# My Title

Joe White[a],[*]

[a]*Department of Economics, Miami University, Oxford, OH, USA*

---

**Abstract**

---

## 1. Introduction

This paper examines how changes in ownership affect staffing levels in U.S. nursing homes. Ownership transitions, through acquisitions, sales, and mergers, have become increasingly common as consolidation in the long-term care sector grows. Much of the current literature studies how consolidation affects prices, costs, and market power, substantially less is known about how ownership changes influence patient outcomes such as nurse staffing. Existing evidence on quality effects is mixed, and staffing, the primary determinant of care quality in nursing homes, has had limited attention.

Healthcare workers make up roughly 10 percent of the workforce and about 17 million people in the United States. The largest healthcare profession in the United States is nursing, and studying their productivity is imperative to understanding and improving patient care. Nursing homes provide a particularly important setting for studying ownership changes. They serve a medically vulnerable population, operate under heavy regulation, and rely intensively on skilled and semi-skilled nursing labor. Changes in ownership may alter managerial priorities, financial constraints, and labor practices in ways that directly affect staffing decisions. Understanding these effects is especially relevant for policymakers and regulators concerned with access to care and patient outcomes.

Using detailed administrative data from CMS, we study the causal effect of ownership changes on nursing hours per patient day. Using staggered ownership transitions and a difference-in-differences event study design, we document significant declines in staffing following ownership changes. These effects vary by staff type, pandemic period, and baseline chain affiliation, highlighting the role of organizational structure and labor market conditions in shaping post change outcomes.

---

[*]Corresponding author

## 2. Institutional Background

## 3. Data and Methods

### 3.1. Data Sources

Three main data sources are utilized, all of which are maintained by the Center for Medicare and Medicaid Services (CMS). The first is the Nursing Home Compare (NHC) Archive, which contains monthly facility level information on ownership, chain affiliation, certification status, and quality measures. The second is Healthcare Cost Report Information System (HCRIS) which contains annual cost reports submitted by every Medicare certified skilled nursing facility (SNF) in the United States. The cost reports contain information about costs, revenue, financial data for each reporting facility, and dates of ownership changes. Lastly, we use the Payroll Based Journal (PBJ) system, which provides facility level, daily staffing information by staff type.

### 3.2. Sample selection

Our sample is constructed at the facility-year-month level. We create our sample using the following criteria. First, we limit our sample to the 48 contiguous U.S. states from the period Jan 2017 - June 2024. Secondly, the Care Compare Archive data is matched to the HCRIS data on a unique nursing home provider number. Because HCRIS only contains information on SNFs, our sample focuses on SNFs instead of all nursing facilities in general.

### 3.3. Identifying Ownership Changes

To our knowledge there are no datasets that have information about ownership changes for all homes within the United States. Therefore, we manually assemble this dataset. We identify ownership changes with two-step procedure using data from HCRIS and CMS. We use the HCRIS Medicare Cost Reports ownership filing data, which contains facility reported ownership change indicators and dates for all certified skilled nursing facilities. We use this as our flag for ownership changes during the period 2017-2024 because facilities are legally required to file these changes, so we have a standardized account of ownership transitions across facilities. However, these indicators may capture administrative or legal changes that do not reflect meaningful transfers of control. To verify whether flagged events correspond to substantive ownership changes, we validate each change of ownership (CHOW) using detailed ownership information from the CMS NHC Archive ownership files. CMS ownership files provide monthly ownership shares for each owner associated with a facility. As a result of the advanced corporate ownership structure detailed previously, it is difficult to discern ownership changes while looking at every ownership level. Therefore, when validating ownership changes, we look at the highest level of owner type in time period $t$. The order, highest to lowest, is defined as indirect owners, direct owners, and then individuals or organizations with a partnership interest. There are three main criteria we use when defining a true change in ownership. First, we define a CHOW when a majority of ownership shares turn over between month

$t-1$ and month $t$. Intuitively, this captures transactions in which control of the facility changes hands, including cases where a new operator acquires most or all shares. Secondly, when ownership percentages are partially or completely missing, we approximate shares by normalizing any reported percentages and, where necessary, assigning equal weights across remaining owners. To reduce false positives from purely legal restructuring, we do not flag a CHOW if at least 80 percent of the ownership stake remains within the same surname family (e.g., a change from Smith Holdings LLC to Smith Family Partners LP). Using true ownership changes, we retain facilities that either (i) have no ownership change in either source or (ii) have exactly one ownership change in each source with the CHOW date within six months of one another. For these facilities, we define the CHOW month as the first month in which the NHC data indicate a change. All dates are converted to calendar months, so changes occurring at the beginning or end of a month are treated identically.

### 3.4. Nursing Staff Levels

Our main outcomes are nursing hours per patient day (HPPD), constructed from the CMS PBJ data. This data includes daily hours worked by each staff type in each facility, together with the daily census from the Minimum Data Set (MDS). The staff types we use in our study are registered nurses (RNs), licensed practical nurses (LPNs), and certified nursing assistants (CNAs). RNs are the most highly trained nursing staff and are responsible for assessment, care planning, medication administration, and supervision of other nursing personnel. LPNs provide basic medical care under the supervision of RNs and physicians, including monitoring vital signs and administering certain treatments. CNAs provide the majority of direct daily care to residents, assisting with activities of daily living. Because these staff types differ in training, scope of practice, and compensation, changes in staffing composition may reflect different organizational responses to ownership changes. We therefore examine each staff category separately, as well as total nursing hours per patient day. For each facility $i$, calendar month $t$, and staff type $s$, we first aggregate daily hours to the monthly level and then compute

$$\mathrm{HPPD}_{i,t}^{(s)} = \frac{H_{i,t}^{(s)}}{R_{i,t}}, \tag{1}$$

where $H_{i,t}^{(s)}$ is the total hours worked by staff type $s$ during month $t$ and $R_{i,t}$ is the total resident days in that facility-month. Total HPPD is defined as the sum of RN, LPN, and CNA HPPD.

We construct two additional measures of data quality. First, we define a monthly coverage ratio

$$C_{i,t} = \frac{\mathrm{DaysReported}_{i,t}}{\mathrm{DaysInMonth}_t},$$

where $\mathrm{DaysReported}_{i,t}$ is the number of distinct days in month $t$ with complete PBJ records for facility $i$. This measure allows us to identify incomplete data

3

in our sample and describe our data completeness. Second, we track gaps in the reporting for each facility. For each observation we compute *gap*, defined as the number of calendar months since the previous PBJ observation for that facility. Finally, we exclude a small number of observations with implausible staffing levels: months in which RN and LPN HPPD are both zero, total HPPD is below 1.5 or above 12, or CNA HPPD exceeds 5.25. These observations are outliers that likely reflect reporting errors rather than genuine variation in staffing. These exclusion criteria align with the criteria outlined in the CMS NHC Technical Users Guide.

### 3.5. Other Controls

We obtain facility characteristics from NHC and HCRIS. The following variables are used as controls in all specifications: Ownership type (government, non-profit, and for-profit), chain affiliation, total number of beds, occupancy rate, payer mix (percent of revenue coming from Medicare, and percent of revenue coming from Medicaid), and acuity. Our measure of acuity is case-mix provided by CMS in the NHC archive data. During our study period CMS altered the methodology used to construct case-mix indices. As a result, raw case-mix values are not directly comparable over time. To address this issue, we measure acuity using relative rankings rather than levels. Specifically, we construct case-mix quartiles within each state and calendar month.

### 3.6. Summary Statistics

Table 1 reports summary statistics for the main analysis sample. The final sample consists of approximately 7,806 unique skilled nursing facilities and 632,231 facility-month observations spanning January 2017 through June 2024. Of these facilities,1589 experience an ownership change during the sample period.

**Table 1:** Panel Summary Statistics

**Panel A**

| Variable | N | Mean | SD | P25 | Median | P75 | Min | Max |
|---|---|---|---|---|---|---|---|---|
| Gap from previous months | 632,231 | 0.026 | 0.396 | 0.000 | 0.000 | 0.000 | 0.000 | 6.000 |
| Coverage ratio | 632,231 | 1.000 | 0.017 | 1.000 | 1.000 | 1.000 | 0.323 | 1.000 |
| RN HPPD | 632,231 | 0.427 | 0.309 | 0.220 | 0.366 | 0.562 | 0.000 | 9.105 |
| LPN HPPD | 632,231 | 0.796 | 0.325 | 0.597 | 0.788 | 0.973 | 0.000 | 6.248 |
| CNA HPPD | 632,231 | 2.100 | 0.548 | 1.741 | 2.049 | 2.413 | 0.000 | 5.248 |
| Total HPPD | 632,231 | 3.324 | 0.793 | 2.809 | 3.241 | 3.721 | 1.500 | 11.675 |
| Beds | 632,231 | 108.431 | 57.876 | 70.000 | 100.000 | 128.000 | 15.000 | 874.000 |
| Occupancy rate (%) | 632,231 | 76.4 | 15.5 | 67.1 | 79.7 | 88.8 | 0.12 | 100.0 |
| % Medicare | 632,231 | 13.0 | 11.1 | 6.1 | 10.0 | 16.0 | 0.00 | 99.4 |
| % Medicaid | 632,231 | 60.5 | 20.2 | 51.4 | 64.7 | 74.4 | 0.01 | 99.7 |

**Panel B**

| Category | Count (CCN) | Share (%) |
|---|---|---|
| For-profit | 5,449 | 69.8 |
| Non-profit | 1,436 | 18.4 |
| Government | 921 | 11.8 |
| Chain facilities | 4,874 | 62.4 |

*Notes:* Rows = 632,231; Facilities = 7,806; Period = 2017/012024/06; Average months per facility = 81.0.

### 3.7. Analytic Approach

In our empirical setting, nursing homes change ownership at different points in time. Recent literature shows that conventional two-way fixed effects (TWFE) estimators can produce biased estimates in staggered adoption settings when treatment effects are heterogeneous over time (Goodman-Bacon, 2021). In particular, TWFE estimates can place negative weights on certain comparisons and implicitly rely on already treated units as controls. To circumvent this, we employ a staggered adoption event study regression design. This approach allows us to estimate treatment effects before and after ownership changes while assessing pre-trends. Specifically we estimate the following event study regression:

$$Y_{it} = \sum_{\tau=-24}^{24} \beta_\tau \, \mathbf{1}\{\text{event\_time}_{it} = \tau\} \cdot \text{Treat}_i + X_{it}'\gamma + \alpha_i + \lambda_t + \varepsilon_{it}. \quad (2)$$

where $Y_{it}$ denotes hours per patient day for facility $i$ in in month $t$. The variable event_time$_{it}$ measures the number of months relative to the month of ownership change. Thus, the coefficients $\beta_\tau$ capture the evolution of staffing outcomes before and after ownership transitions. The vector $X_{it}$ includes time-varying facility characteristics, while $\alpha_i$ and $\lambda_t$ denote facility and month fixed effects, respectively. Identification comes from within-facility changes in staffing

relative to the timing of ownership transitions, using facilities that have not yet experienced a change as controls.

*Parallel Trends.* The identifying assumption underlying our event-study design is that, absent an ownership change, staffing outcomes in treated facilities would have followed the same trends as those in untreated facilities. We assess the plausibility of this assumption by examining estimated coefficients on the pre-treatment event-time indicators and by conducting joint tests of pre-trends. Specifically, we test whether the coefficients on all pre-treatment leads of the ownership change are jointly equal to zero using Wald tests. We report joint tests over the full pre-treatment window as well as over a narrower window closer to the ownership transition, providing a more local assessment of parallel trends. In addition, we report tests for the coefficient on the event-time indicator immediately preceding the reference period to evaluate whether there is evidence of differential trends just prior to treatment.

We estimate equation (2) using only observations outside the anticipatory period immediately preceding ownership changes, thus the event study is centered on period $\tau = -4$. This restriction is motivated by evidence of staffing adjustments prior to formal ownership transitions, which would otherwise violate the parallel trends assumption. Standard errors are clustered at the facility level to account for serial correlation within nursing homes.

While our main specification relies on the event-study design, we also present TWFE estimates to summarize the average post ownership change effect. We estimate the following regression:

$$Y_{it}^k = \beta_{\text{post}}^k \, Post_{it} + X_{it}' \boldsymbol{\gamma}^k + \alpha_i + \lambda_t + \varepsilon_{it}^k, \tag{3}$$

where $Y_{it}^k$ denotes staffing outcome $k$ (RN, LPN, CNA, or total hours per patient day) for facility $i$ in month $t$. The indicator $Post_{it}$ equals one for all months following an ownership change and zero otherwise. Facility fixed effects $\alpha_i$ absorb time invariant differences across nursing homes, while month fixed effects $\lambda_t$ control for common shocks to staffing over time.

The coefficient $\beta_{\text{post}}^k$ captures the average change in staffing following an ownership transition, relative to the pre-change period. All specifications include the same time varying covariates as in the event-study models, and standard errors are clustered at the facility level.

*3.7.1. Subgroup Analysis*

We further examine heterogeneity in the effects of ownership change with two further analyses: the COVID-19 pandemic and baseline chain affiliation.

First, we estimate separate models for the pre-pandemic period (January 2017 through December 2019) and the pandemic period (April 2020 through June 2024). Labor market conditions for nursing staff changed substantially during the COVID-19 pandemic, including heightened turnover, changes in wages, and staffing shortages. Estimating models separately by period allows us to

assess whether ownership changes had differential effects on staffing before and during this structural shift in the nursing labor market.

Second, we examine heterogeneity by baseline chain status. Specifically, we split the sample based on whether a facility belonged to a chain at the beginning of the study period (January 2017). Chain affiliated facilities may differ systematically from independent facilities in access to staffers, internal labor markets, and management practices. Estimating models separately for chain and non-chain facilities allows us to test whether ownership transitions have differential staffing effects across these organizational structures.

In both subgroup analyses, we estimate the same specifications as in equations (2) and (3), maintaining consistent controls, fixed effects, and clustering. These analyses shed light on mechanisms and contextual factors that shape the staffing consequences of ownership changes.

## 4. Results

### 4.1. Main Results

Figure 1 presents event study estimates of the effect of ownership changes on nursing staffing outcomes. Across all staffing categories, we observe no evidence of differential trends prior to ownership changes once anticipatory months are excluded. Following an ownership transition, staffing declines persist for many periods, with the largest and most precise effects occurring for registered nurses (RN), certified nursing assistants (CNA), and total nursing hours per patient day.

RN staffing declines immediately following ownership changes and remains far below pre-transition levels throughout the post-treatment period. CNA staffing exhibits a similar but slightly smaller decline, while total nursing hours mirror the combined reductions across staff types. In contrast, LPN staffing shows limited and statistically insignificant changes over time. These dynamic patterns indicate that ownership transitions lead to sustained reductions in staffing rather than short run adjustments.
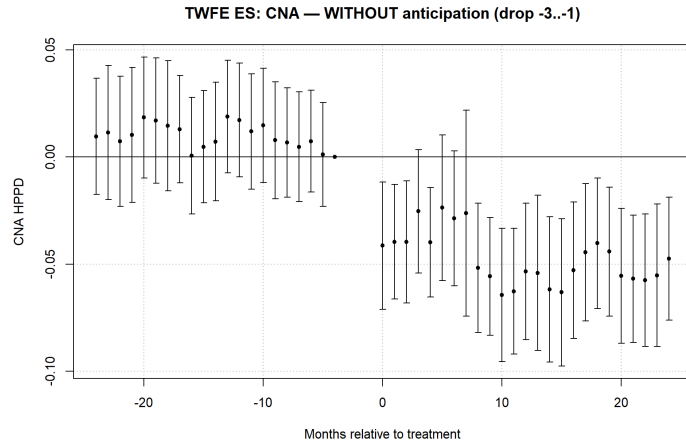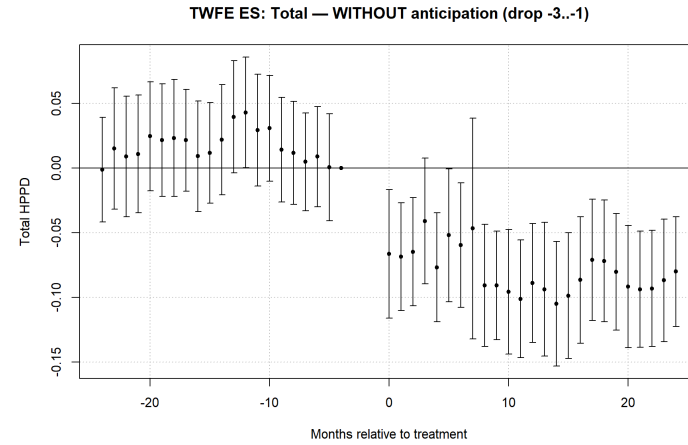
7
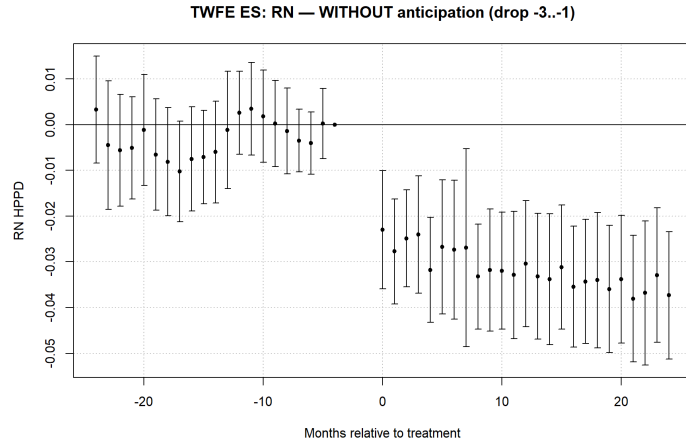
**(a)** Registered Nurses (RN)
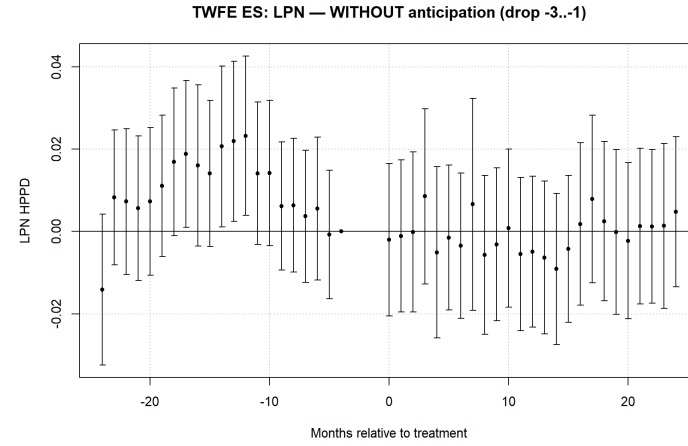
**(b)** Licensed Practical Nurses (LPN)

**(c)** Certified Nursing Assistants (CNA)

**(d)** Total Nursing Hours

**Figure 1:** Dynamic Effects of Ownership Change on Nursing Staffing. Event-study estimates of the effect of a change in ownership on hours per patient day. Points show coefficient estimates relative to the month immediately preceding the ownership change; vertical bars denote 95% confidence intervals. Anticipatory months ($t = -3, -2, -1$) are excluded. All specifications include facility and month fixed effects and baseline covariates.

To summarize these dynamic effects with a single average treatment estimate, Table 2 reports two-way fixed effects estimates of the ownership change effect. Consistent with the event study results, ownership changes lead to meaningful reductions in staffing levels. RN staffing declines by approximately 0.03 hours per patient day (roughly 9 percent), CNA staffing falls by about 0.05 hours (approximately 3 percent), and total staffing decreases by 0.08 hours per patient day (roughly 3 percent). LPN staffing shows no statistically significant change. Estimates are nearly identical when excluding anticipatory months, reinforcing the validity of our event study design.

**Table 2:** Two-Way Fixed Effects Estimates of *post* on Staffing Outcomes (Baseline)

| | Outcomes | | | |
|---|---|---|---|---|
| | **RN** | **LPN** | **CNA** | **Total** |
| **Panel A: Staffing Levels in HPPD** | | | | |
| With anticipation | $-0.031^{***}$ | $0.001$ | $-0.052^{***}$ | $-0.082^{***}$ |
| | $(0.005)$ | $(0.006)$ | $(0.009)$ | $(0.012)$ |
| Without anticipation | $-0.032^{***}$ | $0.000$ | $-0.054^{***}$ | $-0.086^{***}$ |
| | $(0.005)$ | $(0.006)$ | $(0.009)$ | $(0.013)$ |
| **Panel B: Log Staffing Levels in HPPD** | | | | |
| With anticipation | $-0.086^{***}$ | $0.004$ | $-0.026^{***}$ | $-0.025^{***}$ |
| | $(0.016)$ | $(0.009)$ | $(0.005)$ | $(0.004)$ |
| Without anticipation | $-0.092^{***}$ | $0.003$ | $-0.028^{***}$ | $-0.026^{***}$ |
| | $(0.016)$ | $(0.009)$ | $(0.005)$ | $(0.004)$ |

*Notes:* Each cell reports the coefficient on *post* with two-way clustered standard errors (by facility and month) in parentheses. Panel A reports levels (HPPD); Panel B reports logs (HPPD). Samples: *With anticipation* ($N_{\text{levels}} = 632{,}231$; $N_{\text{logs}} = 632{,}231$). *Without anticipation* ($N_{\text{levels}} = 628{,}107$; $N_{\text{logs}} = 628{,}107$).
All specifications include facility and month fixed effects and covariates: *government*, *non-profit*, *chain*, *beds*, *occupancy rate*, *percent Medicare*, *percent Medicaid*, and state case-mix quartile indicators.
Statistical significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

### 4.2. Effects Before and During Covid-19 Pandemic

Table 3 examines whether these effects differ before and after the COVID-19 pandemic. The magnitude of the decline is roughly halved in the post pandemic period for RN, CNA, and total HPPD, suggesting that pandemic induced shifts in labor markets and staffing substantially reduced the impact of ownership transitions.

**Table 3:** TWFE Estimates of *post*: Pre- vs Post-pandemic Periods (Without anticipation)

| | Outcomes | | | |
|---|---|---|---|---|
| | **RN** | **LPN** | **CNA** | **Total** |
| **Panel A: Staffing Levels in HPPD** | | | | |
| Pre-Pandemic Period (2017/01 - 2019/12) | $-0.029^{***}$ | $-0.016^{*}$ | $-0.063^{***}$ | $-0.108^{***}$ |
| | (0.006) | (0.008) | (0.013) | (0.019) |
| Pandemic Period (2020/04 - 2024/06) | $-0.018^{**}$ | $-0.000$ | $-0.041^{***}$ | $-0.058^{***}$ |
| | (0.007) | (0.008) | (0.015) | (0.020) |
| **Panel B: Log Staffing Levels in HPPD** | | | | |
| Pre-Pandemic Period (2017/01 - 2019/12) | $-0.068^{***}$ | $-0.009$ | $-0.032^{***}$ | $-0.033^{***}$ |
| | (0.024) | (0.012) | (0.006) | (0.006) |
| Pandemic Period (2020/04 - 2024/06) | $-0.063^{***}$ | $0.002$ | $-0.020^{**}$ | $-0.019^{***}$ |
| | (0.021) | (0.011) | (0.008) | (0.006) |

*Notes:* Each cell reports the coefficient on *post* with two-way clustered standard errors (by facility and month) in parentheses. Panel A reports levels (HPPD); Panel B reports logs (HPPD). Sample sizes: Row 1 ($N_{\text{levels}} = 255,005$; $N_{\text{logs}} = 255,005$), Row 2 ($N_{\text{levels}} = 359,413$; $N_{\text{logs}} = 359,413$).
All specifications include facility and month fixed effects and covariates: *government*, *non-profit*, *chain*, *beds*, *occupancy rate*, *percent Medicare*, *percent Medicaid*, and state case-mix quartile indicators.
Statistical significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.
Pre-pandemic 2017/01–2019/12; Pandemic 2020/04–2024/06.

*4.3. Effects by Chain Status*

Table 4 compares effects for facilities that began the study period as part of a chain and those that did not. Staffing reductions are noticeably larger among non chain facilities. CNA HPPD falls by 0.034 hours in chain facilities and by 0.079 hours in non chains, and total HPPD declines by 0.062 versus 0.113 hours, respectively. On a log scale, these differences translate into substantially larger percentage declines for non-chain facilities. LPN staffing again shows minimal movement.

**Table 4:** TWFE Estimates of *post*: Chain vs Non-chain Facilities (Jan 2017 Baseline, Without anticipation)

| | Outcomes | | | |
|---|---|---|---|---|
| | **RN** | **LPN** | **CNA** | **Total** |
| **Panel A: Staffing Levels in HPPD** | | | | |
| Chain January 2017 | $-0.031^{***}$ | 0.003 | $-0.034^{***}$ | $-0.062^{***}$ |
| | (0.007) | (0.008) | (0.011) | (0.015) |
| Non-chain January 2017 | $-0.037^{***}$ | 0.003 | $-0.079^{***}$ | $-0.113^{***}$ |
| | (0.009) | (0.011) | (0.019) | (0.025) |
| **Panel B: Log Staffing Levels in HPPD** | | | | |
| Chain January 2017 | $-0.099^{***}$ | 0.008 | $-0.019^{***}$ | $-0.020^{***}$ |
| | (0.021) | (0.013) | (0.006) | (0.005) |
| Non-chain January 2017 | $-0.085^{***}$ | 0.013 | $-0.037^{***}$ | $-0.032^{***}$ |
| | (0.030) | (0.017) | (0.010) | (0.008) |

*Notes:* Each cell reports the coefficient on *post* with two-way clustered standard errors (by facility and month) in parentheses. Panel A reports levels (HPPD); Panel B reports logs (HPPD). Sample sizes: Row 1 ($N_{\text{levels}} = 322,100$; $N_{\text{logs}} = 322,100$), Row 2 ($N_{\text{levels}} = 243,036$; $N_{\text{logs}} = 243,036$).

All specifications include facility and month fixed effects and covariates: *government*, *non-profit*, *chain*, *beds*, *occupancy rate*, *percent Medicare*, *percent Medicaid*, and state case-mix quartile indicators.

Statistical significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Baseline chain classification determined by facility status in January 2017.

Overall, ownership changes lead to sizable and persistent reductions in nursing staffing, with effects concentrated among RN, CNA, and total hours per patient day. These reductions are larger prior to the COVID-19 pandemic and among facilities that were not part of a chain at baseline. Together, the event study and TWFE results indicate that ownership transitions systematically worsen staffing conditions in nursing homes rather than inducing temporary adjustments.

### 4.4. Robustness

This section presents a series of robustness checks designed to assess the validity of our identifying assumptions and the sensitivity of our results to alternative model choices. Across all exercises, the estimated effects of ownership change on staffing remain stable in sign, magnitude, and statistical significance. Most of these will have tables/plots that will go in the appendix.

### 4.4.1. Parallel Trends

Here I will present my table of p-values for the pre-treatment coefficients, and discuss their meaning. Wald Test.

### 4.4.2. Assumptions

In this section I want to present results for event study and TWFE models that rely on different assumptions than the one that I used in my main specification.

### 4.4.3. Event Window and Anticipation

Staffing decisions may adjust in anticipation of formal ownership transitions, potentially violating the parallel trends assumption if these periods are included. To address this concern, we estimate our event study and TWFE models using alternative event windows that exclude months immediately preceding the ownership change and show that the estimated post-treatment effects remain largely unchanged.

### 4.4.4. Gap Size

Because staffing data are occasionally missing or reported with gaps, we examine whether our results are sensitive to the length of time between consecutive observations. Restricting the sample to facilities with smaller reporting gap shows that estimates are similar to those in the full sample, indicating that data gaps do not drive our findings.

### 4.4.5. For Profit vs Not for Profit

Ownership changes may have different implications for staffing depending on a facilitys ownership type. We therefore estimate separate effects for for-profit and not-for-profit nursing homes.

## 5. Discussion