61.6	55.4	3.6
Unpredictable; Increased Frequency (↑ 25%)	1.25	Unpredictable	1,190.9	60.8	57.6	2.9
Unpredictable; Decreased Frequency (↓ 25%)	0.74	Unpredictable	720.1	62.3	53.3	4.6
Perfectly Predictable	0.98	Perfectly	710.4	46.1	56.9	3.7
Perfectly Predictable; Increased Frequency (↑ 25%)	1.24	Perfectly	896.1	46.1	59.9	3.1
Perfectly Predictable; Decreased Frequency (↓ 25%)	0.73	Perfectly	530.5	46.1	54.1	4.7

Notes: Panel (a) presents the hazard functions $h(w)$ characterizing many of the inspection regimes that we consider. We omit the benchmark regime, for which $h(w) = 0$ for all w, as well perfectly-predictable regimes, where the hazard is 100% on a specific week. Panel (b) summarizes our findings for each inspection regime that we consider.

We start with the current inspection regime in Section 9.1 and study its effects on facility effort, patient health, and regulator uncertainty. Then, in Section 9.2, we consider alternative regimes that vary either inspection frequency, predictability, or both.

For each regime, we first solve for facilities' effort policies according to equation (10). This, in conjunction with $h(w)$, implies a long-run distribution of facility effort. The estimated effort-survival relationship in Figure 3 allows us to translate differences in long-run effort distributions into implied effects on patient mortality.[20] We then assess the informational value of each regime. Specifically, the hazard rate characterizing a given regime implies a long-run distribution of time since last inspection (w) and therefore a long-run distribution of uncertainty (U_w) summarized in Figure 7. For each regime, we characterize

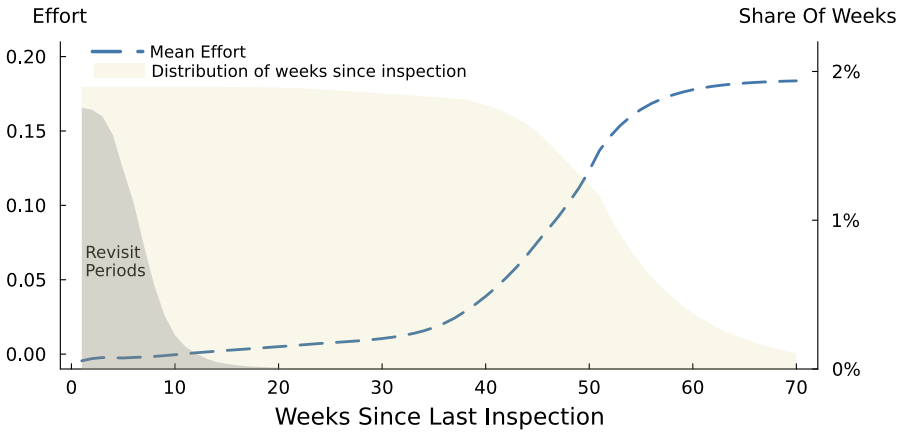
[20]Importantly, because our primary health outcome is 4-week survival, to avoid over-counting we convert effort-induced changes in survival into weekly rates.

the distribution of "information" relative to no inspections, which is simply $1 - U_w$ (Figure H.17).

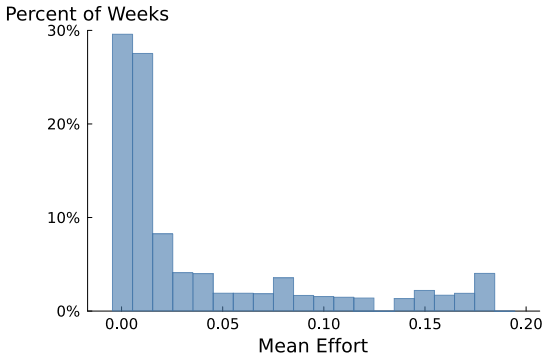
Figure 9 details the set of regimes we consider. Panel (a) depicts the hazard rate of each counterfactual, and Panel (b) summarizes our findings from Sections 9.1 and 9.2. Appendix F provides further details on implementation, as well as additional results.

Figure 10: Effort and Information Under the Current Regime

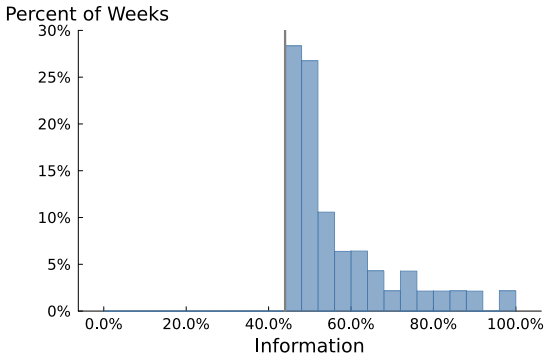
(a) Effort Path and Distribution of w



(b) Distribution of Effort



(c) Distribution of Information



Notes: This figure presents the results from our simulation using the factual hazard rate $h(w)$ at the frequency in the current inspection regime. Panel (a) presents the optimally chosen effort at each week from last inspection, as well as the distribution of weeks since inspection w induced by $h(w)$. Panel (b) combines these two pieces of information, presenting the distribution of e_w^* weighted by the relative frequency of w. Panel (c) presents the corresponding distribution for informational gain $(1 - U_w)$ relative to the no-inspection benchmark. The solid line denotes the minimum level of information a regulator that knows the static component of quality can hold.

9.1 Current Inspection Regime

Figure 10 quantifies the value of the current inspection regime. Panel (a) shows facilities' average optimal effort at each of week w of an inspection cycle under the current inspection regime. As expected, the model implies low levels of effort until nearly $w = 40$, when average effort increases rapidly.[21] The long-run distribution of w implied by the factual hazard is overlaid. The shape of this distribution is quite intuitive: few cycles last long enough to reach high values of w. Note that we also shade the empirical distribution of time that facilities spend in their revisit window, because we make the conservative assumption that behavior during these weeks is unaffected by any inspection regime. Appendix Section F.2 shows that our findings are robust to simply eliminating $w \leq 20$ from all simulations.

Importantly, although average effort rises substantially late in the cycle, most of the long-run distribution of w falls in the earlier weeks, when effort is lower. Panel (b) formalizes this, plotting the long-run distribution of effort implied by panel (a) by weighting the simulated effort policy by the long-run distribution of w. While the resulting distribution features a long tail, the bulk of the mass falls at very low levels of effort.

Finally, panel (c) reports the analogous distribution of information $(1 - U_w)$ weighted by the long-run distribution of w from panel (a). Figure 7 indicates regulators are highly-informed about quality immediately after inspection but that information about the time-varying component of quality decays very quickly. By the middle of a typical inspection cycle, regulators know little about the current value of the dynamic component of quality. Correspondingly, much of the distribution in panel (c) is only modestly above the "floor" of knowing only the static component of quality.

While both effort and information are often low under the current inspection regime, it still represents a considerable improvement over the counterfactual in which no inspections occur. This contrast is summarized in panel (b) of Figure 9. Compared to the benchmark of no inspections, the current inspection regime saves 850.6 lives per year by incentivizing higher facility effort.[22] Likewise, the current inspection regime also provides considerable information about facility quality: on an average week, the regulator has 56.9% information, compared to the 0% they would have if no inspections occurred at all.

[21] Note that this effort curve corresponds to the empirical hazard rate and is therefore identical to the one used to evaluate model fit in panel (b) of Figure 8.

[22] Note that this is likely a conservative estimate, as we are only accounting for the anticipatory behavior before an inspection. We do not explicitly model the deficiency-correction role of inspections nor the effects of publicizing inspection outcomes, both of which prior literature has shown to improve nursing home quality and patient allocation (Cheng, 2023; Chen and Dillender, 2025).

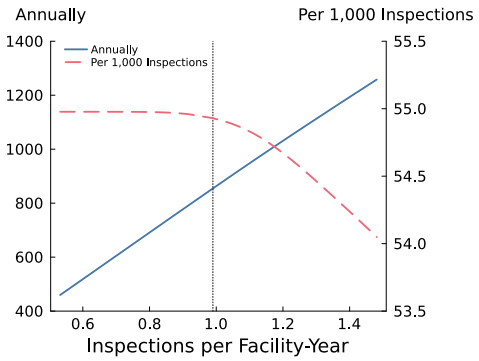
9.2 Alternative Inspection Regimes

We next consider alternative inspection designs that vary the frequency and predictability of inspections, allowing us to disentangle the relative contributions of each in incentivizing effort and informing regulators.

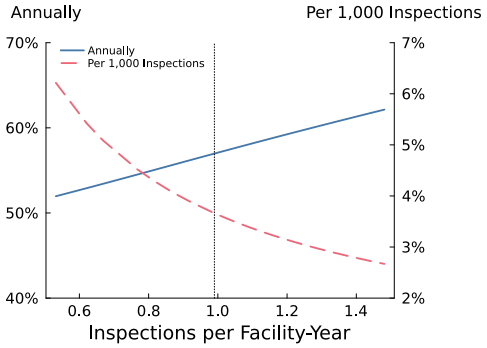
Frequency of Inspections. Policymakers concerned by low-quality care have recently advocated increasing the frequency of nursing home inspections (U.S. Senate, 2023). We use our model to evaluate the importance of inspection frequency while holding predictability fixed. Panel (a) of Figure 9 illustrates how we operationalize this: we hold the shape of the hazard function fixed and shift it left or right to achieve the desired frequency.

Figure 11: Effect of Varying Inspection Frequency

(a) Lives Saved



(b) Information



Notes: This figure summarizes the results of flexibly adjusting the inspection frequency from −50% to +50% of the current rate, while preserving the shape of the hazard function. Panel (a) presents both the annual number of lives saved at each inspection frequency as well as the number saved per 1,000 inspections. Panel (b) presents a similar calculation for the informational gains.

Figure 11 summarizes the impact of alternative regimes ranging from a 50% reduction to a 50% increase in frequency.[23] Panel (a) shows that the number of lives saved consistently increases with the frequency of inspections. For example, a 25% increase in frequency of inspections (0.26 additional inspections per facility-year) saves an additional 215.6 (25.3%) lives. The marginal benefit of additional inspections above the current frequency is slightly decreasing: a 25% increase in frequency reduces the lives saved per 1,000 inspections from 54.9 to 54.4. However, at the factual inspection frequency (approximately 0.99 per facility-year), the current regime remains well within the range of roughly constant marginal benefit.

Panel (b) shows a similar pattern for how inspection frequency increases the regulator's information about quality. With more inspections, the regulator receives more frequent

[23] Figure F.1 presents the full results—analogous to Figure 10—for a 25% increase or decrease in frequency.

signals about quality, and correspondingly, her most recent signal tends to be more up to date. For example, a 25% increase in frequency increases the regulator's information during the average week modestly from 56.9% to 59.7%. Again, we find that the marginal benefits of additional effort are decreasing: a 25% increase in frequency reduces the information per 1,000 inspections from 3.7% to 3.0%.

Predictability of Inspections. An alternative to changing the frequency of inspections is to change their predictability. At one extreme, a *fully unpredictable* inspection regime entails a "flat hazard" under which the probability of inspection is the same in every week. At the other extreme, a *perfectly predictable* inspection regime entails an hazard rate that is zero except during the scheduled inspection week, when it is one.

Figure 12 contrasts the implications of unpredictable and fully-predictable inspection regimes, holding inspection frequency fixed at the current average rate of approximately one inspection every 53 weeks. Panels (a) and (b) show the optimal effort curves and long-run distributions of w under each regime. Under the unpredictable regime, constant inspection risk eliminates incentives for strategic timing of effort provision. Consequently, facilities provide modest, constant effort in all periods. In contrast, under the fully-predictable regime, facilities exert only minimal effort for most of the cycle and concentrate nearly all of their effort in the immediate lead-up to the scheduled inspection.[24] Panel (c) formalizes this comparison by contrasting the long-run distributions of effort under each inspection regime. While the fully-predictable regime does yield some extremely high-effort weeks, the vast majority of weeks have mean effort below those from the unpredictable regime.

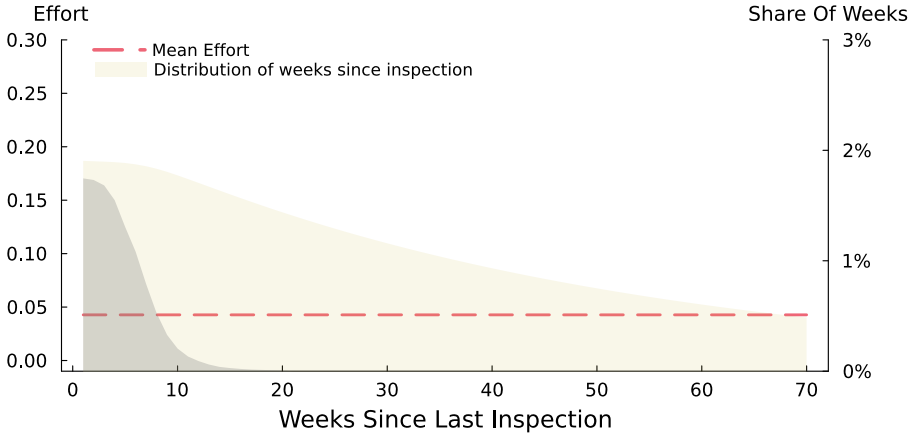
We calculate that effort exerted under the unpredictable regime saves on average 954.5 lives per year relative to no inspections. Importantly, this implies that an unpredictable regime saves 103.9 more lives than the current regime *without adding any additional inspections*. In other words, unpredictability increases the efficiency of inspections to saving 61.6 per 1,000 inspections from 54.9 lives per 1,000 inspections under the current regime. In contrast, a predictable regime saves just 710.4 lives per year and reduces the efficiency of inspections to 46.1 lives saved per 1,000 inspections.

While budget-neutral, unpredictability comes at a cost to the regulator's information about quality. Panel (d) shows that, because some facilities go longer between inspections, more time is spent with low levels of information. The difference, however, is quite modest: the average information under an unpredictable inspection regime is 55.4%, which is only slightly lower than 56.9% under the factual regime. To be indifferent between the two, the regulator would need to value each percentage point of information at roughly 70.0 lives per

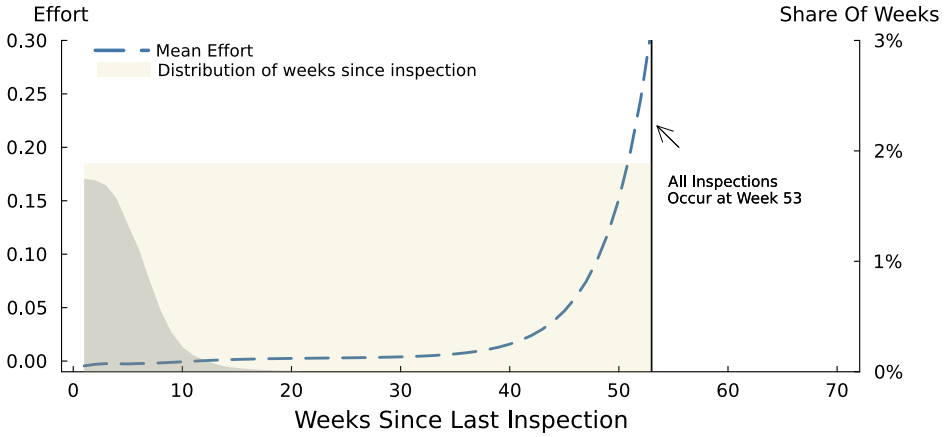
[24] Convex adjustment costs cause facilities to begin ramping up effort slightly prior to the scheduled week of inspection.

Figure 12: Effort and Uncertainty Reduction under Predictable and Unpredictable Regimes

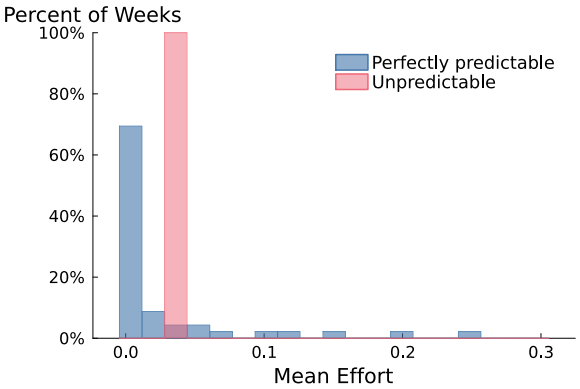
(a) Effort Path and Distribution of w under Unpredictable Regime



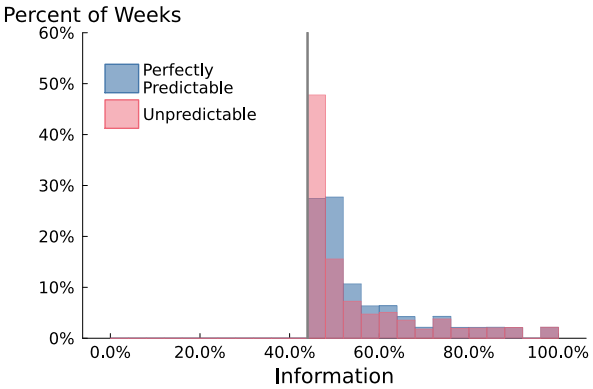
(b) Effort Path and Distribution of w under Predictable Regime



(c) Distribution of Effort



(d) Distribution of Information



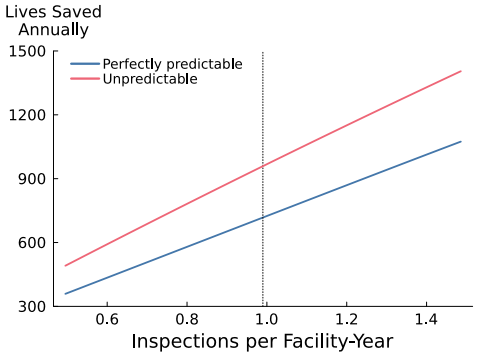
Notes: This figure presents the distribution of effort and informational gain under two regimes. The gray areas in panels (a) and (b) refer to revisit periods. In the unpredictable regime, inspections occur with constant hazard $h(w) = \bar{h}$. In the scheduled regime, inspections occur in the same week each cycle, with $h(w) = 0$ for all $w < 53$ and $h(53) = 1$. The solid line in panel (d) denotes the minimum level of information a regulator that knows the static component of quality can hold.

year.[25] In contrast, the fully-predictable regime delivers a distribution of information that is almost identical to the current regime with an average information that differs by less than 0.1%.

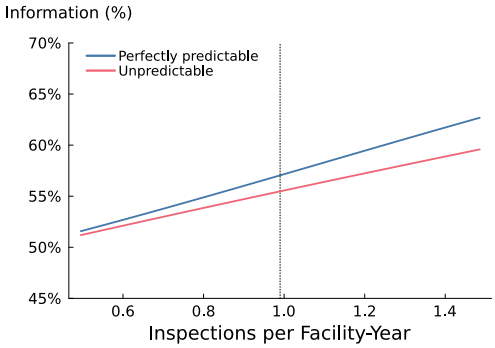
Interaction Between Predictability and Frequency. Finally, we examine how predictability shapes the returns to inspection frequency. Figure 13 plots the results of varying frequency under both unpredictable and perfectly predictable regimes.[26]

Figure 13: Predictability and the Value of Frequency

(a) Lives Saved



(b) Information



Notes: This figure summarizes the results of flexibly adjusting the inspection frequency from −50% to +50% of the current rate, under both the fully unpredictable and fully predictable inspection regimes. Panel (a) presents the annual number of lives saved at each inspection frequency. Panel (b) presents a similar calculation for the informational gain.

Panel (a) shows a clear complementarity between unpredictability and frequency in lives saved: the marginal returns to frequency are larger under the unpredictable regime than under the perfectly-predictable one. Intuitively, unpredictability amplifies the benefits of more frequent inspections, as the additional inspection risk incentivizes slightly higher effort in all periods. In contrast, the incremental inspection risk from increasing frequency under the perfectly-predictable regime is concentrated at a point in the cycle when further increasing efforts would be highly costly, limiting the efficacy of the increase in frequency.

Panel (b) reveals the opposite for the information channel: the marginal returns to frequency are larger under the perfectly-predictable regime than under the unpredictable one. Inspection signals about quality decay over time, and so increasing frequency yields informa-

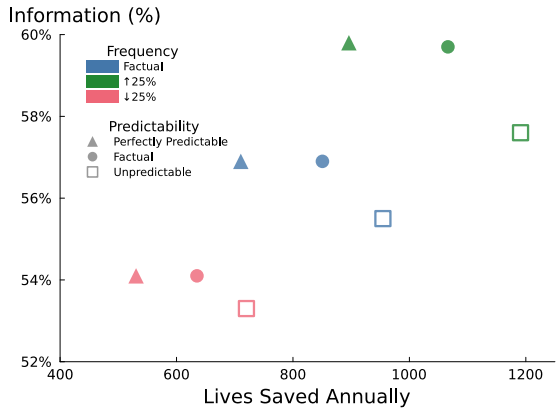
[25] Measuring the precise value of information is outside the scope of this paper and depends on a number of factors, including the regulator's ability to disseminate that information, elasticity of demand to quality measures, and providers' capacity constraints. Nonetheless, recent research finds that information provision does relatively little to shift patients' choice of facilities (Cheng, 2023), suggesting that slightly more accurate quality information would have minimal impact on outcomes.

[26] Figure F.2 presents the analogous results scaled per 1,000 inspections.

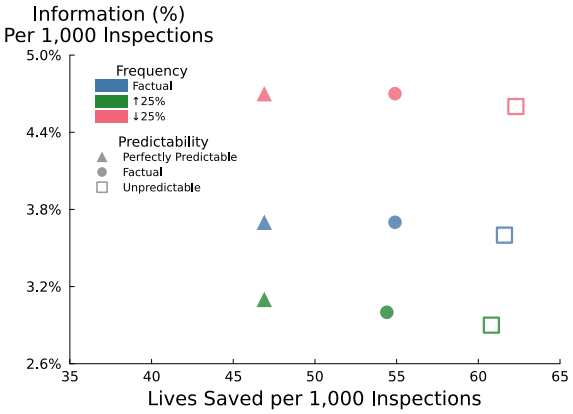
tional gains by replacing late-w periods with early-w periods. With perfect predictability, increasing frequency is highly effective, as it entails removing only the lowest-information w. Increasing frequency under an unpredictable regime, however, shifts the whole distribution of w left and is therefore less effective by nature of being less targeted.

Figure 14: Summary of Information and Lives Saved Under Each Regime

(a) Annually



(b) Per 1,000 Inspections



Notes: This figure summarizes the results of each inspection regime we consider, examining both the information under each inspection regime as well as the number of lives. Panel (a) presents results aggregated annually. Panel (b) presents results scaled per 1,000 inspections.

Summary and Discussion. Figure 14 summarizes the key intuition underlying these results. Frequency and predictability are both policy choices made by regulators. Inspecting facilities more frequently provides greater incentive to exert high levels of effort, while also providing more opportunities to learn about latent quality. However, more inspections are costly, and increasing frequency diminishes their efficiency. By contrast, limiting the predictability of inspections incentivizes firms to exert even more effort while holding frequency constant, increasing their efficiency. However, doing so comes at the cost of informational loss: more irregular inspections may mean less updated information about quality. Finally, there is a clear interaction between frequency and predictability: unpredictable regimes strengthen effort-inducing benefits but weaken the information gains from frequency.

10 Conclusion

This paper examines how the timing and predictability of inspections shape compliance and quality in the context of U.S. nursing homes. Using high-frequency data on staffing and patient outcomes, we document clear cyclical patterns in effort and health: following an inspection, facilities exert little effort while the risk of a new inspection is low and increase

their effort as the next inspection becomes more likely. Patient survival tracks these changes closely, indicating that inspection regimes have direct consequences for patient health.

To quantify the value of different inspection regimes, we develop and estimate a dynamic model in which nursing homes balance effort costs today against the benefits of positive performance when an inspection arrives in the near future. The model captures both the incentive effects of unannounced inspections, as well as the informational role of inspections in providing signals of latent quality. Our results show that less-predictable inspections can increase average effort and patient survival as much as a substantial increase in inspection frequency, while only modestly reducing the informativeness of inspection results. Moreover, unpredictability amplifies the incentive benefits of more frequent oversight.

More generally, our findings highlight predictability as an important dimension of inspection regimes warranting greater attention. Designing inspection regimes that are less predictable can sustain higher agent effort and improve overall compliance without either increasing the cost of conducting inspections or greatly diminishing their informational value to the principal. 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 inspection predictability.

References

Alsan, Marcella and Crystal Yang, "The Hidden Health Care Crisis Behind Bars: A Randomized Trial to Accredit U.S. Jails," Technical Report w33357, National Bureau of Economic Research January 2025.

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.

Bartholomew, David, Martin Knott, and Irini Moustaki, *Latent Variable Models and Factor Analysis: A Unified Approach* Wiley Series in Probability and Statistics, John Wiley & Sons, 2011.

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 here.

Chen, Yiqun and Marcus Dillender, "Government Monitoring of Health Care Quality: Evidence from the Nursing Home Sector," July 2025.

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 here.

_ , "PBJ Public Use Files: Technical Specifications," April 2022. Available here.

_ , "Medicare State Operations Manual - Chapter 7 - Survey and Enforcement Process for Skilled Nursing Facilities and Nursing Facilities," Technical Report 2023.

_ , "Medicare State Operations Manual - Appendix PP - Guidance to Surveyors for Long Term Care Facilities," Technical Report 2024.

_ , "Quality, Safety & Oversight - General Information," December 2024. Available here.

_ , "Care Compare," 2025. Available here.

_ , "Federal Regulatory Groups for Long Term Care," Technical Report April 2025.

_ , "Five-Star Quality Rating System," 2025. Available here.

_ , "Medicare State Operations Manual - Chapter 5 - Complaint Procedures," Technical Report 2025.

_ , "Special Focus Facility ("SFF") Initiative - Background," Technical Report 2025.

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," 2025. NBER Working Paper No. 32258.

_ **and Maggie Shi**, "Screening Through Soft Spending Limits: Evidence from the Medicare Therapy Cap," 2025.

_ **, Andrew Olenski, 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**, "Private equity, Consumers, and Competition: Evidence from the Nursing Home Industry," *NBER Working Paper Series*, 2025.

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.

Grice, James W, "Computing and evaluating factor scores.," *Psychological methods*, 2001, *6* (4), 430.

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, 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.

Konetzka, Tamara R, David C Grabowski, Marcelo Coca Perraillon, and Rachel M Werner, "Nursing home 5-star rating system exacerbates disparities in quality, by payer source," *Health affairs*, 2015, *34* (5), 819–827.

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.

Lin, Wilson, Lauren Xiaoyuan Lu, and Susan Feng Lu, "Do Inspection Delays Lead to Quality Decline? Evidence from U.S. Nursing Homes," Technical Report, SSRN Working Paper June 2025. June 23, 2025.

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.

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, "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. U.S. Senate Committee on Aging.

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 Additional Details on Data Construction

A.1 Sample Construction

Sample Restrictions To define facility inspection cycles, we restrict to standard health inspections, and set each 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 91.9 percent of all nursing home residents. 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.

Construction of Nursing Home Stays We construct patient stays 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 with return not anticipated, when the patient dies, or if another nursing home reports an admission assessment for the patient. If a stay does not have a discharge assessment, an associated 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).

A.2 Factor Analysis for Composite Effort Index

Here we provide a very brief overview of factor analysis, for readers unfamiliar with these techniques. For a more detailed treatment, we direct readers toward Bartholomew et al. (2011).

Setup. Each standardized input choice variable $\tilde{x}_{jt}^k$ is a noisy proxy of the latent variable, effort e_{jt}. We can express each k as:

$$\tilde{x}_{jt}^k = \lambda_k e_{jt} + \varepsilon_{jt}^k$$

The λ_k terms are the factor loadings, which measure how strongly each observed variable reflects latent effort. The ε_{jt}^k terms are idiosyncratic errors, assumed to be uncorrelated across k.

To estimate the factor loadings λ_k, we make the following assumptions:

1. $\mathbb{E}[e_{jt}] = 0$, $\text{Var}(e_{jt}) = 1$ (normalize effort)
2. $\mathbb{E}[\varepsilon_{jt}^k] = 0$, $\mathbb{E}[e_{jt} \cdot \varepsilon_{jt}^k] = 0$ (errors are mean-zero and orthogonal to effort)

3. $\mathbb{E}[\varepsilon_{jt}^k \cdot \varepsilon_{jt}^\ell] = 0$ for $k \neq \ell$ (errors are uncorrelated across variables)

We can then characterize the variance-covariance structure for the standardized observed variables $\tilde{x}_{jt}^k$. To do so, first notice that the variance of each x_{jt}^k is given by:

$$\text{Var}(\tilde{x}_{jt}^k) = \lambda_k^2 \cdot \text{Var}(e_{jt}) + \text{Var}(\varepsilon_{jt}^k) = \lambda_k^2 + \psi_k$$

where $\psi_k \equiv \text{Var}(\varepsilon_{jt}^k)$ is referred to as the "uniqueness" of variable k.

Second, notice as well that the covariance between two variables $\tilde{x}_{jt}^k$ and $\tilde{x}_{jt}^\ell$ (for $k \neq \ell$) is:

$$\text{Cov}(\tilde{x}_{jt}^k, \tilde{x}_{jt}^\ell) = \lambda_k \lambda_\ell$$

because the only shared source of variation is e_{jt} (which by assumption is scaled to have variance 1), and the errors are uncorrelated.

Stack the K standardized variables into a vector:

$$\mathbf{x}_{jt} = \mathbf{\Lambda} e_{jt} + \boldsymbol{\varepsilon}_{jt}$$

where:

- $\mathbf{x}_{jt} \in \mathbb{R}^K$
- $\mathbf{\Lambda} = [\lambda_1, \ldots, \lambda_K]'$ is the $K \times 1$ vector of loadings
- $\boldsymbol{\varepsilon}_{jt} \in \mathbb{R}^K$ is the vector of idiosyncratic terms

The covariance matrix of the observed variables can then be expressed as:

$$\Sigma = \text{Cov}(\mathbf{x}_{jt}) = \text{Cov}(\mathbf{\Lambda} e_{jt} + \varepsilon_{jt}) = \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Psi}$$

where Σ is the $K \times K$ covariance matrix of $\mathbf{x}_{jt}$ and $\mathbf{\Psi}$ is a diagonal matrix with entries ψ_k. This equality follows from the independence assumption between e_{jt} and ε_{jt}^k.

The parameters to be estimated are the factor loadings λ_k and the uniqueness terms ψ_k. Intuitively, there is a model-implied covariance matrix Σ and an empirical estimate of the covariance matrix $\widehat{\Sigma}$, and the model can be estimated by aligning the two.

For simplicity in estimation, we assume normality for the effort e and error ε terms, implying:

$$\mathbf{x}_{jt} \sim N(0, \Sigma)$$

though this assumption can be relaxed. Estimation of Σ then proceeds via MLE. In implementation, we rely on the `psych` package in R, version 2.5.6.

Once we have the factor loadings $\hat{\lambda}_k$, we can recover latent effort e_{jt}. To do so, we rely on the ten Berge factor scoring method (Grice, 2001). Denoting the correlation matrix across inputs $\mathbf{x}_{jt}$ as $\mathbf{R}$, we can express:

$$\hat{e}_{jt} = \mathbf{w}'\mathbf{x}_{jt}$$

where the weight vector $\mathbf{w}$ is:

$$\mathbf{w} = \left(\mathbf{\Lambda}'\boldsymbol{R}^{-1}\mathbf{\Lambda}\right)^{-1/2}\mathbf{R}^{-1}\mathbf{\Lambda}$$

Factor Loadings. Table A.1 reports the factor loading estimates as well as the cross-correlations between the standardized inputs to the composite effort index.

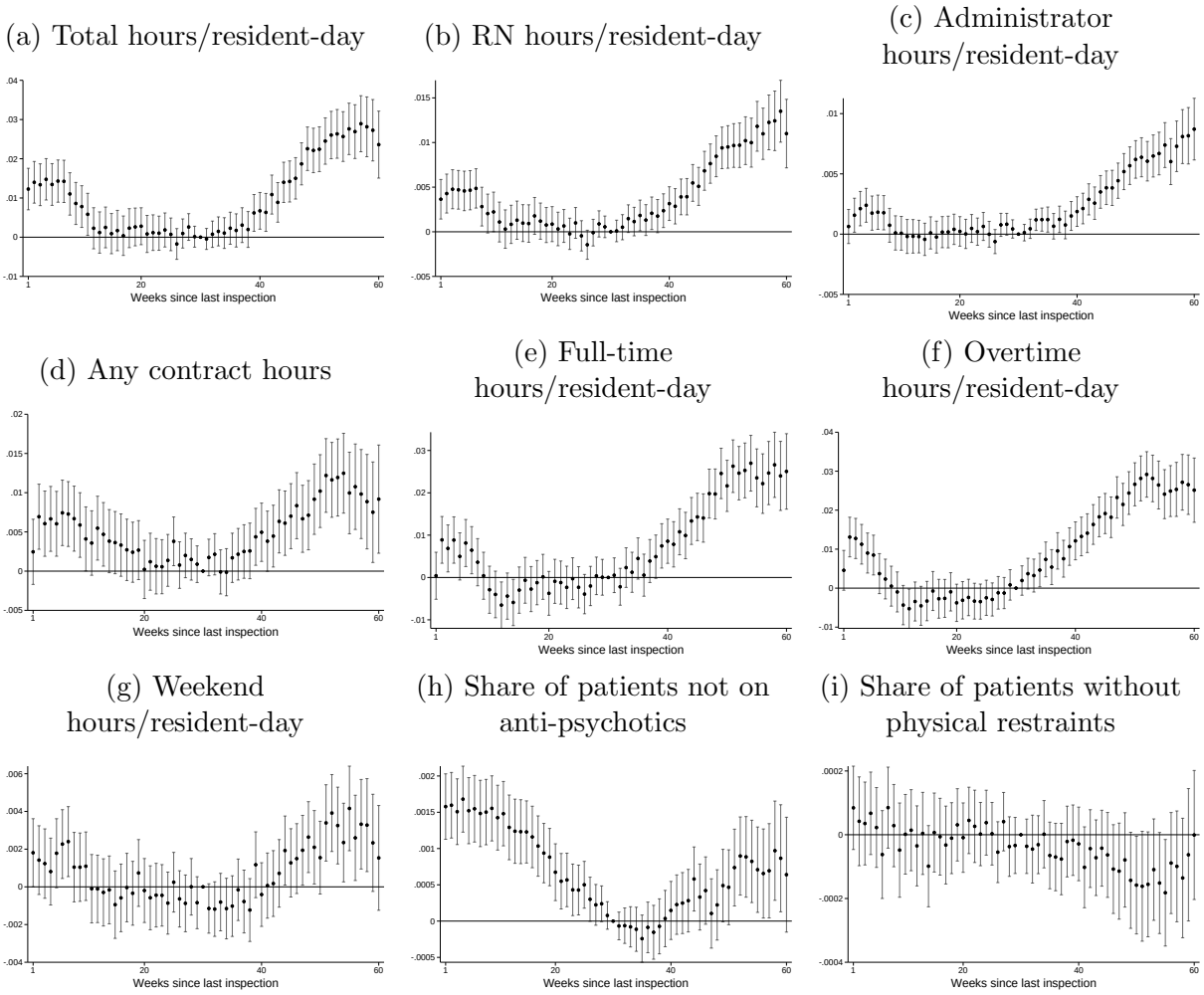
Table A.1: Factor Loadings and Cross-correlation Table

		Cross-correlations								
	Factor loading (λ_k)	Hours per resident	RN hours	Admin hours	Any contract	FT hours	Overtime hours	Weekend hours	% No Antipsych	% No Restraint
Hours/resident	0.987	1.000								
RN hours/resident	0.671	0.665	1.000							
Admin hours/resident	0.420	0.418	0.434	1.000						
Any Contract	0.001	0.000	0.008	0.006	1.000					
Full-time hours/resident	0.726	0.714	0.407	0.291	-0.001	1.000				
Overtime hours/resident	0.417	0.399	0.153	0.131	0.052	0.741	1.000			
Weekend hours/resident	0.886	0.876	0.585	0.303	-0.004	0.644	0.369	1.000		
% No antipsychotics	0.294	0.290	0.248	0.156	0.016	0.180	0.074	0.264	1.000	
% No restraint	-0.057	-0.056	-0.059	-0.027	0.008	-0.036	-0.009	-0.052	0.003	1.000

Notes: This table reports the factors loading from the factor analysis used to construct the composite effort index (Section 3.1 and Appendix Section A.2), as well as the cross-correlations between the subcomponents. Staffing measures are derived from the PBJ and measured as hours per resident-day. Patient care inputs are derived from assessments in MDS 3.0. Sample is the facility effort sample (2017-2019).

Decomposing Cyclicality. To investigate the drivers of the effort estimates in Figure 2, we repeat our inspection cycle analysis on each component of the composite effort index. Figure A.1 plots the estimates from equations (2) and (3) on (a) total hours per resident-day, (b) RN hours per resident-day, (c) administrator hours per resident-day, (d) an indicator for any contractor hours, (e) full-time hours per resident-day, (f) overtime hours per resident-day, (g) weekend hours per resident-day, (h) the share of patients not on anti-psychotics, and (i) the share of patients without physical restraints. All of these components display a similar "U-shape" as in Figure 2, where effort decreases in the middle of the inspection cycle and increases again as the next inspection nears. The share of patients without physical restraints shows the most muted response, which is consistent with its relatively low factor loading (Table A.1).

Figure A.1: Components of Effort Index by Weeks Since Last Inspection



Notes: This figure plots the coefficients from estimating equation (2) on the components of the effort index. Panel (a) plots staffing hours per resident-day (PBJ), panel (b) plots RN hours per resident-day (PBJ), panel (c) plots administrator hours per resident-day (PBJ), panel (d) plots an indicator for whether the nursing home reported any contractor hours (PBJ), panel (e) plots full-time hours per resident-day (PBJ), panel (f) plots overtime hours per resident-day (PBJ), panel (g) plots weekend hours per resident-day (PBJ), panel (h) plots the share of patients not on anti-psychotics (MDS), and panel (i) plots the share of patients without physical restraints (MDS). The coefficients for week 0 are estimated but not reported.

B Additional Empirical Analyses

B.1 Inspection Predictability

Although nursing homes are aware of their inspection hazard throughout the cycle — understanding the broad window during which an inspection is likely to occur — they are unlikely to know the exact timing of their next inspection, particularly several months in advance. Legally, inspection dates are unannounced, and it is operationally infeasible for survey agencies to schedule inspections far ahead of time. State survey agencies are chronically under-funded (U.S. Senate, 2023), and inspectors must also conduct unplanned visits, including revisits and complaint-based inspections.

Revisit inspections are unpredictable because the number required depends on how quickly each facility corrects its deficiencies. Complaint inspections are similarly difficult to anticipate: states are required to investigate shortly after receiving a complaint, and in the most severe cases, an inspection must begin within two days (CMS, 2025d). Coupled with high turnover and vacancy rates among survey agency staff, these factors complicate advance scheduling and reduce the predictability of inspection timing.

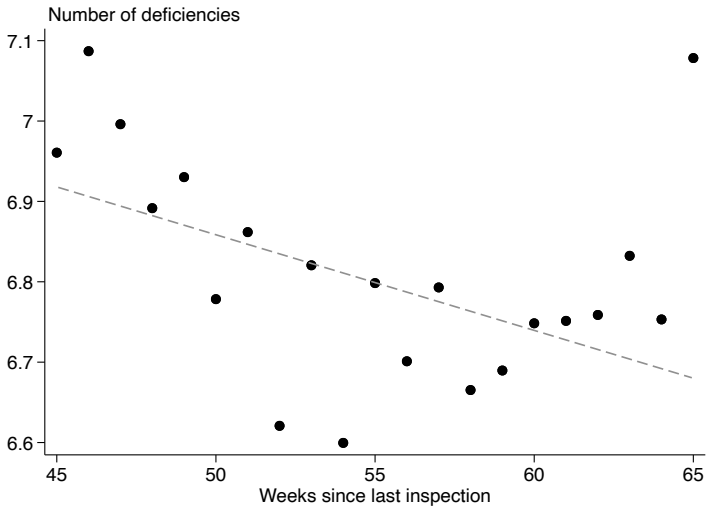
We provide two pieces of empirical evidence consistent with nursing homes not knowing their exact inspection date. First, facilities subject to earlier-than-expected inspections perform worse on those inspections. Figure B.1 shows that, after controlling for facility, state-year, and calendar-week fixed effects, shorter inspection cycles are associated with more deficiencies (panel a), and facilities whose current cycle is shorter than their previous one perform worse (panel b).

Second, facility behavior immediately around inspection dates also suggests limited foresight. Figure B.3 panel (a) plots an event study of daily staffing around the inspection start date, acquired through a Freedom of Information Act request. Staffing rises sharply during the inspection but not before, indicating that facilities are unaware of the inspection date even days in advance. Panel (b) contrasts these results with an event study centered on the publicly reported inspection end date. What appears as an "anticipatory" spike based on end dates in fact occurs during the inspection itself.

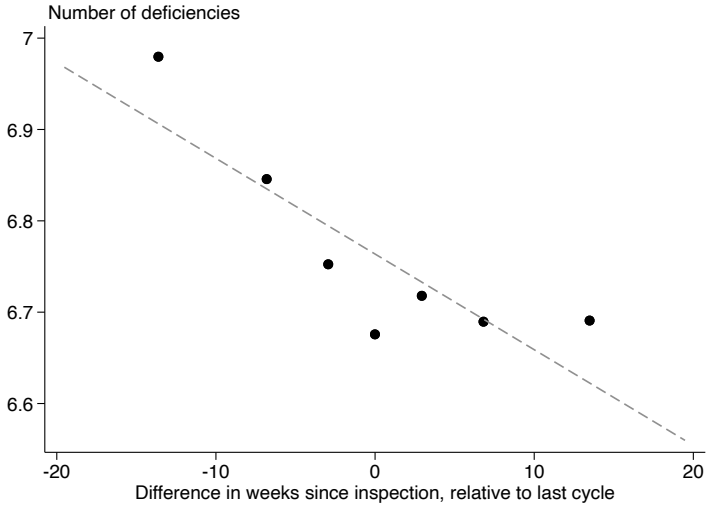
Together, these findings indicate that facilities do not know their exact inspection date ahead of time. Consequently, changes in facility behavior in the months leading up to an inspection should reflect responses to an increasing inspection hazard, rather than preparation for a known inspection date.

Figure B.1: Relationship Between Time Since Last Survey and Inspection Outcomes

(a) Deficiency Count vs. Time Since Last Inspection



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

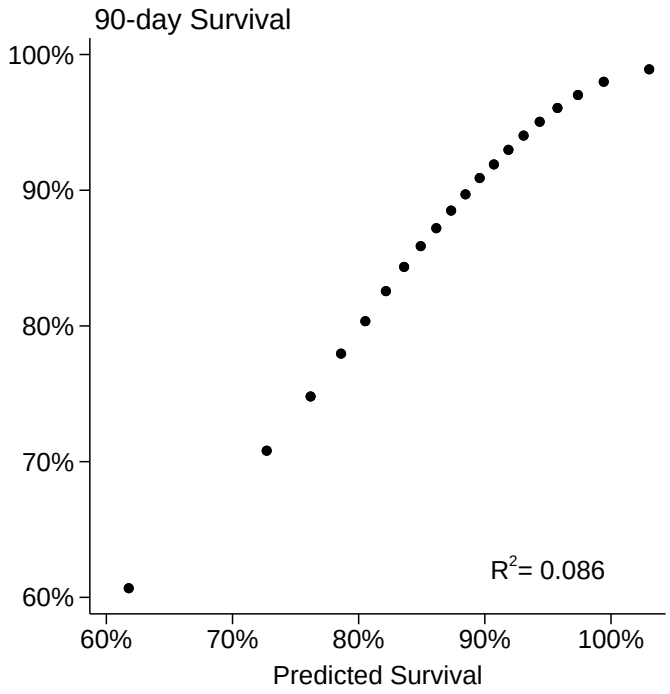


Notes: This figure plots binscatters of the relationships between the number of deficiencies in each inspection and (a) the number of weeks since a facility's last inspection, and (b) the difference in the number of weeks since the last inspection, relative to the previous inspection cycle. State-year, facility, and calendar week fixed effects have been residualized out. The sample in (a) is restricted to cycles with 45-65 weeks between inspections and the average number of deficiencies is 16.8. The coefficient on weeks since last inspection in the associated regression, clustered at the facility-level, is −0.011 (SE: 0.004). The sample in (b) is restricted to cycles where the weeks since last inspection ranges from -20 to 20 and is less than 65, and the average number of deficiencies is 16.8. The coefficient on weeks since last inspection in the associated regression, clustered at the facility-level, is −0.007 (SE: 0.002). Both samples are restricted to facilities not participating in the Special Focus Facility program.

B.2 Predicted Survival Model

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 survival 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. There is a strong monotonic relationship between actual and predicted 90-day survival (Figure B.2), and a bivariate regression of the two has an R^2 of 0.086.

Figure B.2: Predicted vs. Actual Survival



Notes: This figure plots a binscatter of predicted vs. actual 90-day survival 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.

B.3 Daily Event Studies Around Inspections and Revisits

In this section, we complement our week-level analysis with a day-level event studies around two key points in the inspection cycle: the inspection date and the end date of the last revisit. Regression equation (19) summarizes this event study design for facility j on day t:

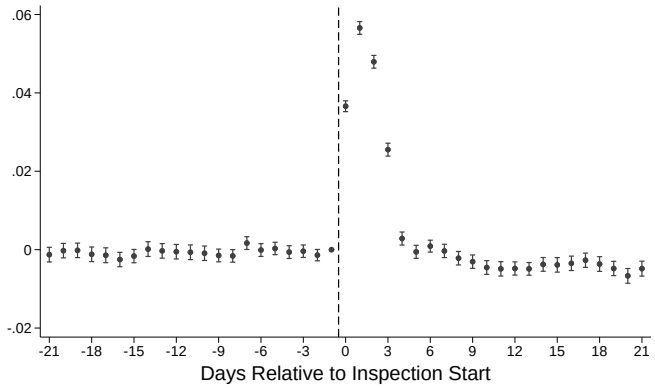
$$Y_{jt} = \sum_{r \in [-21,-1)}^{[0,21]} \beta_r \, \mathbb{1}(\text{Day Rel. Event}_{jt} = r) + \alpha_j + \lambda_t + \varepsilon_{jt}, \tag{18}$$

where β_r is the coefficient on an indicator variable for being r days away from the inspection or the end of the last revisit.

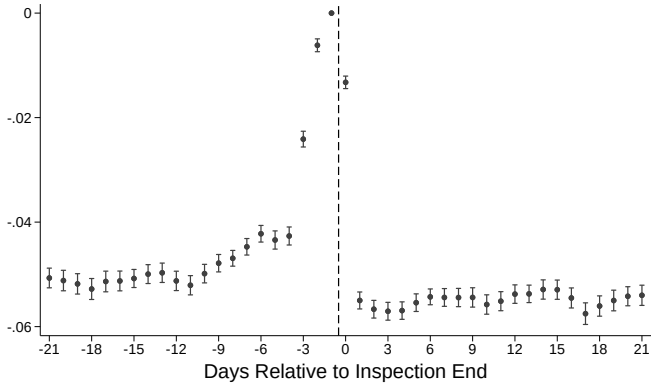
Inspection Start and End Dates Figure B.3 plots the daily event study around the inspection start and end dates. These dates often differ because inspections last, on average, 4.9 days. Panel (a) plots the coefficients in a specification centered around the inspection *start* date — this date is not typically publicly reported, and was acquired through a Freedom of Information Act request. There is no evidence of an anticipatory increase in effort; instead the spike in effort occurs during the inspection itself. It is concentrated in the first four days after the inspection starts, consistent with the average inspection duration. In contrast, panel (b) plots a daily event study around the publicly-reported inspection dates, which are inspection *end* dates.

Figure B.3: Daily Event Studies Around Inspection Start and End Dates

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



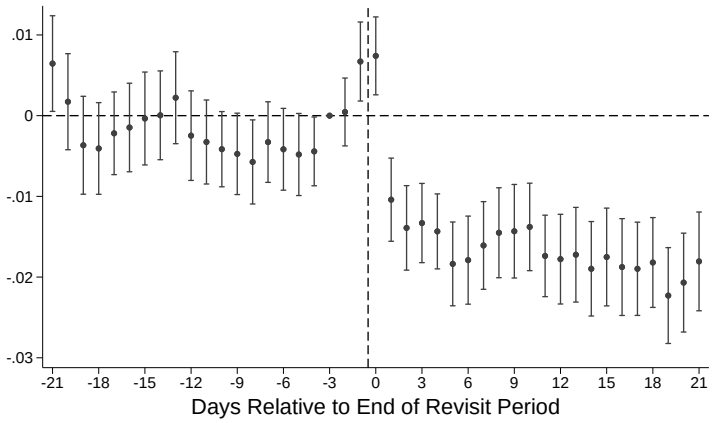
(b) Staffing Hours per Resident-Day, relative to Inspection End Date



Notes: This figure plots the estimates from equation (18), centered around (a) the inspection start date and (b) the inspection end date, for inspections in 2017-2019. Survey start dates are acquired from a Freedom of Information Act request.

Revisits Figure B.4 plots the daily event study around the end of the last revisit. Staffing is relatively elevated leading up to the last revisit and then spikes on the day(s) of the revisit, similar to the spike in effort during the initial visit observed in Figure 2a. Once the last revisit concludes, facility effort drops down to be *below* the pre-visit level. This sharp reduction in effort as soon as the likelihood of a revisit drops to zero further supports the notion that facilities choose their effort based on their perceived inspection hazard.

Figure B.4: Staffing Hours per Resident-Day, Relative to End of Last Revisit



Notes: This figure plots the estimates from equation (18), centered around the end date of a facility's last revisit within an inspection cycle, for inspections in 2017-2019.

B.4 Heterogeneity by Facility and Patient Characteristics

In this section we examine heterogeneity in both effort responses as well as patient survival, across different facility and patient characteristics.

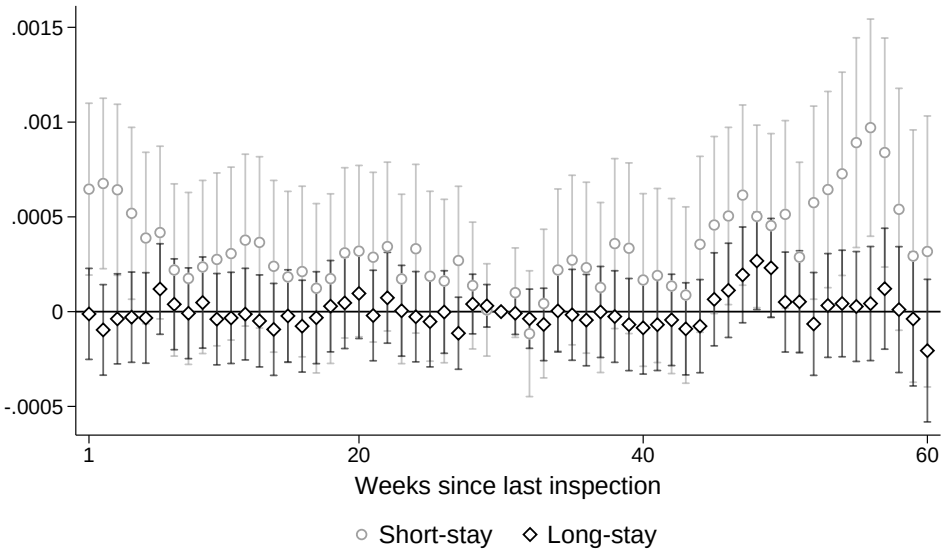
In Figure B.6, we consider heterogeneous effort responses. We first split states by the average (panel a) and variance (panel b) of their inspection cycles. Facilities in states with shorter inspection cycles tend to increase effort more as the next inspection approaches. This is consistent with these facilities perceiving a faster rise in inspection hazard as they get further from their last inspection. Facilities in states with higher-variance cycles ramp *down* effort slower after each inspection, but increase it at the end of the cycle at a similar pace as those in lower-variance states. This could be explained by high-variance states being more likely to return to a facility shortly after a previous inspection, thus keeping facilities "on their toes" more earlier in the cycle. We find no differences in the effort response by facility size (panel c), Medicaid share (panel d), or black share (panel e). For-profit facilities appear to have a slightly larger effort response. There is no difference by a facility's private equity ownership status. Facilities that are part of a chain appear to ramp down effort slower after inspections, but ramp effort back up at the same rate. Facilities that on average have lower deficiency scores (panel i) – that is, higher-quality facilities – are slightly more responsive. Finally, facilities that had a deficiency in their most recent inspection cycle (panel j) exhibit a larger effort response. Overall, the heterogeneity results indicate that facilities subject to higher inspection hazards, who tend to do well on inspections, and who may face increased pressure to perform well exhibit larger effort responses.

Turning to the effects on survival in Figure B.7, we are unable to detect heterogeneous effects on any characteristic other than by for-profit status – the survival improvements are concentrated in non-profit facilities (panel f), which is consistent with their larger effort response.

We also consider heterogeneity by patient characteristics on survival in Figure B.8.[27] The survival of patients with low predicted survival rates (panel a), with an Alzheimer's or Dementia diagnosis at admission (panel b), with a hospital stay prior to admission (panel c), and who are male (panel g) is more responsive to inspection-induced effort. However, we find few differences by Medicaid status at admission (panel d), age (panel e), and race (panel f).

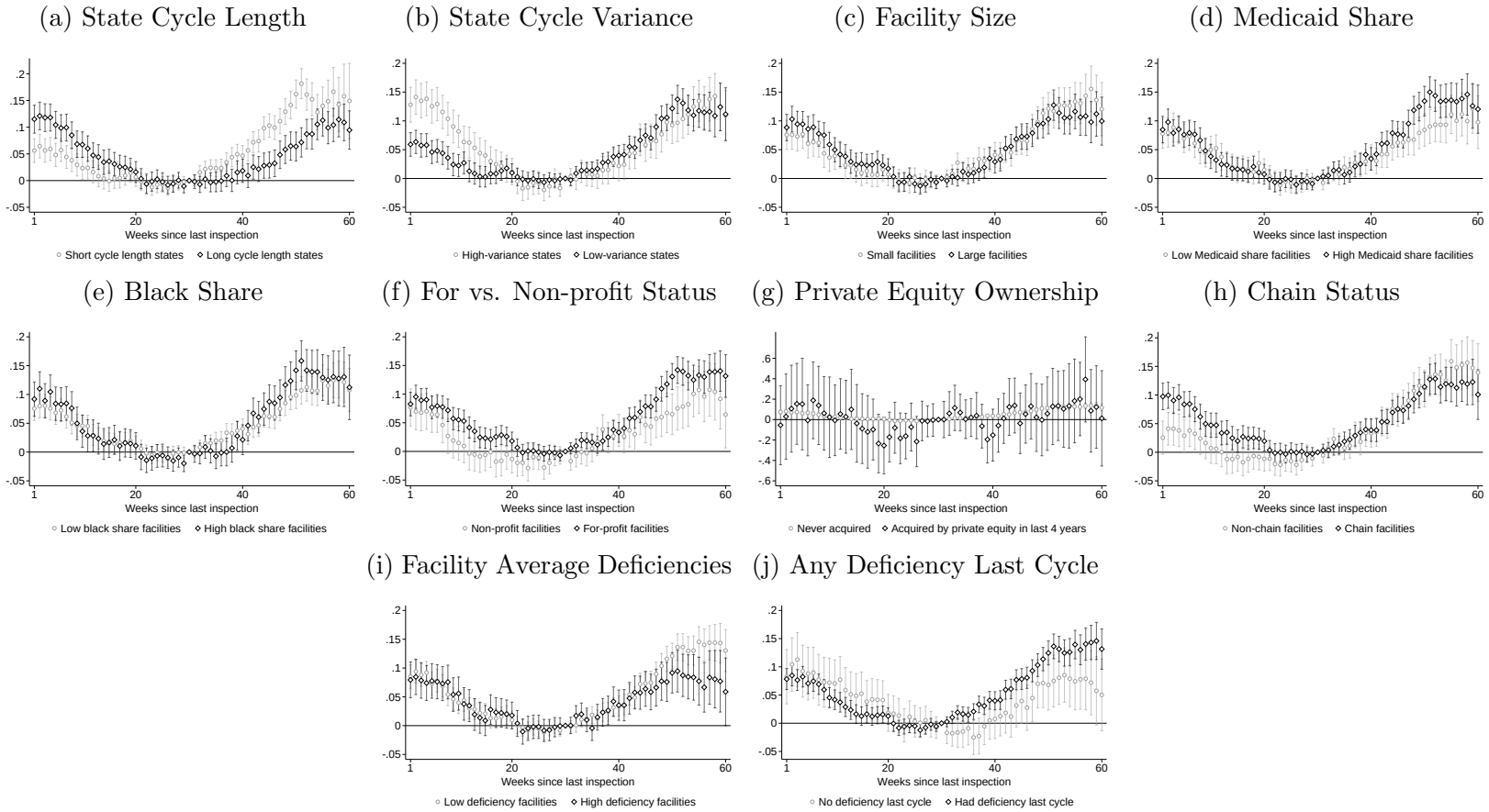
[27] We can only look at heterogeneity by patient characteristics on survival, which is at the patient-level, but not on effort, which is at the facility-level.

Figure B.5: Patient Survival by Weeks Since Last Inspection, By Week of Stay



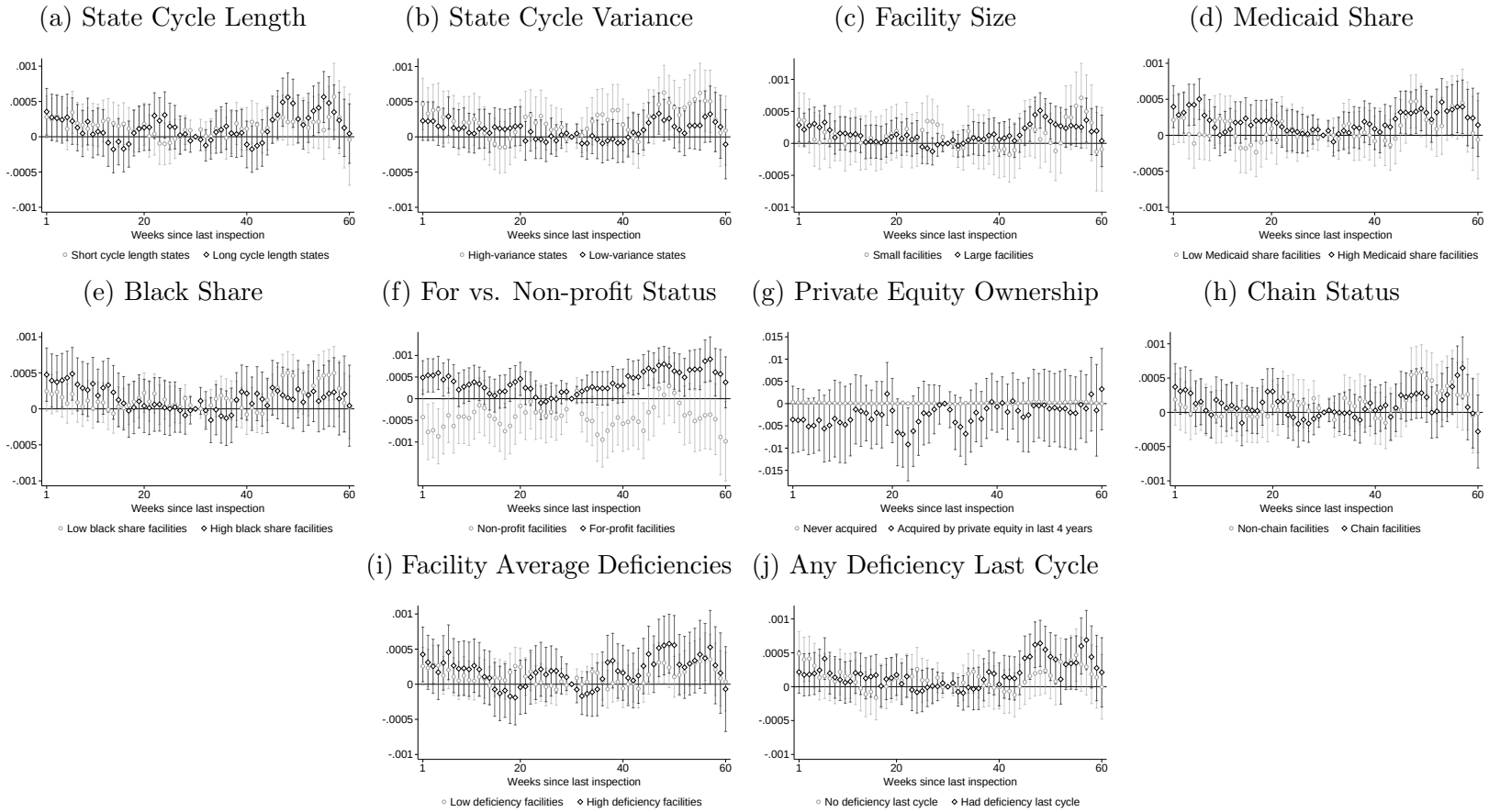
Notes: This figure plots the estimates from equation (3), split by short stays (i.e., in their first 12 weeks of the stay) or long stays (greater than 13 weeks from admission). The coefficients for week 0 are estimated but not reported.

Figure B.6: Heterogeneity in Effort by Weeks Since Last Inspection



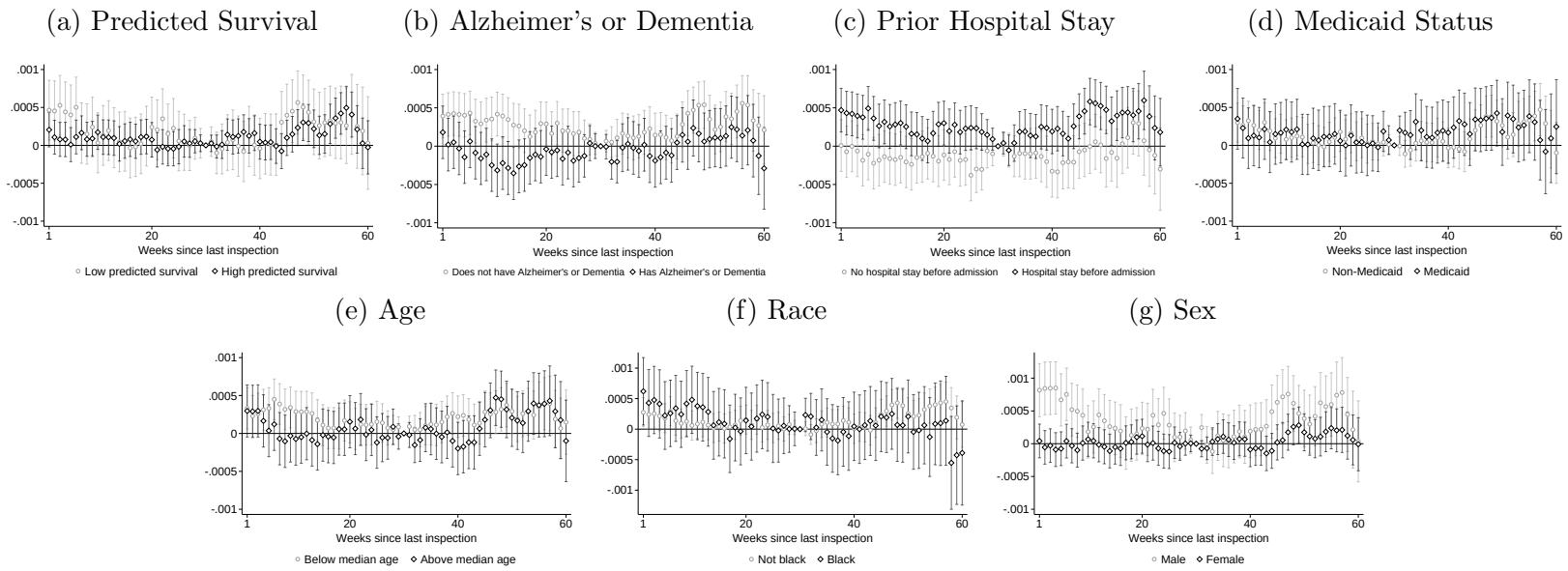
Notes: This figure plots the estimates from equation (2) on the effort composite index, split by (a) states with above vs. below median average inspection cycle length in 2017-2019, (b) states with above vs. below median inspection cycle variance in 2017-2019, (c) facilities with above vs. below-median bed counts in 2017, (d) facilities with above vs. below average Medicaid shares (with respect to their 2017 state average), (e) facilities with above vs. below average black shares (with respect to their 2017 state average), (f) for-profit vs. non-profit facilities, (g) private equity ownership status, (h) chain affiliation (i) facilities with above vs. below average deficiency scores (with respect to their state 2017 average), and (j) inspection cycles with and without a deficiency in the last cycle. For-profit status, private equity ownership status, and chain affiliation status are defined at the facility week level. The coefficients for week 0 are estimated but not reported.

Figure B.7: Heterogeneity in Survival by Weeks Since Last Inspection



Notes: This figure plots the estimates from equation (3) on four week survival, split by (a) states with above vs. below median average inspection cycle length in 2017-2019, (b) states with above vs. below median inspection cycle variance in 2017-2019, (c) facilities with above vs. below-median bed counts in 2017, (d) facilities with above vs. below average Medicaid shares (with respect to their 2017 state average), (e) facilities with above vs. below average black shares (with respect to their 2017 state average), (f) for-profit vs. non-profit facilities, (g) private equity ownership status, (h) chain affiliation, (i) facilities with above vs. below average deficiency scores (with respect to their state 2017 average), and (j) inspection cycles with and without a deficiency in the last cycle. For-profit status, private equity ownership status, and chain affiliation status are defined at the facility week level. The coefficients for week 0 are estimated but not reported.

Figure B.8: Heterogeneity in Survival Based on Patient Characteristics by Weeks Since Last Inspection



Notes: This figure plots the estimates from equation (3) on four week survival, split by (a) stays with individuals below and above the median predicted survival rate in 2013-2019, (b) stays with individuals with and without Alzheimer's disease or dementia at admission, (c) whether an individual stayed in a hospital within 1 month prior to admission to a facility, (d) Medicaid status, (e) stays with individuals below and above median age, (f) Black vs non-Black patients, and (g) sex. The coefficients for week 0 are estimated but not reported.

B.5 Additional Health Outcomes

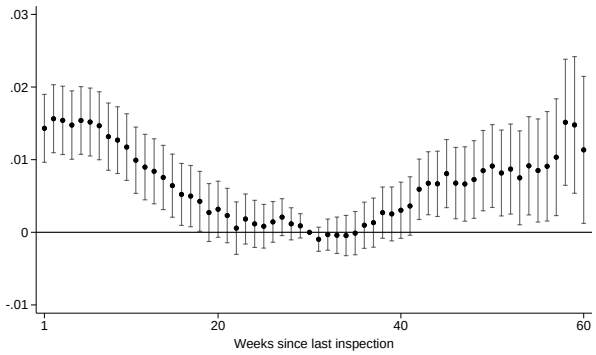
In this appendix, we extend the analysis of patient health outcomes to consider non-survival measures. As with our analysis of patient survival in the main text, each outcome corresponds to estimates from equation (3).

Figure B.9 considers independence with the activities of daily living (ADL). ADLs are measures of a resident's independence in performing tasks related to personal care, including transfers, 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. The activity scores are summed to construct a "composite" ADL score. Because higher scores indicates worse health, we negate them to report independence in ADLs. The negated ADL scores show similar patterns as effort and survival — it is lowest in the middle of the inspection cycle, when nursing homes are subject to the least oversight. This is true in both composite form (panel a), as well as in each component measure: transferring (panel b), bed mobility (panel c), toileting (panel d), and eating (panel e).

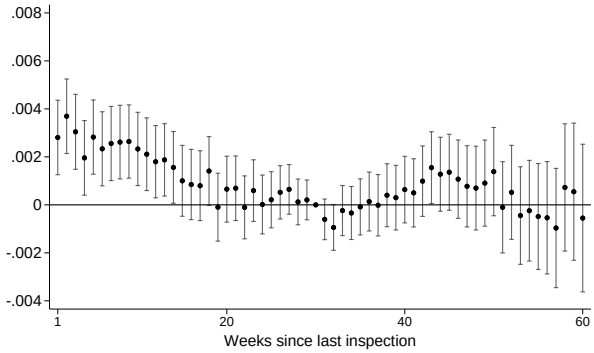
Figure B.10 presents patterns for other health outcomes. We find a similar, albeit more muted patterns when assessing the rates of pressure ulcers (panel a), another common nursing home quality measure. We find very slight evidence for increases in the share of patients discharged to a hospital (panel b). This could be because facilities putting in more effort are able to recognize patient deterioration more readily, or it could reflect selective discharge. While the former reflects effort that improves patient health, the latter reflects cream-skimming behavior that does not.

Figure B.9: Independence with the Activities of Daily Living

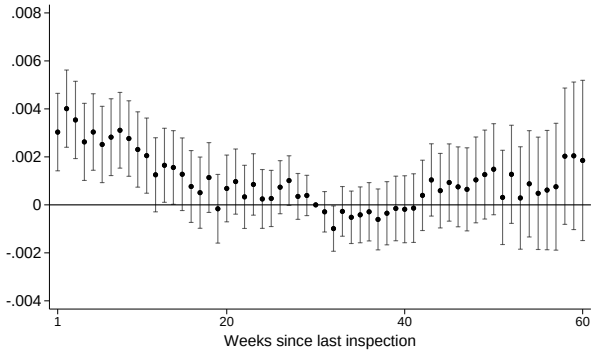
(a) Composite Score



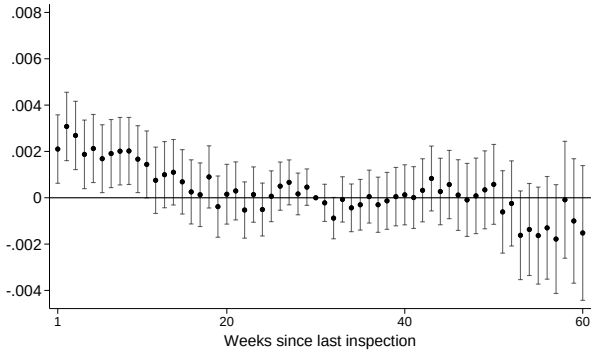
(b) Transfer Score



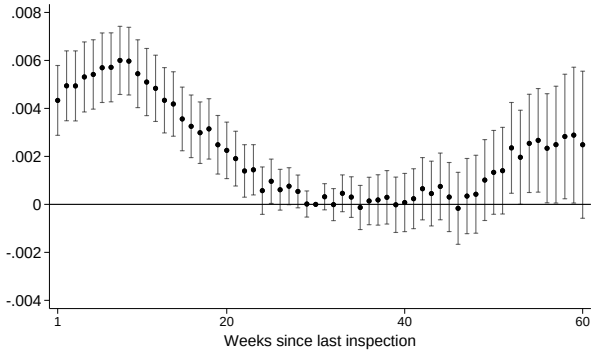
(c) Bed Mobility Score



(d) Toilet Use Score



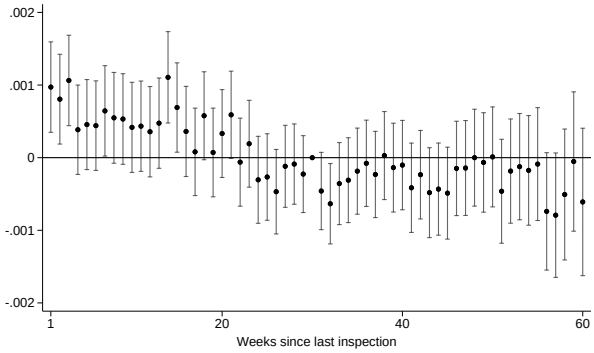
(e) Eating Score



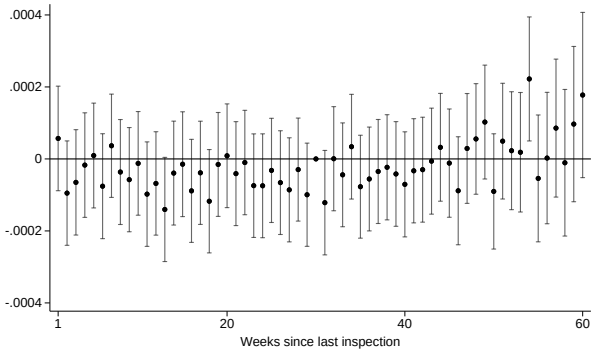
Notes: This figure plots the coefficients from estimating equation (3) on health outcomes measured in the MDS or MEDPAR for patients in the facility in a given inspection week. All outcomes are normalized so that a higher value is associated with a better outcome. Panel (a) reports the composite ADL score, which is the sum of the component scores. Panel (b) reports the transferring component score. Panel (c) reports the bed mobility component score. Panel (d) reports the toileting component score. Panel (e) reports the eating component score. The coefficients for week 0 are estimated but not reported.

Figure B.10: Additional Health Outcomes by Weeks Since Last Inspection

(a) Patients Without Pressure Ulcers



(b) Share of Patients Discharged to Hospital



Notes: This figure plots the coefficients from estimating equation (3) on health outcomes measured in the MDS or MEDPAR 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 if patients are without pressure ulcers, and panel (b) plots the share of patients with a discharge to the hospital (including both return anticipated and not anticipated). The coefficients for week 0 are estimated but not reported.

B.6 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 (19) summarizes this stacked event study design for facility j in calendar week t and index inspection event c:

$$Y_{jct} = \sum_{r \in [-30,-1)}^{[0,30]} \beta_r \; \mathbb{1}(\text{Week Rel. Inspection}_{jct} = r) + \alpha_j + \lambda_t + \delta_c + \varepsilon_{jct}, \tag{19}$$

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 (2) 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 (3), we also consider the corresponding specification for patient i:

$$Y_{ict} = \beta_0 + \sum_{r \in [0,30)}^{(30,60]} \beta_r \; \mathbb{1}(\text{Week Rel. Inspection}_{j(i),c(i),t} = r) + \gamma X_i + \alpha_{j(i)} + \lambda_t + \delta_{c(i)} + \varepsilon_{ict} \tag{20}$$

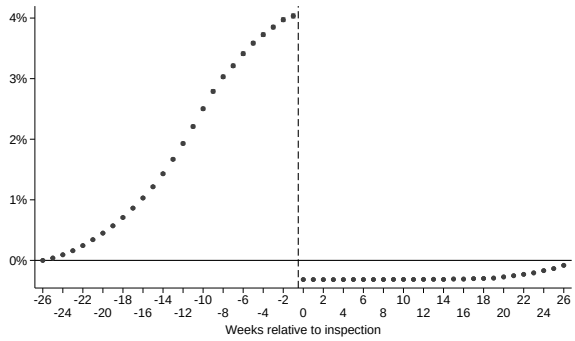
where, as before, X_i controls for patient i's predicted mortality at admission.

Results. Our findings are consistent with the results from Section 4. Appendix Figure B.11 plots the event studies results for the inspection hazard, facility effort, 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.[28] 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).

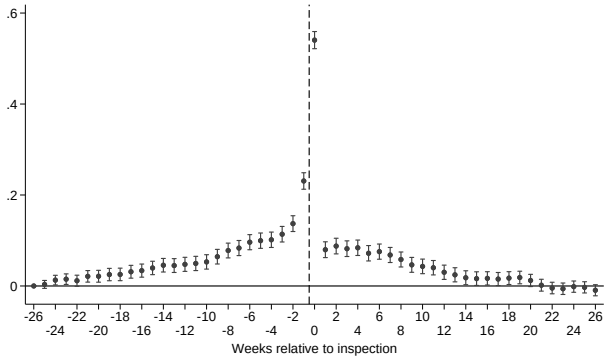
[28] Appendix Figure B.3 plots daily event studies in the 21 days before and after the inspection, centered around either the start date (a), which is available for a subset of years through FOIA request, or the end date (b), which is available for our entire sample. These indicate that the staffing spike begins once the inspection has begun.

Figure B.11: Event Studies: Inspection Hazard, Facility Effort, and Patient Survival

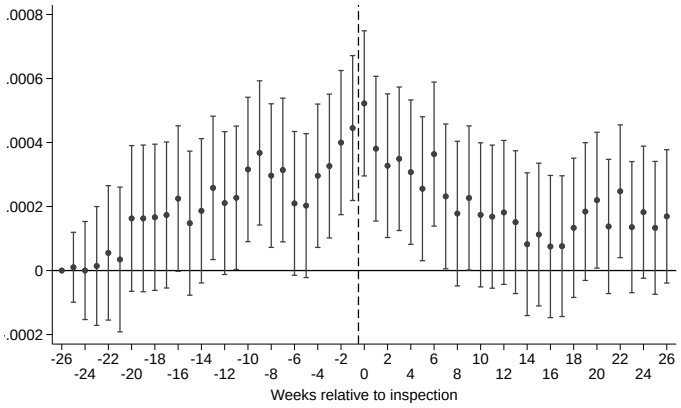
(a) Inspection Hazard



(b) Facility Effort



(c) 4-Week Survival Rate



Notes: This figure plots the estimates from equations (19) and (20) on (a) the inspection hazard rate, (b) standardized facility effort, and (c) 4-week patient survival rate. The effort index is a composite measure combining labor and non-labor inputs, and is described in further detail in Section 3.1 and Appendix Section A.2.

C Estimation of Inspection Performance Parameters

C.1 Derivation of Distributions

Our estimation routine makes use of the long-run distributions for several random variables for ease in simulation. In this section, we derive these distributions, and in doing so, provide a proof for Theorem 2.

Starting from the definition of latent quality in equation (11), notice we can express $q_{j,t}$ as a function of any prior value $q_{j,t-\Delta}$ for any arbitrary lag Δ:

$$\begin{aligned}
q_{j,t} &= \varphi_j + \rho q_{j,t-1} + \epsilon_{j,t} \\
&= \varphi_j + \rho\left(\varphi_j + \rho q_{j,t-2} + \epsilon_{j,t-1}\right) + \epsilon_{j,t} \\
&= \varphi_j + \rho\left(\varphi_j + \rho\left(\varphi_j + \rho q_{j,t-3} + \epsilon_{j,t-2}\right) + \epsilon_{j,t-1}\right) + \epsilon_{j,t} \\
&= \varphi_j \sum_{k=0}^{\Delta-1} \rho^k + \rho^{\Delta} q_{j,t-\Delta} + \sum_{k=0}^{\Delta-1} \rho^k \epsilon_{j,t-k} \\
&= \varphi_j \frac{1-\rho^{\Delta}}{1-\rho} + \rho^{\Delta} q_{j,t-\Delta} + \sum_{k=0}^{\Delta-1} \rho^k \epsilon_{j,t-k}. \qquad (21)
\end{aligned}$$

As each $\epsilon_{j,t-k}$ is i.i.d. normal with variance σ_ϵ^2, it follows that the conditional distribution of $q_{j,t} \mid q_{j,t-\Delta}, \varphi_j$ is similarly normal:

$$q_{j,t} | q_{j,t-\Delta}, \varphi_j \sim \mathcal{N}\left(\frac{1-\rho^{\Delta}}{1-\rho}\varphi_j + \rho^{\Delta} q_{j,t-\Delta}, \frac{1-\rho^{2\Delta}}{1-\rho^2}\sigma_\epsilon^2\right)$$

where the variance term follows from the finite geometric sum of the non-deterministic term in equation (21):

$$\text{Var}\left(\sum_{k=0}^{\Delta-1} \rho^k \epsilon_{j,t-k}\right) = \sum_{k=0}^{\Delta-1} \rho^{2k}\, \sigma_\epsilon^2 = \sigma_\epsilon^2 \sum_{k=0}^{\Delta-1} \rho^{2k} = \sigma_\epsilon^2 \frac{1-\rho^{2\Delta}}{1-\rho^2} \quad \text{for } |\rho| < 1.$$

To consider the long-run distribution, we simply let $\Delta \to \infty$:

$$q_{j,t} | \varphi_j \sim \mathcal{N}\left(\frac{1}{1-\rho}\varphi_j, \frac{\sigma_\epsilon^2}{1-\rho^2}\right)$$

Because the conditional distribution $q_{j,t} \mid \varphi_j$ is normal, and φ_j is normal, it follows that the

marginal distribution of $q_{j,t}$ is also normal:

$$q_{j,t} \sim \mathcal{N}\left(0, \frac{\sigma_\varphi^2}{(1-\rho)^2} + \frac{\sigma_\epsilon^2}{1-\rho^2}\right). \tag{22}$$

We will employ the following useful shorthands in the remaining derivations:

$$\sigma_q^2 \equiv \operatorname{Var}(q_{j,t}), \qquad V_\varphi \equiv \frac{\sigma_\varphi^2}{(1-\rho)^2}, \qquad V_\epsilon \equiv \frac{\sigma_\epsilon^2}{1-\rho^2}.$$

The variance of $q_{j,t}$ can be expressed as the sum of two components:

$$\sigma_q^2 = V_\varphi + V_\epsilon$$

where V_φ represents the long-run static cross-facility variation, and V_ϵ captures long-run within-facility variation in quality.

We have so far derived the marginal distribution of $q_{j,t}$, as well as the conditional distributions $q_{j,t} \mid q_{j,t-\Delta}, \varphi_j$ and $q_{j,t}|\varphi_j$. Because our estimation procedure considers the covariance between subsequent inspections as a set of moments, we still require the distribution $q_{j,t}|q_{j,t-\Delta}$.

Notice that because (22) applies to any arbitrary t, $q_{j,t-\Delta}$ is likewise normal. Because φ_j is normal and $q_{j,t-\Delta}$ is normal conditional on φ_j, it follows that φ_j and $q_{j,t-\Delta}$ are bivariate normal. A consequence of this is that:

$$\varphi_j = \underbrace{\mathbb{E}[\varphi_j]}_{0} + \frac{\operatorname{Cov}(\varphi_j, q_{j,t-\Delta})}{\operatorname{Var}(q_{j,t-\Delta})}\left(q_{j,t-\Delta} - \underbrace{\mathbb{E}[q_{j,t-\Delta}]}_{0}\right) + v_{jt-\Delta}, \tag{23}$$

where $\mathbb{E}[v_{jt-\Delta} q_{j,t-\Delta}] = 0$.[29] Note further that:

$$\begin{aligned}
\operatorname{Cov}(\varphi_j, q_{j,t-\Delta}) &= \operatorname{Cov}\left(\varphi_j, \frac{\varphi_j}{1-\rho} + \sum_{\tau=-\infty}^{t-\Delta} \rho^{t-\Delta-\tau}\epsilon_{j\tau}\right) \\
&= \frac{1}{1-\rho}\sigma_\varphi^2 + \sum_{\tau=-\infty}^{t-\Delta} \rho^{t-\Delta-\tau} \underbrace{\operatorname{Cov}(\varphi_j, \epsilon_{j\tau})}_{0} \\
&= \frac{1}{1-\rho}\sigma_\varphi^2 = (1-\rho)V_\varphi
\end{aligned} \tag{24}$$

[29] In fact, $v_{jt-\Delta} \sim \mathcal{N}\left(0, \left(1 - \operatorname{corr}(\varphi_j, q_{j,t-\Delta})^2\right) \operatorname{Var}(\varphi_j)\right)$.

Finally, we can now express the covariance between current and lag quality, $\text{Cov}(\, q_{j,t}, q_{j,t-\Delta}\,)$:

$$\begin{aligned}
\text{Cov}(q_{j,t}, q_{j,t-\Delta}) &= \text{Cov}\Big(\varphi_j \frac{1-\rho^{\Delta}}{1-\rho} + \rho^{\Delta} q_{j,t-\Delta} + \sum_{k=0}^{\Delta-1} \rho^k \epsilon_{j,t-k},\ q_{j,t-\Delta}\Big) \\
&= \frac{1-\rho^{\Delta}}{1-\rho} \underbrace{\text{Cov}(\varphi_j, q_{j,t-\Delta})}_{(1-\rho)\, V_\varphi} + \rho^{\Delta} \underbrace{\text{Var}(q_{j,t-\Delta})}_{\sigma_q^2 = V_\varphi + V_\epsilon} + 0 \\
&= (1-\rho^{\Delta})\, V_\varphi + \rho^{\Delta}\, (V_\varphi + V_\epsilon) \\
&= V_\varphi + \rho^{\Delta} V_\epsilon.
\end{aligned}$$

which is sufficient to determine the variance-covariance matrix of the distribution described in Theorem 2.

C.2 Moments

Because the facility-level accumulated effort measure $s_{j,t}$ is only observed for 2017–2019, we impute effort in all years using the *average* accumulated effort $\bar{s}_\Delta$ associated with each inspection cycle length Δ. This substitution makes $\bar{s}_\Delta$ a deterministic function of Δ, so that all deficiency indicators are conditionally independent given $(q_{j,t}, \Delta)$, and expectations are taken only over the distribution of latent quality $q_{j,t}$ (and Δ, when applicable).

We employ three sets of moments. The first set contains the conditional mean deficiency rates for each cycle length Δ: $\mathbb{E}[x^d_{j,t} \mid \Delta] = \bar{x}^d_\Delta$, described in Theorem 1.

The second set includes the (unconditional) cross-sectional covariances between each deficiency d and the sum of all *other* deficiencies $d' \neq d$. That is, we include for each deficiency d:

$$\text{Cov}\left(x^d_{j,t}, \sum_{d' \neq d} x^{d'}_{j,t}\right) = \sum_{d' \neq d} \text{Cov}\left(x^d_{j,t}, x^{d'}_{j,t}\right)$$

For D possible deficiencies, this yields a vector of D moments. Note that $x^d_{j,t}$ and $x^{d'}_{j,t}$ are independent Bernoulli draws conditional on $q_{j,t}$ and the cycle length Δ. That is, $q_{j,t}$ is the only random term that drives any covariance between deficiencies drawn from the same inspection.

The third set of moments include the (conditional) autocovariances of deficiencies across subsequent inspections. These moments capture the time series variation of deficiencies:

$$\text{Cov}\left(x^d_{j,t}, \sum_{f} x^f_{j,t-\Delta} \mid \Delta\right) = \sum_{f} \text{Cov}\left(x^d_{j,t}, x^f_{j,t-\Delta} \mid \Delta\right)$$

where Δ denotes the number of periods elapsed between the inspection at time t and the most recent prior inspection. For $|\mathcal{G}|$ unique inspection gaps, this grouping adds a further $|\mathcal{G}| \times D$ moments.

Constructing model-implied analogues of these moments is straightforward, given the long-run distributions derived in Appendix Section C.1. For a guess of parameters, we draw $(q_{j,t}, q_{j,t-\Delta})$ from a bivariate normal distribution, where $\text{Var}(q_{j,t}) = \text{Var}(q_{j,t-\Delta}) = \sigma_q^2$ and $\text{Cov}(q_{j,t}, q_{j,t-\Delta}) = V_\varphi + \rho^\Delta V_\epsilon$.[30] We can accordingly construct deficiency probabilities using the normal cumulative distribution function and the observed (fixed) accumulated efforts $p^d_{jt} = \Phi\big(\kappa_d + \gamma_d(q_{j,t} + \lambda \bar{s}_\Delta)\big)$; likewise with $p^d_{jt-\Delta}$. These model-implied probabilities are sufficient to construct each of our moments.

Notice that we can express the second set of moments (the cross-sectional covariances) in terms of these probabilities:

$$\begin{aligned}
\text{Cov}\big(x^d_{j,t}, x^{d'}_{j,t}\big) &= \mathbb{E}\big[x^d_{j,t} x^{d'}_{j,t}\big] \;-\; \mathbb{E}\big[x^d_{j,t}\big]\,\mathbb{E}\big[x^{d'}_{j,t}\big] \\
&= \mathbb{E}_q\Big[\,\mathbb{E}\big[x^d_{j,t} x^{d'}_{j,t} \mid q_{j,t}\big]\Big] \;-\; \mathbb{E}_q\Big[\,\mathbb{E}\big[x^d_{j,t} \mid q_{j,t}\big]\Big]\,\mathbb{E}_q\Big[\,\mathbb{E}\big[x^{d'}_{j,t} \mid q_{j,t}\big]\Big] \\
&= \mathbb{E}_q\big[p^d_{jt}\, p^{d'}_{jt}\big] \;-\; \mathbb{E}_q\big[p^d_{jt}\big]\,\mathbb{E}_q\big[p^{d'}_{jt}\big]
\end{aligned}$$

which follows from the law of iterated expectations and the conditional independence of $\mathbf{x}_{j,t}$ given $q_{j,t}$. A similar derivation applies to $\text{Cov}(x^d_{j,t}, x^d_{j,t-\Delta})$.

C.3 Proof of Theorem 1

Here we prove the invertibility of the model-implied gap-conditional deficiency rate $\mathbb{E}[x^d_{j,t} \mid \Delta]$ stated in Theorem 1.

Notice the following: for a given deficiency d and realized quality $q_{j,t}$, we can express the model-implied probability of a deficiency as:

$$\mathbb{E}[x^d_{j,t} \mid \Delta] = \mathbb{E}_q[\mathbb{E}[x^d_{j,t}|q_{j,t}, \Delta]] = \mathbb{E}_q[p^d_{jt} \mid \Delta] = \mathbb{E}_q[\Phi\big(\kappa_d + \gamma_d(q_{j,t} + \lambda \bar{s}_\Delta)\big)]$$

where $\mathbb{E}_q[\cdot]$ denotes the expectation over latent quality $q_{j,t}$, which recall has the following distribution:

$$q_{j,t} \sim \mathcal{N}\left(0, \frac{\sigma^2_\varphi}{(1-\rho)^2} + \frac{\sigma^2_\epsilon}{1-\rho^2}\right) \equiv \mathcal{N}(0, \sigma^2_q)$$

[30] Note that we hold the draws fixed across candidate parameter values, scaling the draws via a Cholesky decomposition. Doing so allows us to excise simulation noise from the objective function.

By definition:

$$\begin{aligned}\Phi(\kappa_d + \gamma_d(q_{j,t} + \lambda\bar{s}_\Delta) &= \Pr\big(Z \le \kappa_d + \gamma_d(q_{j,t} + \lambda\bar{s}_\Delta) \mid q_{j,t}\big)\\ \mathbb{E}_q[\Phi(\kappa_d + \gamma_d(q_{j,t} + \lambda\bar{s}_\Delta))] &= \Pr\big(Z \le \kappa_d + \gamma_d(q_{j,t} + \lambda\bar{s}_\Delta)\big)\\ &= \Pr(Z - \gamma_d q_{j,t} \le \kappa_d + \gamma_d\lambda\bar{s}_\Delta)\end{aligned}$$

for some standard normal $Z \sim \mathcal{N}(0,1)$ that is independent of $q_{j,t}$. Because Z and $q_{j,t}$ are both independent normals, their linear combination $Y := Z - \gamma_d q_{j,t}$ is also normal, with $\mathbb{E}[Y] = 0$ and

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{Var}(Z) + \mathrm{Var}(\gamma_d q_{j,t})\\ &= 1 + \gamma_d^2\sigma_q^2\end{aligned}$$

Hence, $Y \sim \mathcal{N}(0, 1 + \gamma_d^2\sigma_q^2)$, and so:

$$\Pr(Y \le \kappa_d + \gamma_d\lambda\bar{s}_\Delta) = \Phi\left(\frac{\kappa_d + \gamma_d\lambda\bar{s}_\Delta}{\sqrt{1 + \gamma_d^2\sigma_q^2}}\right)$$

Equivalently:

$$\mathbb{E}[x_{j,t}^d \mid \Delta] = \mathbb{E}_q\left[\Phi\big(\kappa_d + \gamma_d(q_{j,t} + \lambda\bar{s}_\Delta)\big)\right] = \Phi\left(\frac{\kappa_d + \gamma_d\lambda\bar{s}_\Delta}{\sqrt{1 + \gamma_d^2\sigma_q^2}}\right)$$

Let the sample mean $\bar{x}_\Delta^d$ denote the point estimator for $\mathbb{E}[x_{j,t}^d \mid \Delta]$. Plugging this in and rearranging yields:

$$\Phi^{-1}(\bar{x}_\Delta^d) \times \sqrt{1 + \gamma_d^2\sigma_q^2} = \kappa_d + \lambda(\gamma_d\bar{s}_\Delta) \tag{25}$$

Equation (25) shows that the empirical deficiency rate $\bar{x}_\Delta^d$ at each gap Δ identifies a linear combination of the parameters $(\kappa_d, \gamma_d, \lambda)$, after accounting for σ_q^2. Since $\bar{s}_\Delta$ varies with Δ but not across deficiencies, we can stack this equation across all (d, Δ) pairs to form a linear system that identifies the deficiency-specific intercept κ_d and the common effort-loading parameter λ, given a value for the remaining parameters $(\rho, \sigma_\varphi^2, \sigma_\epsilon^2, \gamma_d)$.

C.4 Estimation Procedure

Our sample of inspection results includes indicators for up to 205 possible deficiencies. Each deficiency requires its own slope on quality and effort (γ_d) and intercept (κ_d). While we are able to concentrate out κ_d, via Theorem 1, estimating the vector of γ_d parameters is

computationally challenging. We employ a two-step procedure to relieve this computational burden.

In the first step, we jointly estimate all non-deficiency-specific parameters $\{\lambda, \sigma_\varphi, \sigma_\epsilon, \rho, \delta\}$ and the deficiency-specific parameters $\{\gamma_d, \kappa_d\}$ for the 5 most common deficiencies documented in Appendix Table G.1. To do so, we perform a grid search over δ, searching over $\{\gamma_d, \sigma_\varphi, \sigma_\epsilon, \rho\}$ to minimize the GMM objective function, while concentrating out λ and κ_d as described above. We select the optimal δ as the value that minimizes the loss function.

In the second step, we hold these estimates fixed and iteratively estimate the remaining deficiency-specific parameters $\{\gamma_d, \kappa_d\}$. We add each marginal deficiency d one at a time to our initial set. Some deficiencies fail to converge at plausible parameter values; we exclude these from the remaining analysis. Our final set of estimates includes $D = 160$ total deficiencies. These selected deficiencies cover 98.9% of all deficiency citations issued during our sample.

C.5 Computing Expected Uncertainty

We characterize the regulator's uncertainty over quality as the conditional variance of quality given the most recently observed inspection outcomes $\mathbf{x}_{j,t-w}$ (which occurred w periods ago) and the static component of quality, φ_j: $\text{Var}(q_{j,t}|\mathbf{x}_{j,t-w}, \varphi_j)$. We condition on the static component of quality, φ_j, as in the long-run a regulator will learn this term under *any* inspection regime. We summarize this uncertainty by taking the expectation over φ_j and all possible prior inspection outcomes for a given w: $\mathbb{E}\left[\text{Var}(q_{j,t}|\mathbf{x}_{j,t-w}, \varphi_j)\right]$.

We can derive a convenient expression for this term by applying the law of total variance to the unconditional variance of quality:

$$\begin{aligned} \text{Var}(q_{j,t}) &= \mathbb{E}[\text{Var}(q_{j,t}|\mathbf{x}_{j,t-w}, \varphi_j)] + \text{Var}(\mathbb{E}[q_{j,t}|\mathbf{x}_{j,t-w}, \varphi_j]) \\ \implies \mathbb{E}[\text{Var}(q_{j,t}|\mathbf{x}_{j,t-w}, \varphi_j)] &= \text{Var}(q_{j,t}) - \text{Var}(\mathbb{E}[q_{j,t}|\mathbf{x}_{j,t-w}, \varphi_j]) \end{aligned} \tag{26}$$

Notice that $\text{Var}(q_{j,t}) = \sigma_q^2$ is derived above. We can compute $\text{Var}(\mathbb{E}[q_{j,t}|\mathbf{x}_{j,t-w}, \varphi_j])$ via simulation. Applying the law of iterated expectations to the inner term:

$$\mathbb{E}\left[q_{j,t} \,\middle|\, \mathbf{x}_{j,t-w}, \varphi_j\right] = \mathbb{E}\left[\mathbb{E}\left[q_{j,t} \,\middle|\, q_{j,t-w}\right] \,\middle|\, \mathbf{x}_{j,t-w}, \varphi_j\right] \tag{27}$$

From Appendix Section C.1 we have:

$$\mathbb{E}\big[q_{j,t} \mid \mathbf{x}_{j,t-w}, \varphi_j\big] = \underbrace{\left(\rho^w + \frac{V_\varphi}{\sigma_q^2}\left(1 - \rho^w\right)\right)}_{\beta_w} \mathbb{E}\big[q_{j,t-w} \mid \mathbf{x}_{j,t-w}, \varphi_j\big],$$

Plugging this into equation (26) gives

$$\mathbb{E}\left[\text{Var}(q_{j,t} \mid \mathbf{x}_{j,t-w}, \varphi_j)\right] = \sigma_q^2 - \beta_w^2 \,\text{Var}\big(\mathbb{E}[q_{j,t-w} \mid \mathbf{x}_{j,t-w}, \varphi_j]\big)$$

with β_w defined as above.

All that is left is to compute is $\text{Var}(\mathbb{E}[q_{j,t}|\mathbf{x}_{j,t}, \varphi_j])$. Recognizing that $\text{Var}(\mathbb{E}[q_{j,t-w}|\mathbf{x}_{j,t-w}, \varphi_j]) = \text{Var}(\mathbb{E}[q_{j,t}|\mathbf{x}_{j,t}, \varphi_j])$, we can do so easily via simulation: first draw N latent qualities from the known marginal distribution $q \sim \mathcal{N}(0, \sigma_q^2)$ and cycle lengths Δ from the empirical distribution. For each q_i, we can compute $p_i^d = \Phi(\kappa_d + \gamma_d(q_i + \lambda \bar{s}_\Delta))$, and simulate the vector of deficiencies where $x_i^d \sim \text{Bernoulli}(p_i^d)$. We can then take each unique deficiency vector $\mathbf{x}$, compute $\mathbb{E}[q_i|\mathbf{x}_i, \varphi_i]$, and aggregate across unique $\mathbf{x}_i$ and φ_i to compute $\text{Var}(\mathbb{E}[q|\mathbf{x}, \varphi])$. In practice, we use $N = 1,000,000$, though the results are not sensitive to this choice.

For interpretability, we express "uncertainty" over latent quality as this variance term scaled by the unconditional variance of quality:

$$U_w = \mathbb{E}\big[\text{Var}(q_{j,t} \mid \mathbf{x}_{j,t-w}, \varphi_j)\big]/\sigma_q^2$$

This uncertainty measure does not asymptote to 100% as $w \to \infty$, i.e. as cycle lengths grow arbitrarily long. This is because there is a time-invariant facility-specific component of quality: the intercept φ_j in equation (11), which limits the extent to which $\text{Var}(q_{j,t} \mid \mathbf{x}_{j,t-w})$ can converge to the unconditional variance σ_q^2.

D Alternative Model Estimates

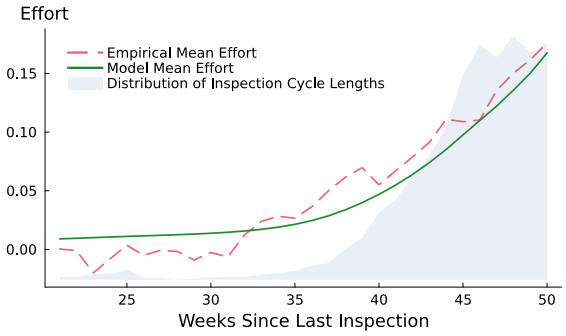
States can vary widely in their inspection timing. This means that the hazard function $h(w)$ may have different shape across different states. Furthermore, our reduced form results show that the effort response varies by state cycle length (Figure B.6a). Accordingly, in this appendix we consider the results from estimating our model when pooling together states with similar inspection schedules. Estimating separately may improve the fit of our model, as we do not need to force $h(w)$ to be the same across states. First, we pool states by their relative frequency of inspections (Figure D.1). States that have a relative high frequency of inspections are termed "fast" states, whereas states that have less frequent inspections are labeled "slow" states. We identify states as "fast" or "slow" depending whether their mean inspection frequencies is above or below the 50th percentile across states. We also examine the results from estimating our model on a single large state, California (Figure D.2). We find results that are very consistent with those reported in the main text.

Figure D.1: Effort Choice Parameters

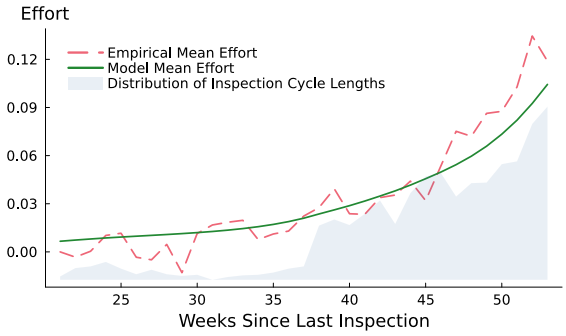
(a) Estimated Model Parameters

Parameter	Estimate (SE)		Description
	Fast States	Slow States	
α_1	−0.041 (0.011)	−0.013 (0.016)	Linear effort cost
α_2	3.107 (0.070)	2.610 (0.082)	Quadratic effort cost
α_3	19.745 (2.685)	30.759 (7.521)	Adjustment cost
σ_e	0.601 (0.002)	0.642 (0.002)	SD of effort noise

(b) Model Fit: Fast States



(c) Model Fit: Slow States



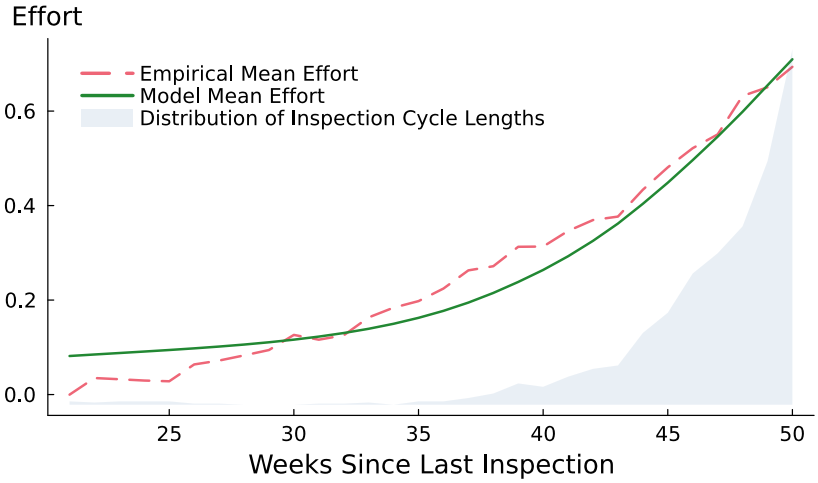
Notes: Panel (a) presents estimates of the effort parameters for facilities in relatively fast and slow states. Panels (b) and (c) present the model fit in both groups, plotting both the mean nursing home efforts over time as well as the model-implied optimal effort. The figures also include the empirical distributions of inspection cycle lengths in each group.

Figure D.2: Effort Choice Parameters: California

(a) Estimated Model Parameters

Parameter	Estimate	SE	Description
α_1	-0.051	0.006	Linear effort cost
α_2	0.399	0.010	Quadratic effort cost
α_3	6.971	0.652	Adjustment cost
σ_e	0.559	0.004	SD of effort noise

(b) Model Fit: California



Notes: Panel (a) presents estimates of the effort parameters for facilities in California. Panel (b) presents the model fit, plotting both the mean nursing home effort over time as well as the model-implied optimal effort. The figure also includes the empirical distribution of inspection cycle length.

E Additional Details on Hazards

Monotonic Hazard To recover the empirical hazard rate from the inspection data, we compute the proportion of facilities inspected in each week conditional on not having been inspected earlier. Let n_w be the number of cycles with length exactly w, and S_w be the number of facilities that are not yet inspected at the beginning of week w. The empirical hazard rate at w is as follows:

$$h(w) = \frac{n_w}{S_w}, \text{ where } S_1 = \sum_{\tau=1}^{W} n_\tau, S_w = S_{w-1} - n_{w-1} \ (w \geq 2).$$

We then impose a monotonicity restriction: whenever $h(w) < h(w-1)$, we set $h(w) = h(w-1)$.

Long-run Hazard We observe in the data that efforts start to flatten out from $w = 51$. For $w > \overline{w} = 51$, we compute a long-run hazard, denoted as $h(\overline{w})$. Specifically, we calculate $h(\overline{w})$ as the ratio of the total number of inspections to the total number of SNF-weeks observed beyond $\overline{w}$. That is,

$$h(\overline{w}) = \frac{\text{Total inspections beyond } \overline{w}}{\text{Total SNF-weeks observed beyond } \overline{w}}.$$

F Additional Details on Counterfactuals

F.1 Details on Counterfactuals

Unpredictable Inspections. One of the primary adjustments we make to inspection timing is to consider fully unpredictable inspections. Doing so is equivalent to imposing a constant inspection hazard 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.[31]

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:

$$E[L] = \sum_{w=1}^{T} \left(\prod_{\tau=1}^{w-1} (1 - h_\tau) \right) \tag{28}$$

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 $\tilde{h}(t) := \bar{h} = 1/E[L]$. (More precisely, to account for the revisit window, the constant hazard rates are calculated as $\bar{h} = 1/(E[L] - E[\tau_k])$.) We then resolve 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.

Adjusting Inspection Frequency. Several counterfactual inspection regimes involve increasing inspection frequency without changing the shape of the hazard rate. Doing so requires an alternative hazard function $\tilde{h}(w)$ such that the implied mean time between inspections is some fixed percentage lower or higher than the one implied by the monotone empirical rate.

To find such an increased frequency hazard function $\tilde{h}(w)$ without changing the original function's 'shape', we employ the following iterative procedure. Starting from the original empirical hazard $h(w)$ (a vector), we remove one period from the start and add one period

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

of $\max h(w)$ to the end. We then compute the new $\widetilde{E[L]}$ implied by this modified hazard vector using equation (28). We iterate this procedure until the modified mean time between inspections $\widetilde{E[L]} \approx \alpha E[L]$, where for example $\alpha = 0.8$ implies a 20% decrease in the expected cycle length, or equivalently, a 25% increase in the mean inspection frequency. Doing so 'shifts' the hazard function to the left or to the right, while maintaining the function's shape beyond the initial excised region. Figure F.1 presents the results summarized in the main text: a 25% increase in inspection frequency, along with a 25% decrease.

No Inspections. The benchmark against which all our inspection regimes are compared is the "no inspections" case. This benchmark allows us to evaluate the value of inspections. To compute the model-implied effort (and therefore lives saved) under this benchmark, we impose $h(w) = 0$ for all periods w and re-solve the Bellman equation, holding our estimates of the parameters fixed, and simulate counterfactual efforts implied by this policy.

Fully Scheduled Inspections. To consider fully scheduled inspections, which are perfectly predictable from the vantage point of facilities, we impose the following hazard rate. Facilities face no inspection risk prior to week 53: $h(w) = 0$ for all periods $w < 53$. In week 53 (the median inspection length), we set $h(53) = 1$, ensuring that facilities are inspected in the same period every cycle, and negating any of anticipatory benefits coming from the unpredictability of inspections.

Calculating Lives Saved. For each inspection regime, we can solve the Bellman equation with the given hazard function $h(w)$. The solution yields the average optimal effort $e(w)$ and we can estimate the lives saved by leveraging the effort-survival relationship. However, the revisit window complicates the analysis, as the hazard function faced by facilities — and therefore their chosen efforts — depend on when the cycle starts.

Denote τ as the start of an inspection cycle (i.e., the week following the revisit, or, in the cases of no revisits, the week after the inspection). $e(w, \tau)$ is defined as the optimal average effort in week w conditional on the stock being *reset to 0* at τ[32]. The share of time spent in week w given the cycle starts at τ is:

$$P(w, \tau) = P(w|\tau)P(\tau)$$

where $P(\tau)$ is the empirical probability distribution of the cycle starts derived from the survey start dates and $P(w|\tau)$ is calculated analytically based on the hazard function $h(w|\tau)$:

[32]During the revisit window, $e(w, \tau)$ is set to be the long-run effort in the no-inspection counterfactual.

$$h(w \mid \tau) = \begin{cases} 0, & \text{if } w < \tau, \\ h(w), & \text{if } w \geq \tau. \end{cases}$$

Note that $h(w|\tau) = 0$ if $w < \tau$. This ensures that no inspection will occur within the revisit window.

Denote the average number of patients in a week from 2017-2019 as N, and the weekly effect of effort on patient survival as β. The total number of lives saved annually is then:

$$\eta = 52 \times N \times \beta \sum_{w,\tau} e(w,\tau) P(w,\tau)$$

For comparison, we also re-estimate the counterfactuals without the revisit window entirely. In this case, $h'(w) = h(w)$ for $w \geq 21$. That means the hazards in all inspection regimes except for the unpredictable regime are effectively shifted to the left by 20 periods, and the constant hazard rates are recalculated by the shifted hazard function $h'(w)$, as shown in Figure F.3 panel (a). And the lives saved annually can be simplified as:

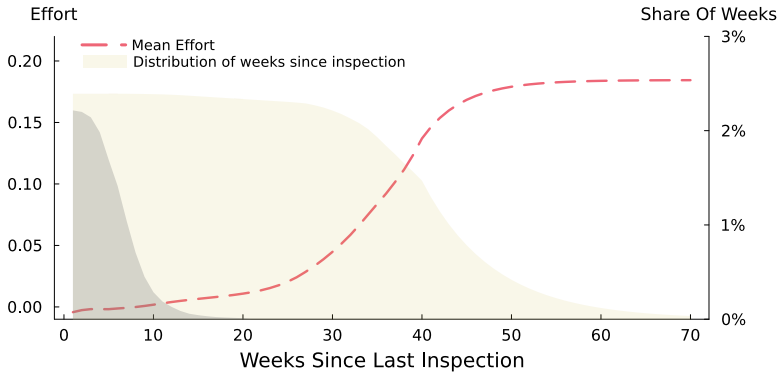
$$\eta' = 52 \times N \times \beta \sum_{w} e'(w) P'(w)$$

We find highly similar results, shown in Figure F.3 panel (b).

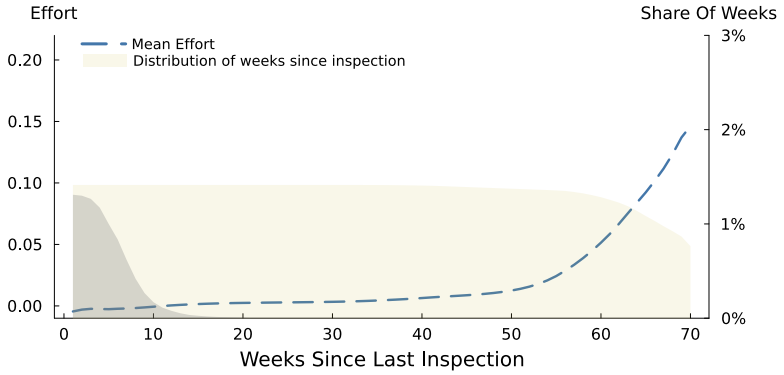
F.2 Additional Counterfactual Results

Figure F.1: Adjusting Inspection Frequency

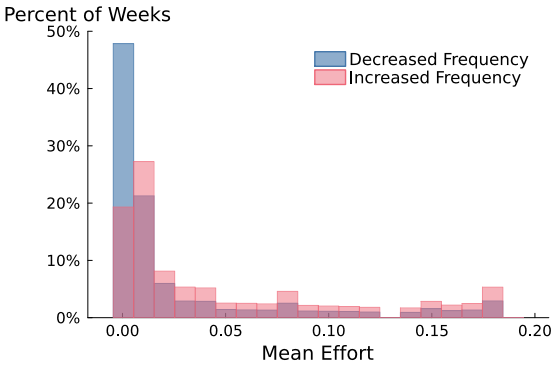
(a) Effort Path and Distribution of w under Increased Frequency



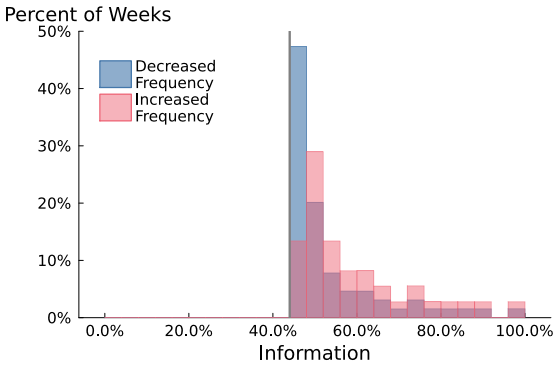
(b) Effort Path and Distribution of w under Decreased Frequency



(c) Distribution of Effort



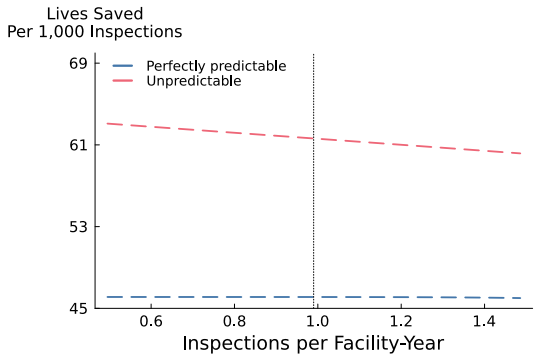
(d) Distribution of Information



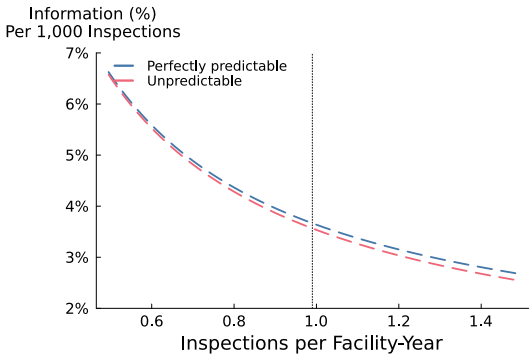
Notes: Figure summarizes the results of two benchmark inspection frequencies, analogous to Figure 10. We consider a 25% increase in inspection frequency as well as a 25% decrease in frequency relative to the current rate. The gray areas in panels (a) and (b) refer to revisit periods. Panel (c) presents the corresponding distributions of model-implied effort. Panel (d) corresponding distributions of information.

Figure F.2: Predictability and the Value of Frequency

(a) Lives Saved Per 1,000 Inspections



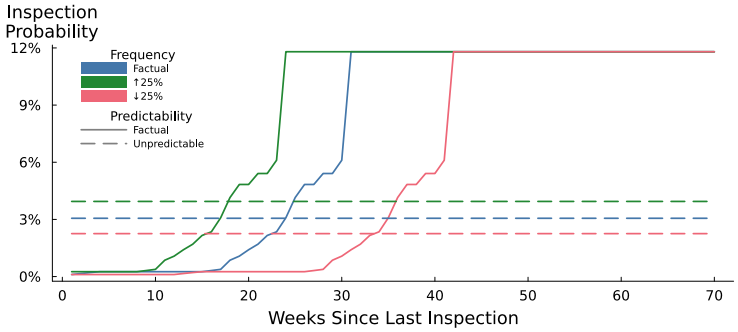
(b) Information



Notes: Figure summarizes the results of flexibly adjusting the inspection frequency from −50% to +50% of the baseline rate, under both the fully unpredictable and fully predictable inspection regimes. Panel (a) presents the number of lives saved at each inspection frequency. Panel (b) presents a similar calculation for the informational gain.

Figure F.3: Alternative Inspection Regimes without Revisit Window

(a) Inspection Hazard Rates



(b) Summary of Results

	Frequency (Per Facility-Year)	Predictability	Lives Saved Annually	Lives Saved Per 1,000 Inspections	Information (%)	Information (%) Per 1,000 Inspections
No Inspections (Benchmark)	0	-	0	-	0	-
Current Regime	1.59	Factual	1,351.9	54.1	60.6	2.4
Increase Frequency (↑ 25%)	2.05	Factual	1,673.4	52.0	63.4	2.0
Decrease Frequency (↓ 25%)	1.17	Factual	1,030.5	56.0	57.6	3.1
Unpredictable	1.59	Unpredictable	1,507.1	60.3	58.4	2.3
Unpredictable; Increased Frequency (↑ 25%)	2.05	Unpredictable	1,906.5	59.3	60.9	1.9
Unpredictable; Decreased Frequency (↓ 25%)	1.17	Unpredictable	1,130.5	61.4	55.8	3.0
Perfectly Predictable	1.58	Perfectly	1,139.6	46.0	60.9	2.5
Perfectly Predictable; Increased Frequency (↑ 25%)	2.08	Perfectly	1,499.3	45.9	64.7	2.0
Perfectly Predictable; Decreased Frequency (↓ 25%)	1.18	Perfectly	855.2	46.0	57.6	3.1

Notes: This figure summarizes the counterfactual results assuming there is *no revisit window*. Panel (a) presents the hazard functions $h(w)$ that characterize each of the inspection regimes considered. We plot the hazard rates corresponding to baseline frequency as well as 25% increases and decreases in frequency. Both factual and unpredictable hazard rates are plotted for each frequency. We omit the benchmark hazard rate, at which $h(w) = 0$ for all w. Panel (b) summarizes the results of each inspection regime considered in panel (a), in addition to the fully scheduled inspection regime, in which $h(w) = 0$ for $w \neq 33$, and $h(33) = 1$.

G Appendix Tables

Table G.1: Top 20 Deficiencies (2006-2019)

Deficiency	Description	Share (%)
812	Food Procurement, Store/Prepare/Serve - Sanitary	4.79
880	Infection Prevention & Control	4.42
813	Personal Food Policy	4.19
689	Free of Accident Hazards/Supervision/Devices	4.13
684	Quality of Care	3.39
698	Dialysis	3.16
697	Pain Management	3.15
744	Treatment /Service for Dementia	3.09
675	Quality of Life	3.06
656	Develop/Implement Comprehensive Care Plan	2.97
761	Label/Store Drugs & Biologicals	2.51
658	Services Provided Meet Professional Standards	2.5
757	Drug Regimen is Free From Unnecessary Drugs	2.47
686	Treatment/Svcs to Prevent/Heal Pressure Ulcers	1.99
690	Bowel/Bladder Incontinence, Catheter, UTI	1.92
657	Care Plan Timing and Revision	1.75
659	Qualified Persons	1.64
677	ADL Care Provided for Dependent Residents	1.55
641	Accuracy of Assessments	1.53
921	Safe/Functional/Sanitary/ Comfortable Environment	1.37

Notes:

This table lists the top 20 most common deficiencies among inspections in 2006-2019, ranked by the share of inspections they are cited in. To account for a 2017 reform which redefined deficiency definitions, we crosswalk the pre-reform deficiency tags into post-reform ones following CMS (2024a).

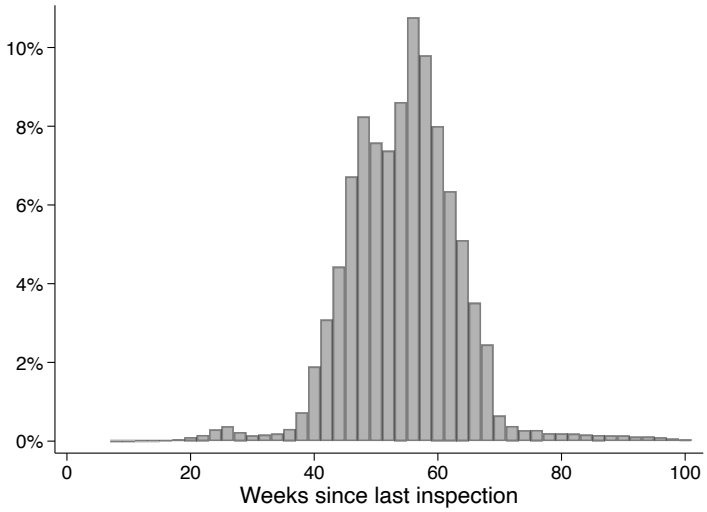
Table G.2: Summary Statistics

	Mean	Std. Dev	50%	25%	75%
Facility Effort Sample: 2017-2019					
Staffing Characteristics					
Total hours	2,232	1,418	1,952	1,295	2,805
Hours per resident-day	3.77	0.96	3.63	3.19	4.14
RN hours per resident-day	0.67	0.52	0.57	0.38	0.81
Admin hours per resident-day	0.29	0.20	0.26	0.18	0.37
Share with contract hours	33%	-	-	-	-
Full-time hours per resident-day	2.42	0.72	2.35	1.99	2.76
Overtime hours per resident-day	1.43	0.60	1.38	1.03	1.77
Weekend hours per resident-day	0.96	0.26	0.92	0.81	1.06
Facility Occupancy Measures					
Weekly census	70.9	52.8	65	38	97
Weekly admissions	2.7	3.4	2	0	4
Weekly discharges	3.1	7.1	2	0	4
Facility Ownership Characteristics					
Share for-profit	73.8%	-	-	-	-
Share chain owned	66.5%	-	-	-	-
Share private equity owned	0.5%	-	-	-	-
N Inspection cycles = 40,654	N Facilities = 15,715		N Facility-weeks = 1,904,128		
Patient Care and Health Outcomes Sample: 2013-2019					
Patient Characteristics					
Age	80.8	11.7	82.7	73.6	89.4
Predicted survival rate at admission	0.88	0.08	0.89	0.83	0.93
Share black	12.1%	-	-	-	-
Share female	65.4%	-	-	-	-
Patient Care Measures					
Share not on anti-psychotics	82.8%	-	-	-	-
Share without physical restraints	98.7%	-	-	-	-
Patient Health Outcomes					
ADL score (negative)	−7.6	4.8	−8	−12	−4
4-week survival rate	95.9%	-	-	-	-
Share without pressure ulcers	90.9%	-	-	-	-
Share without fall or fracture	96.8%	-	-	-	-
Share discharged to a hospital	1.68%	-	-	-	-
Patient Characteristics at Admission					
Share with Alzeimer's or Dementia	23.9%	-	-	-	-
Share on Medicaid	10.1%	-	-	-	-
Share with hospital stay in last 30 days	82.1%	-	-	-	-
Number of chronic conditions	6.2	2.8	6	4	8
N Inspection cycles = 118,209	N Facilities = 16,580				
N Stays = 16,949,743	N Patient-weeks = 351,304,403				
Inspection Outcomes Sample: 2006-2019					
Inspection Characteristics					
Weeks between inspections	55.0	5.5	55	50	59
Average deficiencies	7.0	6.4	5	2	10
Share with deficiency	87.1%	-	-	-	-
N Inspection cycles = 156,341	N Facilities = 17,399				

Notes: Patient care and health outcomes sample is limited to the 91.9% of all nursing home residents who are in the MBSF and have 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. The Inspection Outcomes sample is restricted to cycles whose lengths fall between 45 and 65 weeks.

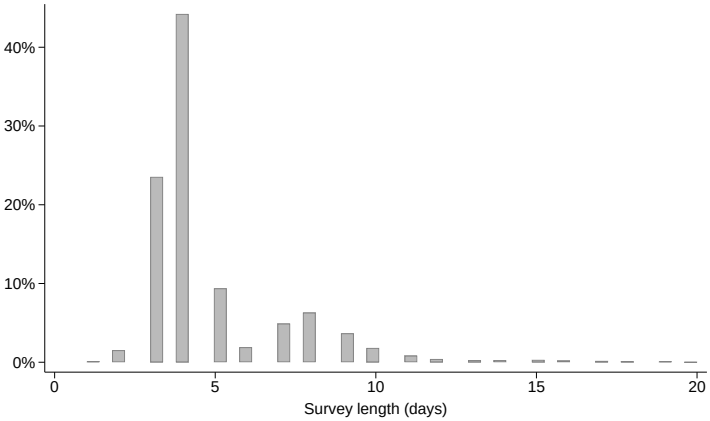
H Appendix Figures

Figure H.1: Histogram of Weeks Since Last Inspection



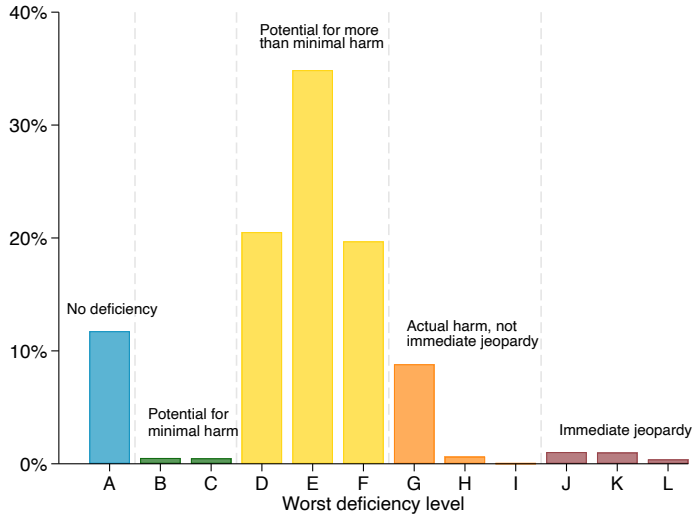
Notes: This figure plots a histogram of weeks since last inspection for standard health surveys in 2013-2019. The mean is 54.4 and the standard deviation is 9.9.

Figure H.2: Distribution of Inspection Length (2014-2019)



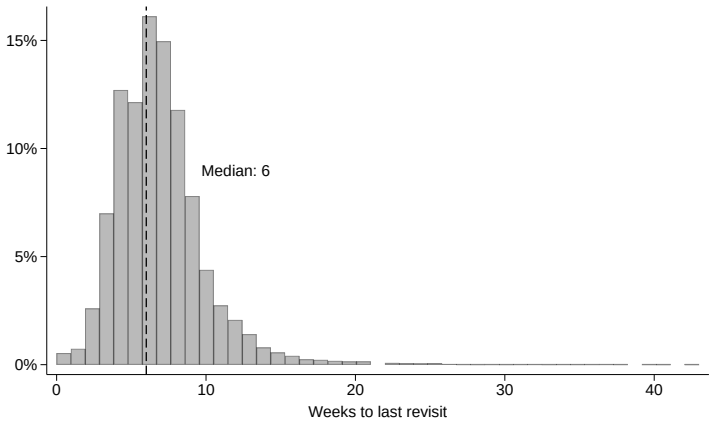
This figure plots the distribution of inspection survey length for inspections from 2014-2019. Inspection length is calculated as the difference between the last day of a survey and the first day and derived from data made available through a Freedom of Information Act request. The average inspection length is 4.9, the 25th percentile is 3, the median is 4, and 75th percentile is 5.

Figure H.3: Distribution of Highest Deficiency Category, 2006-2019



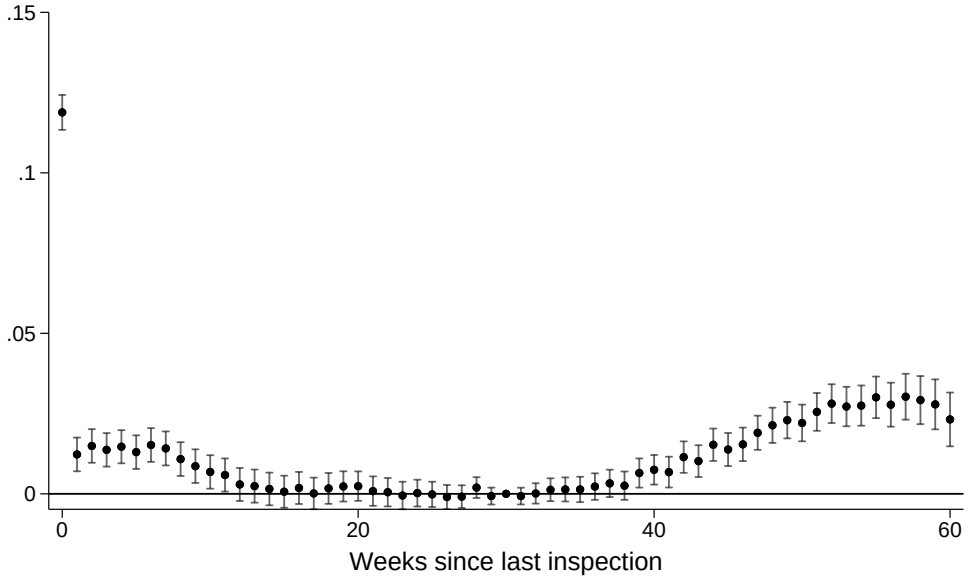
Notes: This figure plots a histogram of the worst deficiency level of all deficiencies found in an inspection in 2006–2019. Level A corresponds to no deficiency or an isolated deficiency with no actual harm with potential for minimal harm was found, while levels B-C correspond to a pattern of or widespread deficiencies, respectively, that pose no actual harm with potential for minimal harm. Levels D-F correspond to isolated, pattern of, and widespread deficiencies 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 with actual harm that is not immediate jeopardy. Levels J-K correspond to isolated, pattern of, and widespread deficiencies with immediate jeopardy to resident health or safety (CMS, 2023).

Figure H.4: Distribution of Weeks to Last Revisit



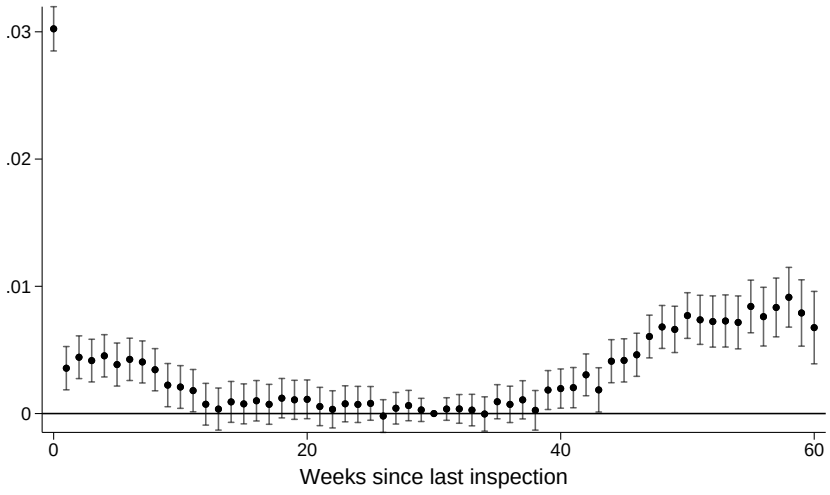
Notes: This figure plots the histogram of weeks between a facility's initial inspection and their last revisit inspection, among the 95.1% of inspections in 2017-2019 with at least one revisit.

Figure H.5: Effort (Across-Facility Std. Dev.) By Weeks Since Last Inspection



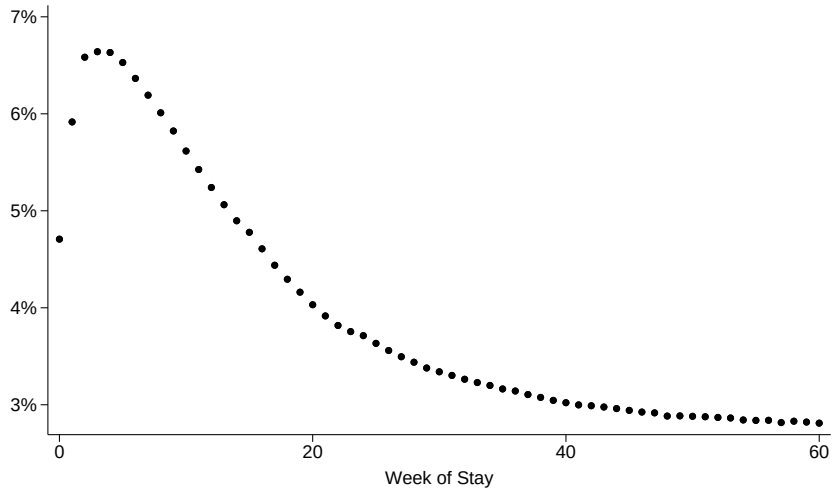
Notes: This figure plots the coefficients from estimating equation (2) on the composite effort index, where the effort measure is defined in terms of across-facility standard deviations, which are constructed using distribution of effort across all facilities in the middle of their inspection cycles (weeks 21-30). In contrast, the main results in Figure 2 plot the coefficients in terms of within-facility standard deviations, which are constructed relative to each individual facility's distribution of effort in the middle of their inspection cycle. The effort index is a composite measure combining labor and non-labor inputs, and is described in further detail in Section 3.1 and Appendix Section A.2.

Figure H.6: Log Total Hours by Weeks Since Last Inspection



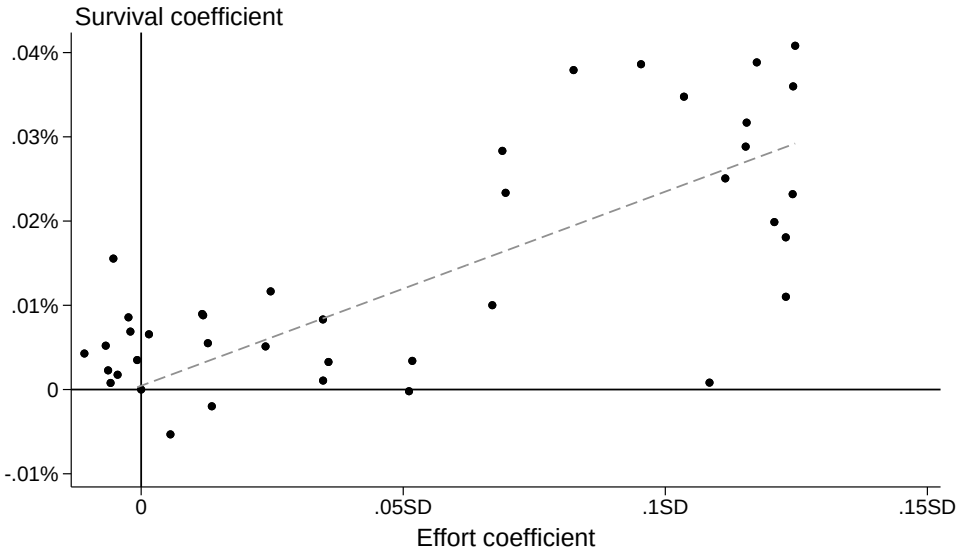
Notes: This figure plots the estimates from equation (2) on log total hours per week.

Figure H.7: Patient Mortality, by Weeks of Stay



Notes: This figure plots the 4-week patient mortality rate by week-of-stay.

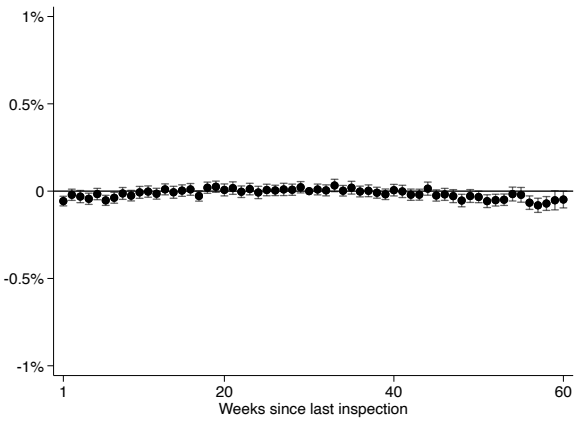
Figure H.8: Comparison of Inspection Hazard, Effort, and Survival Coefficients, Restricted to $\geq$ 21 Weeks from Inspection



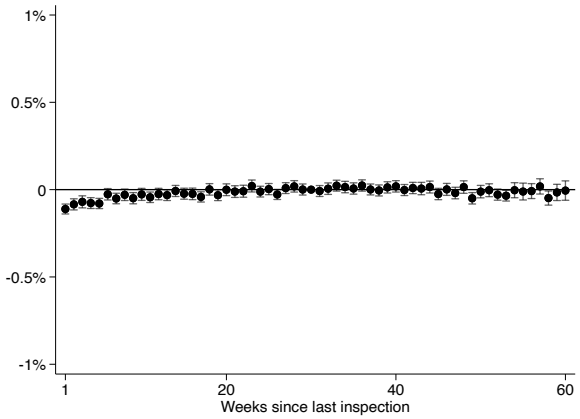
Notes: Figure H.8 plots the relationship between the effort and survival coefficients in Figure 2. Coefficients are derived from estimating equations (2) and (3) on the full sample, and the coefficients plotted are restricted to weeks from inspection $w \geq 21$. The slope of the line of best fit for these weeks is 0.002 (SE 0.0003). The effort index is a composite measure combining labor and non-labor inputs, and is described in further detail in Section 3.1 and Appendix Section A.2.

Figure H.9: Patient Admissions and Discharges by Weeks Since Last Inspection

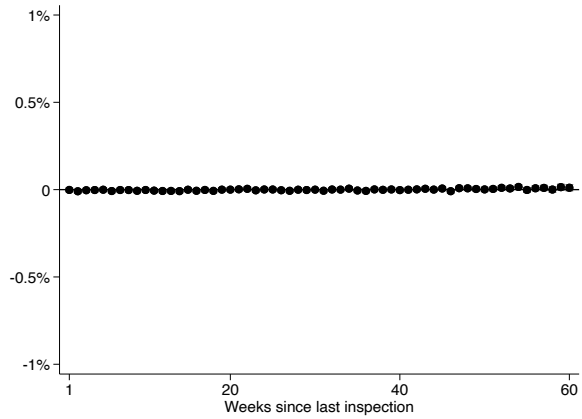
(a) Admissions



(b) Discharges with no return anticipated



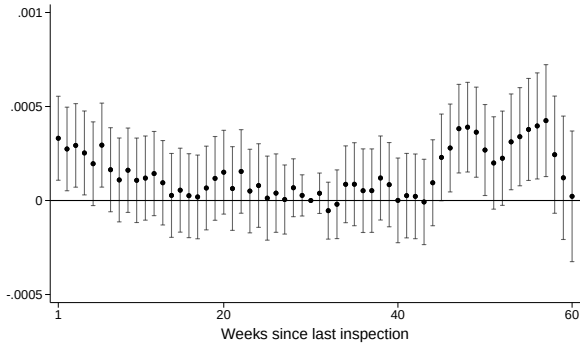
(c) Discharges with return anticipated



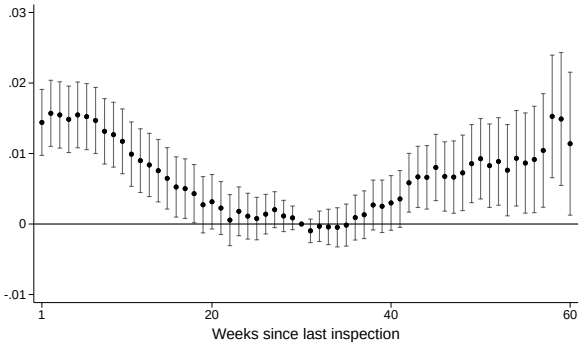
Notes: This figure plots the coefficients from estimating equation (2) on (a) admissions, (b) discharges with no return anticipated, and (c) discharges with return anticipated. All values are normalized in percent terms relative to the average patient count in the omitted week. The coefficients for week 0 are estimated but not reported.

Figure H.10: Survival Rate and Additional Health Outcomes by Weeks Since Last Inspection, without Predicted Survival Controls

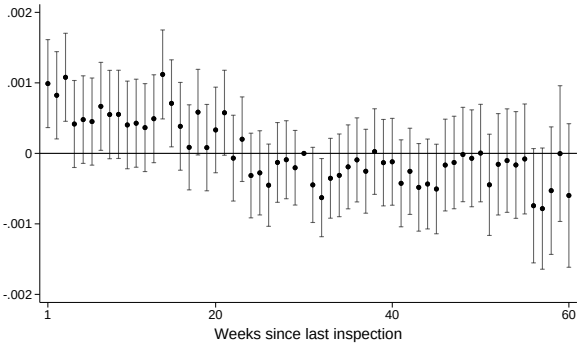
(a) 4-Week Survival Rate



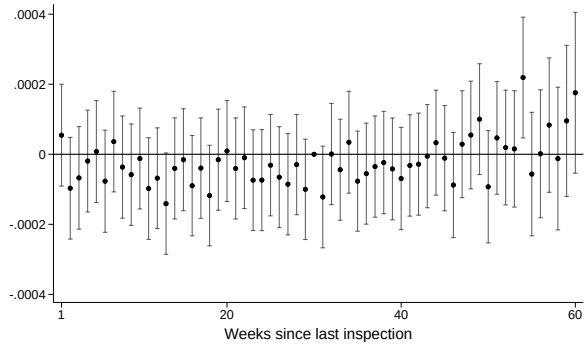
(b) (Negative) Activities of Daily Living Score



(c) Patients Without Pressure Ulcers



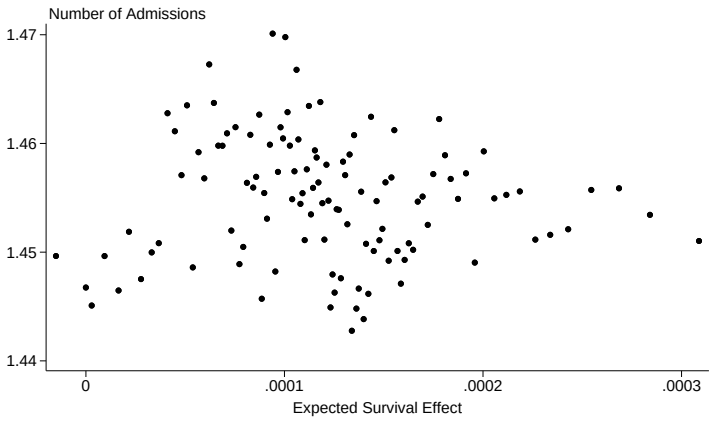
(d) Patient Discharged to Hospital



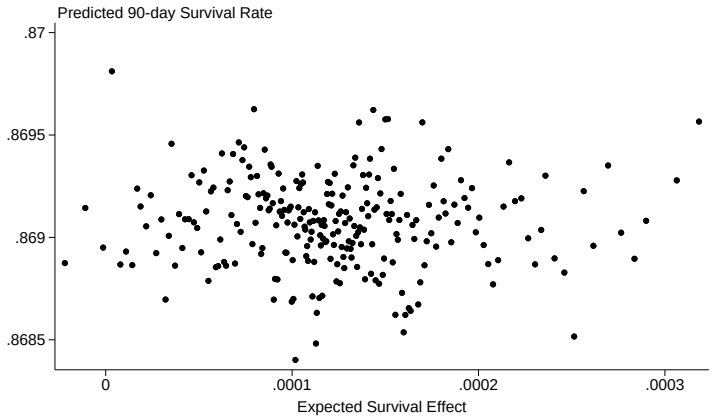
Notes: This figure plots the estimates from equation (3), excluding the control for predicted patient survival. The coefficients for week 0 are estimated but not reported.

Figure H.11: Zip-Code Level Analysis: Relationship between Patient Pool and Expected Survival Effect

(a) Number of Admissions

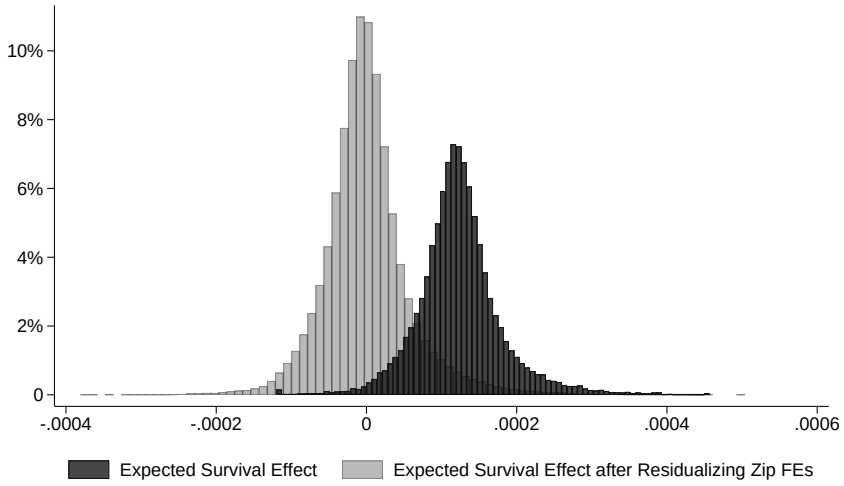


(b) Predicted Survival of Nursing Home patients



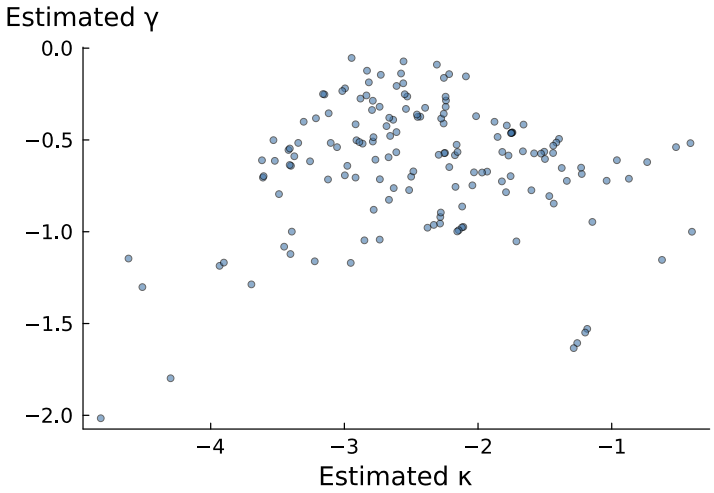
Notes: This figure plots the relationships between the expected survival effect, $\widehat{z}_{it}$ experienced by patient i in calendar week t and (a) the number of admissions from a zip code in that week and the (b) the 90-day predicted survival rate for patients in a nursing home from a zip code in that week, after residualizing out zip code and calendar week fixed effects. The expected survival effect is defined in equation (6). See Section B.2 for a description of how the predicted 90-day survival rate is constructed.

Figure H.12: Distribution of Expected Survival Effect



Notes: This figure plots histograms of the expected survival effect z_{it} before and after residualizing out patient zip code fixed effects. The results indicate that there is considerable variation within zip code and across-time.

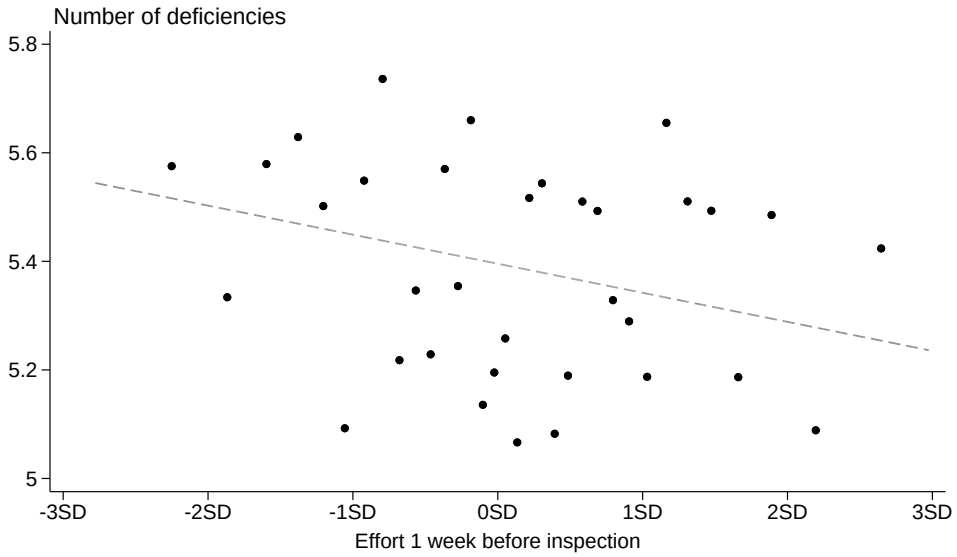
Figure H.13: Estimates of γ_d and κ_d



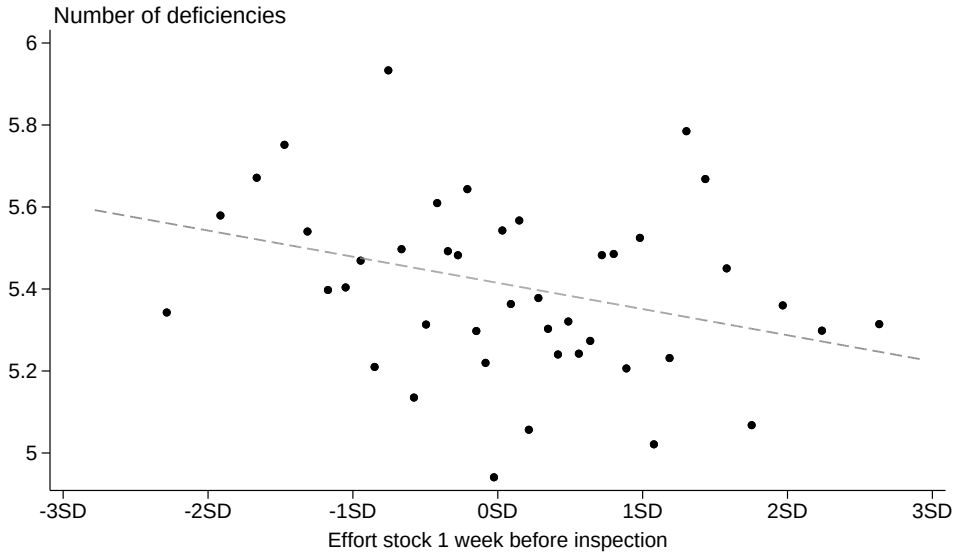
Notes: This figure plots for each deficiency d, the corresponding estimates of γ_d (the sensitivity of deficiency d to quality and effort) and κ_d (deficiency d's baseline prevalence).

Figure H.14: Relationship Between Effort and Stock and Inspection Outcomes

(a) Deficiency Count vs. Effort in Week Before Inspection

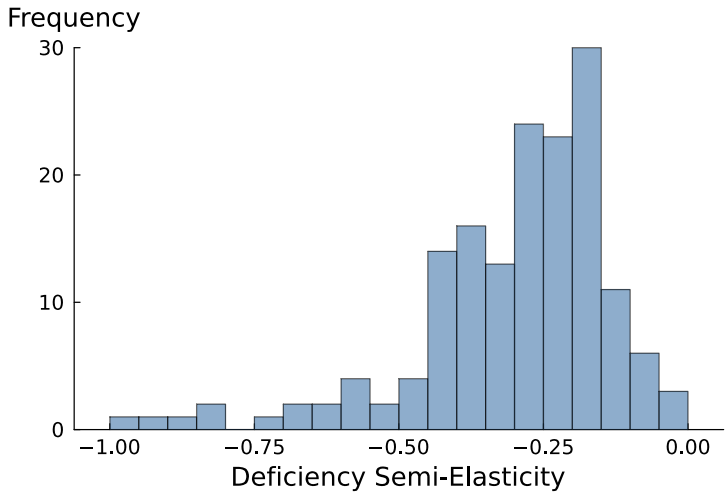


(b) Deficiency Count vs. Effort Stock in Week Before Inspection



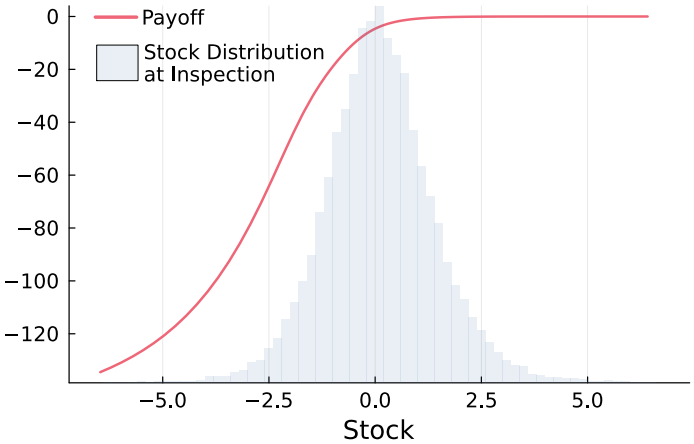
Notes: This figure plots binscatters of the relationships between the number of deficiencies in each inspection and (a) the composite effort index in the week prior to inspection, and (b) the stock of effort in the week prior to inspection. State-year, facility, and inspection week fixed effects have been residualized out. The average number of deficiencies is 5.5. The coefficient on effort in the regression for panel (a), clustered at the facility-level, is −0.034 (SE: 0.026), and the coefficient on the stock of effort in the regression for panel (b) is −0.055 (SE: 0.027). See Section 6.1 for a description of how effort stock is constructed.

Figure H.15: Deficiency Semi-Elasticities



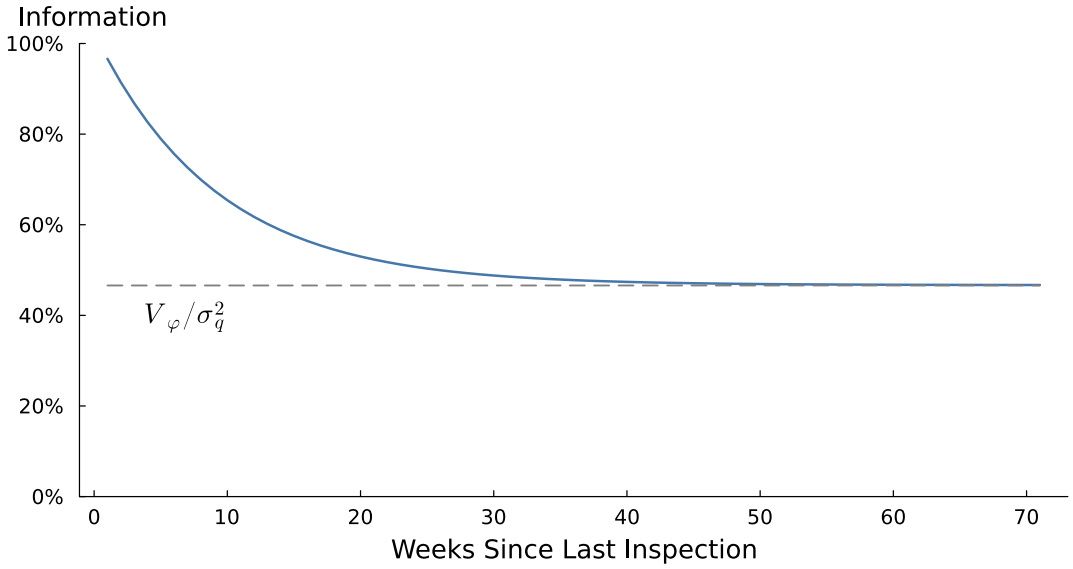
Notes: This figure plots the distribution of instantaneous semi-elasticities of each deficiency d, at mean quality, when moving from initial effort stock 0 to the average stock at the end of the cycle.

Figure H.16: Payoff Function $\pi(s_{j,k,w})$



Notes: This figure presents the payoff function $\pi(s_{j,k,w})$ a facility expects to receive if inspected w weeks after the prior inspection. For clarity, we also overlay the distribution of realized $s_{j,k,w}$ occurring at inspection time.

Figure H.17: Information Over Quality



Notes: This figure characterizes the regulator's information over quality, given by $1-U_w$. Derivation of this uncertainty measure provided in Appendix C.5.