A Additional detail on data and methods

The PBJ are daily employee-level staffing data for each facility, and are required reporting for all direct care staff. For each employee working at a given facility on a given day, the PBJ reports the staff type, contract type, and number of hours worked. We use the data from 2018Q1-2023Q2.

We limit our analysis to CNAs, LPNs, and RNs, including those with administrative duties. We construct weekly measures (Monday-Sunday) for each facility of the hours worked, number of unique employees who worked at a facility, number of unique employees on payroll, and new hires and departures. To construct a consistent measure of new hires and departures over time, we implement a 12-week look-back (look-forward) for new hires (departures)—that is, any employee who worked their first week after 12 weeks of not working is considered a new hire, and any employee who worked their last week before at least 12 weeks of not working is considered a departure. An employee is "on payroll" for all weeks starting with their hire date, and ending with their departure date, i.e. the number of employees on payroll will include both employees who worked a shift that week and employees who were "absent" that week.

In order to identify each employee's hire date and separation date, it is necessary to have continuous reporting by facilities. Thus, for our outcome measures regarding new hires, separations, and payroll, we limit our sample to the last period of continuous reporting for each facility. In addition, although the PBJ data are generally of high quality due the potential for government audit, there are some instances in the data where nearly all of a facility's employee IDs changed in a given week and were replaced by an entire new staff of employee IDs. These changes likely reflect errors in the data (caused by, for example, a new payroll software) rather than real turnover. We identify these "software changes" as instances where the total nursing staff count changed by no more than 25% relative to the previous week but more than 50% of the IDs were new and again keep for each facility only the period following the last software change, if any changes were identified. Together, these changes exclude 10.3% of facility-weeks from our sample.

Finally, even after removing obvious software changes, there are some instances of implausibly high turnover at the start of new quarters. To avoid these outlier values creating potential bias in our results, we exclude the week of the first day of each quarter from our analysis for all dependent variables that count unique employees. We exclude that "first-week-of-the-quarter" week and the week prior for our measure of new hires, and we excluded the "first-week-of-the-quarter" week and the week following for our measure of departures.

B Additional detail on Illinois's Medicaid reform

B.1 Timing of the reform

HB0246 was introduced in the House on January 25, 2021. The bill passed both houses on April 7, 2022 and Governor Pritzker signed the bill into law on May 31, 2022.

To allow for staffing data to become available, the staffing incentive payment uses a three-quarter lag, meaning payments made in 2022Q3 (the first quarter after the reform) were based on staffing data from 2021Q4. Because this means that for the first two quarters of payments made following the reform (2022Q3 and 2022Q4 payments), the relevant staffing quarters had already passed prior to the reform's passage, the bill phased in the staffing incentive by paying facilities the greater of their actual staffing payment and the incentive payment for a facility that staffed at 85% of STRIVE for those two quarters (i.e. facilities that staffed below 85% STRIVE in those two quarters were paid as though they had staffed at 85% STRIVE). The full staffing incentive schedule took effect for 2022Q2 staffing; for this reason, we call April 1, 2022 the "effective" date of the reform.

B.2 Entire reform package

The bill allocated $717 million in additional funding to nursing homes.[8]

The $717 million budgeted amount includes the following, by far the largest of which is the STRIVE incentive payment:

1. **STRIVE incentive program, $360 million:** See Section 2.2 for details.

2. **Voluntary CNA subsidy program, $85 million:** Facilities that choose to participate in this component must establish a CNA pay scale that pays CNAs with more years of industry experience higher wages: the scale must increase wages for CNAs with one year of experience by at least $1.50 per hour (relative to CNAs with less than one year of experience), and offer an additional $1 per hour for each additional year of experience beyond the first year, up to a maximum of $6.50 for CNAs with at least 6 years of experience. Medicaid will fully subsidize wage increases of these amounts; facilities can increase wages more steeply with experience, but those wage increases would not be subsidized. Facilities can also choose to participate in a promotion pay scale that increases hourly wages by $1.50 per hour for CNA who are promoted, capped at 15% of the facility's CNA workforce.

[8] Because these reforms are partially funded through an increase in a tax on nursing homes, the estimated net increase is about $465 million per year.

3. **Quality Incentive Payments, $70 million:** This component divides $17.5M each quarter among nursing homes based on their long-stay quality star rating. The incentive payments are calculated such that facilities receiving 1 star receive no payment, and facilities receiving 2-5 stars are paid proportionally to the number of extra stars above 1.

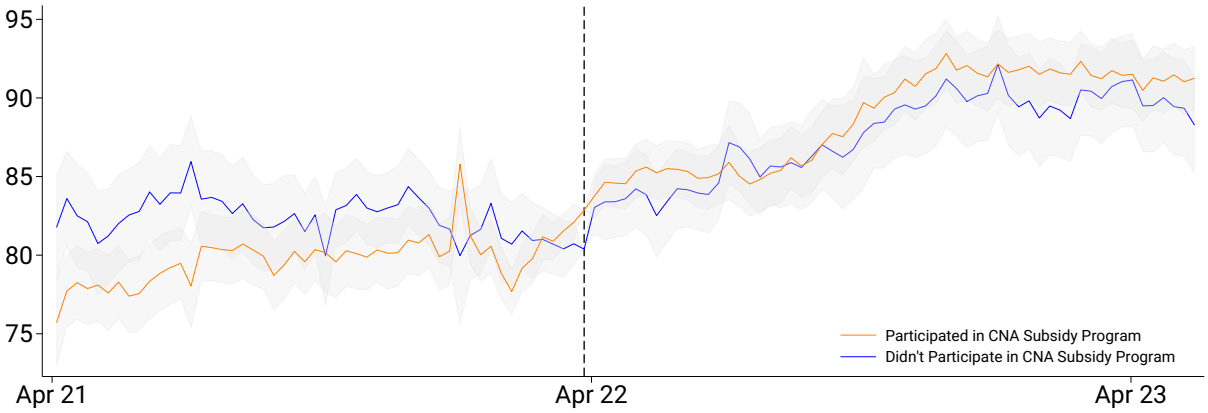
4. **Across-the board per-diem increase, $202 million:** This component increases the base per-day reimbursement rate by $7 per resident for all facilities, and by an additional $4 per day for facilities in which Medicaid covers at least 70% of their residents.

The main challenge to attributing our findings to the STRIVE incentive program is that the reform also included a voluntary CNA experience program. However, it is unlikely that the experience program is the primary driver of our results for several reasons. First, while we cannot measure the year-of-experience variable that is used to determine the CNA payscale, the fact that we see substantial hiring rather than greater retention or greater hours suggests that firms were not only increasing staffing among the more experienced workers. Second, not all facilities chose to participate in the CNA subsidy program. As of April 2023, 402 Illinois facilities (54%) had elected to participate in this program. In Figure B.1 we compare the change in target staffing for these 402 facilities to the change in target staffing for the remaining facilities that did not participate, splitting the sample by whether the facility was a high- or low-Medicaid facility pre-reform.[9] Focusing on high-Medicaid facilities, we find that the increases in staffing are very similar for facilities participating in the program and facilities that did not participate in the program.

We include all components of the reform in additional cost estimates in Table B.1, where we replicate the cost effectiveness analysis of Section 4.3 using the *total* change in Medicaid reimbursement as the cost to the government (rather than only the STRIVE incentive payments). Given the size of the "level" shifts included in the package, the full reform was much less cost effective than only the incentive program, even for high-Medicaid facilities.

[9]High-Medicaid facilities were disproportionately likely to opt in to the policy: 72% of high-Medicaid facilities opted in, compared to only 48% of low-Medicaid facilities.

(a) High-Medicaid facilities



(b) Low-Medicaid facilities

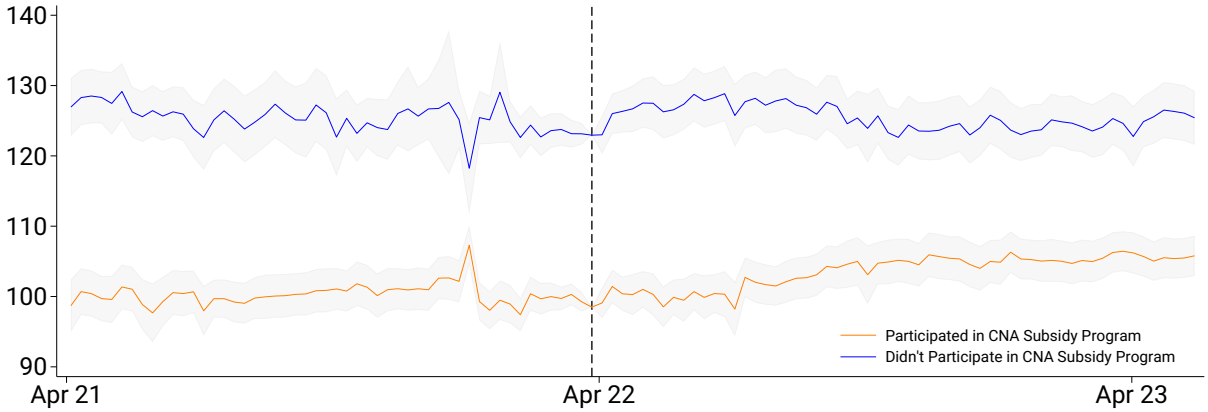


Figure B.1: Event study of staffing, by CNA incentive participation

Notes: Regression coefficients and 95% confidence intervals from event study regression. Each point estimate is added to the baseline average value for IL facilities in the pre-treatment period. Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the rate reform: April 1, 2022.

	All Facilities (1)	High-Medicaid (2)	Low-Medicaid (3)
Additional Staffing Expenditure ($)			
CNA	28,083.14	49,189.38	13,238.61
LPN	7,722.45	15,412.13	2,314.10
RN	-777.53	15,696.08	-12,363.84
Total Value of Additional Staffing ($)	36,447.68	80,497.07	3,319.62
Additional Total Payment from Reform ($)	196,310.78	244,430.66	160,121.44
Implied Passthrough (%)	18.57	32.93	2.07
Medicaid Utilization (%)	62.07	78.47	49.75
Implied Medicaid-Only Passthrough (%)	14.69	26.52	1.11

Table B.1: Assessment of Cost of Total 2022 Reform

Notes: Table provides a cost effectiveness calculation from the entirety of the 2022 reform. The top panel contains the additional staffing expenditure in 2023Q1 implied by the point estimates. The market price of the marginal hours is calculated using wages and benefits from Medicare cost reports. Hourly costs for CNAs, LPNs, and RNs are $23.40, $38.52 and $45.19, respectively. To calculate the additional total payment of the reform, we compute the facility-level changes in Medicaid reimbursement between 2022Q2 and 2023Q1. The implied passthrough is the ratio of the value of the marginal staffing expenditure over the additional quarterly payment under the reform. The Medicaid-only passthrough considers only the benefit of additional staffing accrued to Medicaid patients. Separate estimates are given for all facilities (column (1)), high-Medicaid facilities (column (2)), and low-Medicaid facilities (column (3)).

C Labor demand elasticity calculation

In this section, we describe the procedure to calculate the labor demand elasticity at each facility implied by the treatment effect estimates.

The facility-specific elasticity, ϵ_i, is the percentage change in labor divided by the percentage change in the effective wage. The numerator is straightforward to compute. For each facility i, the post-treatment staffing level y_i is observed. To determine the percentage change in y_i, we calculate the counterfactual staffing level $\tilde{y}_i = y_i - \hat{\beta}$ that facility i would have had in the absence of treatment, where $\hat{\beta}$ is the pooled treatment effect estimated in Section 4.1. We allow high- and low-Medicaid facilities to vary in their treatment effects, as described in the main text. The relative change in labor is then given by $(y_i - \tilde{y}_i)/\tilde{y}_i$.

The percentage change in effective wages is determined by two components: the observed wage cost of moving from $\tilde{y}_i$ to y_i and the marginal incentive payment the facility receives for doing so. Recall that y_i is a ratio (the level of staffing relative to the target level per resident day based on the facility's patient case mix). To convert this ratio to marginal hours, we multiply y_i by the staffing the facility reports in the post-treatment period by the number of observed patient days, to compute the number of marginal staff hours induced by the policy. To determine the marginal cost of these hours, we use facility-level CNA/LPN/RN wages reported in the federal HCRIS cost report data for the years 2021 and 2022. Given the noise inherent to the HCRIS data, we winsorize wages at the 1^{st} and 99^{th} percentiles. We compute the facility-level average hourly wages using each facility's skill mix (CNA/LPN/RN shares of total direct care hours). Denote the observed wage cost of the marginal hours as w_i.

Finally, we can compute the additional incentive payments each facility receives as a result of changing their staffing from $\tilde{y}_i$ to y_i using the known reimbursement schedule shown in Figure 1. We take the product of the incentive payments and the number of Medicaid days (using the baseline Medicaid shares used throughout). This provides the marginal incentive payments from the additional hours, denoted p_i. The effective wage cost for facility i moving from $\tilde{y}_i$ to y_i is then given by $w_i - p_i$, i.e. the wage cost net of the marginal incentive payment. Accordingly, the percentage change in wages is given by: $\big((w_i - p_i) - w_i\big)/w_i) = -p_i/w_i$.

The labor demand elasticity for facility i is then:

$$\epsilon_i = \frac{(\tilde{y}_i - y_i)/\tilde{y}_i}{-p_i/w_i}$$

The distribution of ϵ_i is shown in Figure C.1. The figure is truncated at -0.55, approximately the 2^{nd} percentile. The mean elasticity is -0.142, well within the range of conventional estimates, particularly for short-term responses and for the US labor market (Lichter

et al., 2015). This mean elasticity includes considerable heterogeneity: high-Medicaid facilities are more elastic, with a mean elasticity of −0.200, compared to only −0.097 for low-Medicaid facilities.

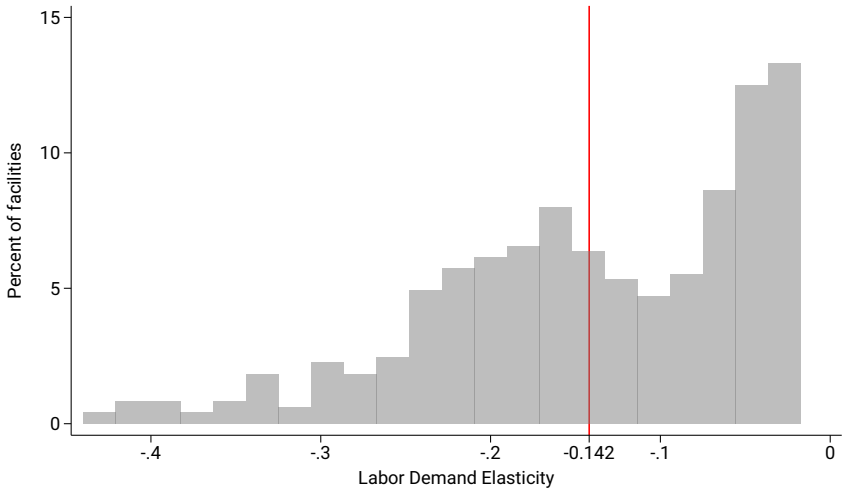


Figure C.1: Distribution of Implied Labor Demand Elasticities

Notes: Figure presents estimates of the labor demand elasticities implied by the change in labor and change in effective wages at each firm. Vertical red line denotes the overall mean elasticity.

D Effects of the reform on clinical quality

In light of the large staffing increases documented in the paper, this section examines the impact of the rate reform on patient health.

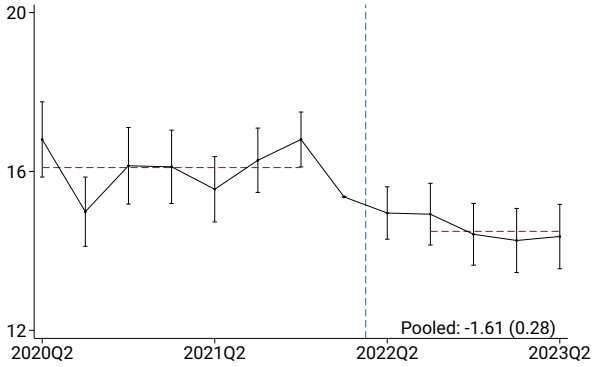
The quarterly NHC data provide several measures of resident health, however, many of these measures are unlikely to be sensitive to changes in direct staffing measures, especially in the short-term. Shen et al. (2023) show that measures related to functioning of long-stay residents (in contrast to short-term patients whose stays are primarily focused on rehabilitation, and more process-based measures such as the use of antipsychotic medications) are particularly sensitive to changes in staffing, and so we focus our attention on these measures. Specifically, we examine the share of long-stay residents whose need for assistance with the activities of daily living (ADL) increased, as well as the share of long-stay residents whose mobility worsened. Since both outcomes are measured quarterly, we modify equation (1) to include quarter, rather than week, fixed effects. Since the quarterly data provide fewer observations over a given time period, we also expand the sample to include data beginning in 2020Q2 to allow us to better assess the plausibility of the parallel trends assumption.

Figure D.1 shows modest improvements in both measures of resident health. The share of long-stay residents whose need for ADL help increased over the prior period fell by 1.61 percentage points at the end of the sample period. Similarly, the share of long-stay residents whose mobility worsened fell by 1.47 percentage points by the end of the sample period. These estimates are suggest potentially larger effects for high-Medicaid facilities, for whom the improvements in these quality measures were 2.05 and 1.50 percentage points, respectively, compared to 1.14 and 1.62 percentage points for low-Medicaid facilities.

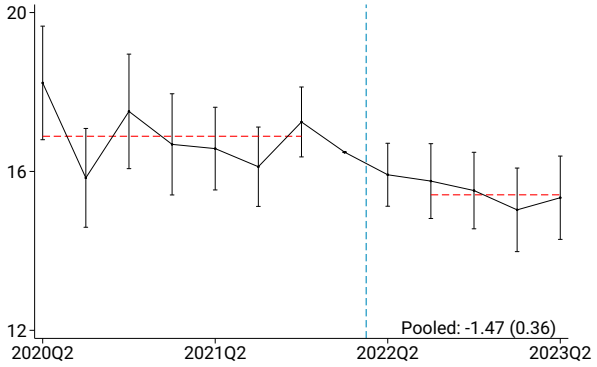
Prior literature finds a wide range of effects of the relationship between staffing and resident health: a 1% increase in staffing has been shown to improve outcomes by anywhere from 0.002-0.012 standard deviations.[10] Our estimates indicate that the reform led to a 5.4% increase in the staffing level, a 1.6 percentage point decline in the share of patients who need help with ADLs and a 1.5 percentage point decline in the share whose mobility worsened. These results correspond to a 1% increase in target staffing leading to 0.033 and 0.030 standard deviation reduction in these two clinical outcomes, respectively. Given that the reform we study was explicitly intended to increase staffing and improve clinical outcomes, it is unsurprising that we find larger results that past studies.

[10]For instance, two recent studies, Furtado and Ortega (2023) and Grabowski et al. (2023), both examine the impact of immigration on nursing home staffing and clinical outcomes. They find that a 1% increase in staffing corresponds to increases in quality outcomes of 0.002 and 0.012 standard deviations, respectively.

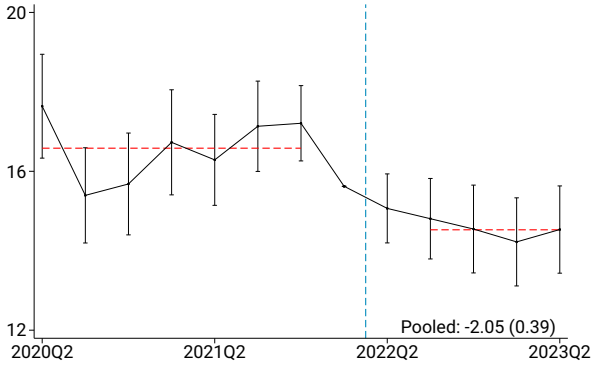
(a) Share of residents whose need for ADL help increased



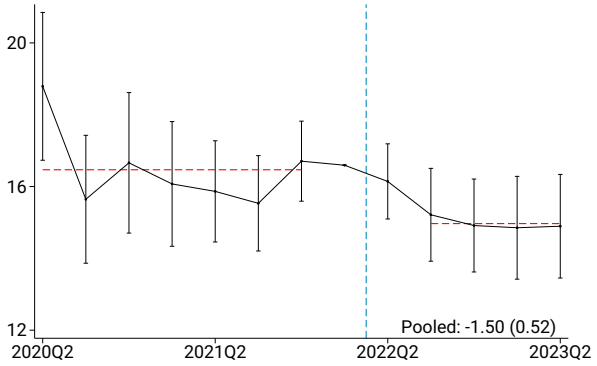
(b) Share of residents whose mobility worsened



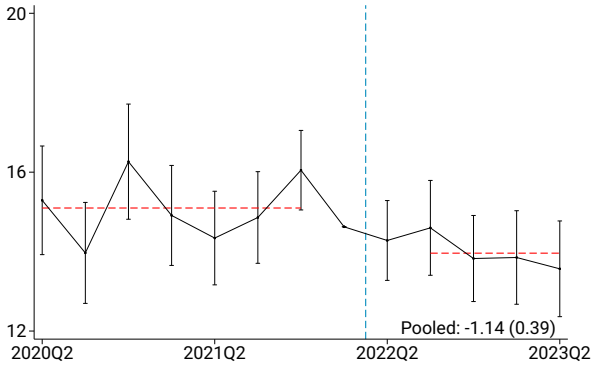
(c) Share of long-stay residents whose need for ADL help increased, high-Medicaid facilities



(d) Share of long-stay residents whose mobility worsened, high-Medicaid facilities



(e) Share of long-stay residents whose need for ADL help increased, low-Medicaid facilities



(f) Share of long-stay residents whose mobility worsened, low-Medicaid facilities

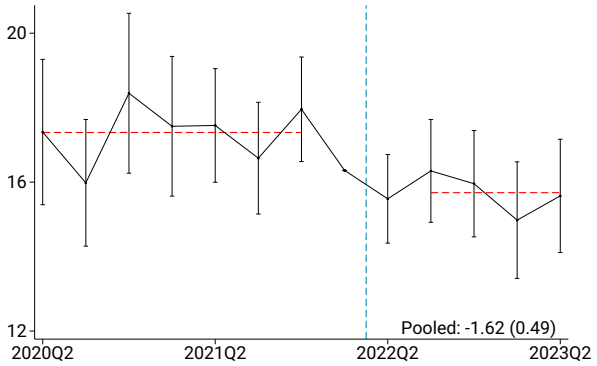


Figure D.1: Event study of staffing-sensitive clinical quality measures

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Vertical bars denote 95% confidence interval with robust standard errors clustered by facility. Dependent variables are quarterly facility quality measures from NHC. The left column is the share of long-stay residents whose need for ADL help increased relative to the previous quarter. The right column is the share of long-stay residents whose mobility worsened relative to the previous quarter. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

E Additional results

This section includes additional analyses that supplement the analyses in the main text.

1. Table E.1 shows summary statistics of the Illinois and non-Illinois facilities included in our analysis.

2. Figures E.1 and E.2 replicate the analysis in Figure 4 (extensive and intensive margins of staffing adjustment) but for all facilities and low-Medicaid facilities, respectively. In general, the staffing results are larger in high-Medicaid facilities than low-Medicaid facilities.

3. Figure E.3 replicates the staffing adjustment margins analysis in Figure 4, focusing on CNAs, the occupation most affected by the reform.

4. Figure E.4 includes additional outcomes.

   (a) Figure E.4a and E.4b confirm that the increase in staffing levels is not the result of facilities limiting or changing their admissions practices to reduce their census or case-mix and thus inflate their staffing ratios. We observe a slight increase in average resident count in Illinois after the policy, and no change in case mix.

   (b) Figure E.4c shows that the policy did not change the share of staff hours that are worked on weekends, and Figure E.4d shows that it did not change the share of overtime hours (hours after the 40th hour in a week for a given employee).

   (c) Figures E.4e, E.4f complement the findings in Figure 4 of an increase in both part-time and full-time staff by decomposing the increase in staffing hours by part-time (< 35 hours per week) and full-time (≥ 35 hours per week) workers, finding that full-time workers are contributing slightly more to the overall increase. Figure E.4g shows that the share of hours coming from part-time workers is unchanged, and Figure E.4h shows that the share of workers on staff who were part-time workers increased slightly as a result of the reform.

   (d) Figures E.4i and E.4j complement the findings in Figure 4 by decomposing the overall staffing increase into hours worked by contract workers and hours worked by non-contract workers. We find that hours worked by non-contract workers explain almost all of the increase. Nonetheless, the increase in contract staff is related to a slight increase in the share of hours worked by contract workers and a slightly larger increase in the share of workers who are contract workers.

5. Figure E.5 presents a piecewise analogue to Figure 3, with separate quadratic polynomials estimated over different regions of the payment schedule.

6. Figure E.6 illustrates the time series variation in daily Medicaid rates over our sample period.

7. Figure E.7 presents an event study of the 2019 reform using a continuous treatment design. This design leverages the fact that the 2019 reform contained two components: one component of the 2019 reform increased the per-diem for Medicaid residents by $4.55 a day, and a second component increased the "support rate" – that is, allocated funds to facilities based on *historical* costs. Since this second component varied across facilities, we construct an "exposure" variable equal to the additional payment per resident that a facility would be expected to receive as a result of the rate increase, equal to the facility's share of residents who are on Medicaid multiplied by the change in the Medicaid per-diem for that facility between 2019Q2 and 2019Q3. We then estimate an identical event study specification of the form:

$$y_{it} = \sum_{\tau \neq -1} \beta^{\tau} (IL_i \times \text{Exposure}_i \times d_t^{\tau}) + \alpha_i + \alpha_t + \varepsilon_{it} \tag{2}$$

8. Figure E.8 presents the histograms of marginal hours induced by the reform used for the cost calculation conducted in Section 4.3.

9. Table E.2 presents the cost calculations underlying the assessment in Section 4.3.

	Illinois		Non-Illinois	
	Pre-Policy	Post-Policy	Pre-Policy	Post-Policy
	(1)	(2)	(3)	(4)
STRIVE Ratio	99.20	102.94	114.25	113.48
CNA STRIVE	56.89	61.59	64.77	65.49
LPN STRIVE	19.54	20.01	28.55	28.34
RN STRIVE	22.78	21.33	20.93	19.65
Total Payroll	66.31	77.77	69.05	73.97
Total Staff Working Any Shift	54.76	63.46	57.84	61.73
Hours per Employee	31.34	30.46	32.13	31.64
Staff Working $<$ 20 hours	14.70	19.31	13.78	15.66
Staff Working 20-34 hours	16.72	18.83	17.78	18.47
Staff Working $\geq$ 35 hours	23.34	25.32	26.28	27.60
Weekly New Hires	3.13	4.78	2.66	3.23
Weekly Separations	3.03	4.74	2.66	3.24
Share Hours by Contract Staff	7.31	9.79	8.07	9.81
% ADL Need Increased	13.87	12.43	14.66	14.62
% Mobility Worsened	17.40	13.77	17.62	14.92
N	697	694	14,397	14,286

Table E.1: Summary Statistics

Notes: Table provides summary statistics for all facilities studied, from the pre- and post-policy periods for each outcome. The pre-period ranges from 2021Q2 through 2022Q1. The post-period ranges from 2022Q2-2023Q2.

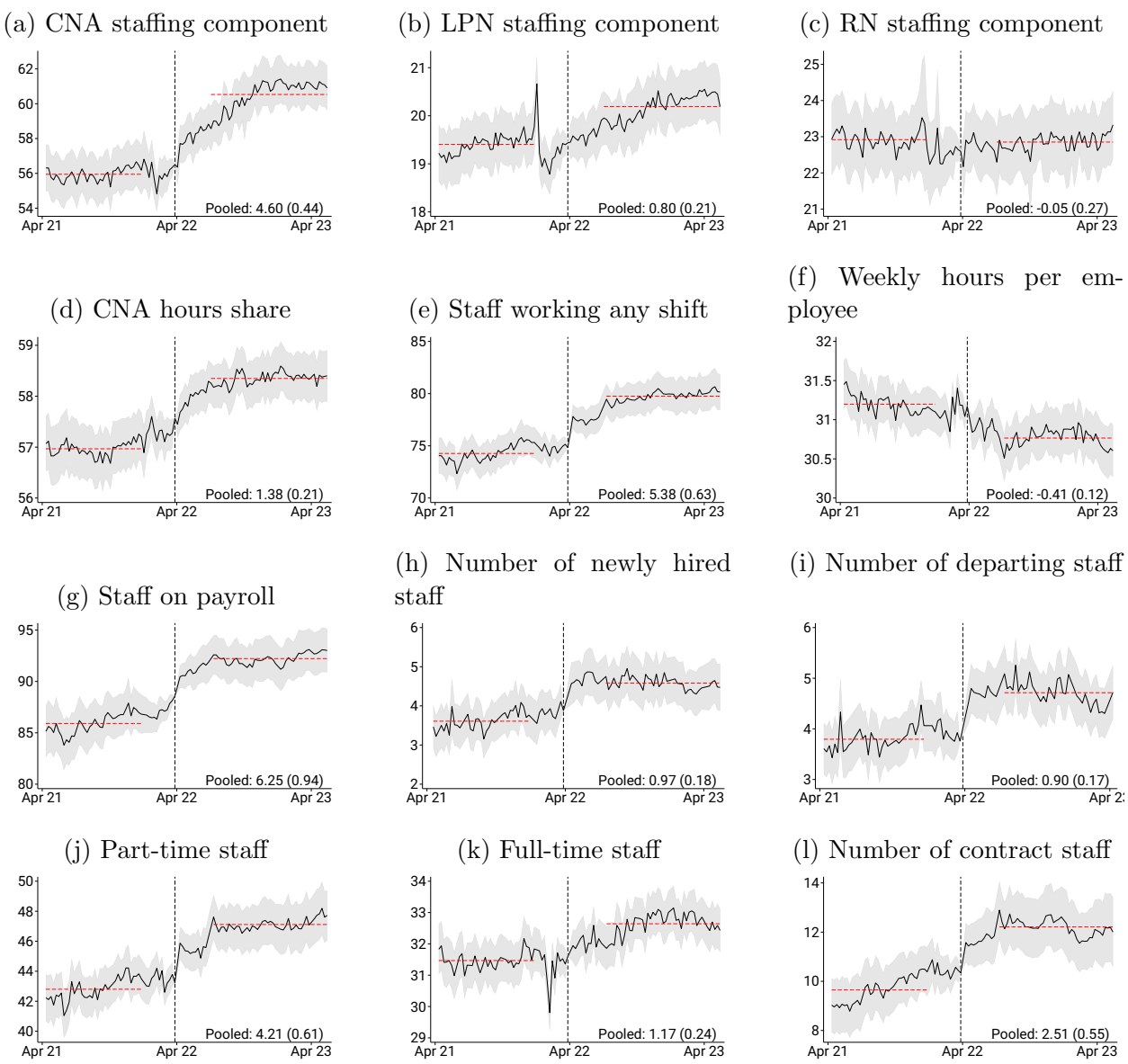


Figure E.1: Event study of different margins of staffing adjustment, all facilities

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Employee counts are scaled by the facility's average daily resident census (per 100 residents). Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

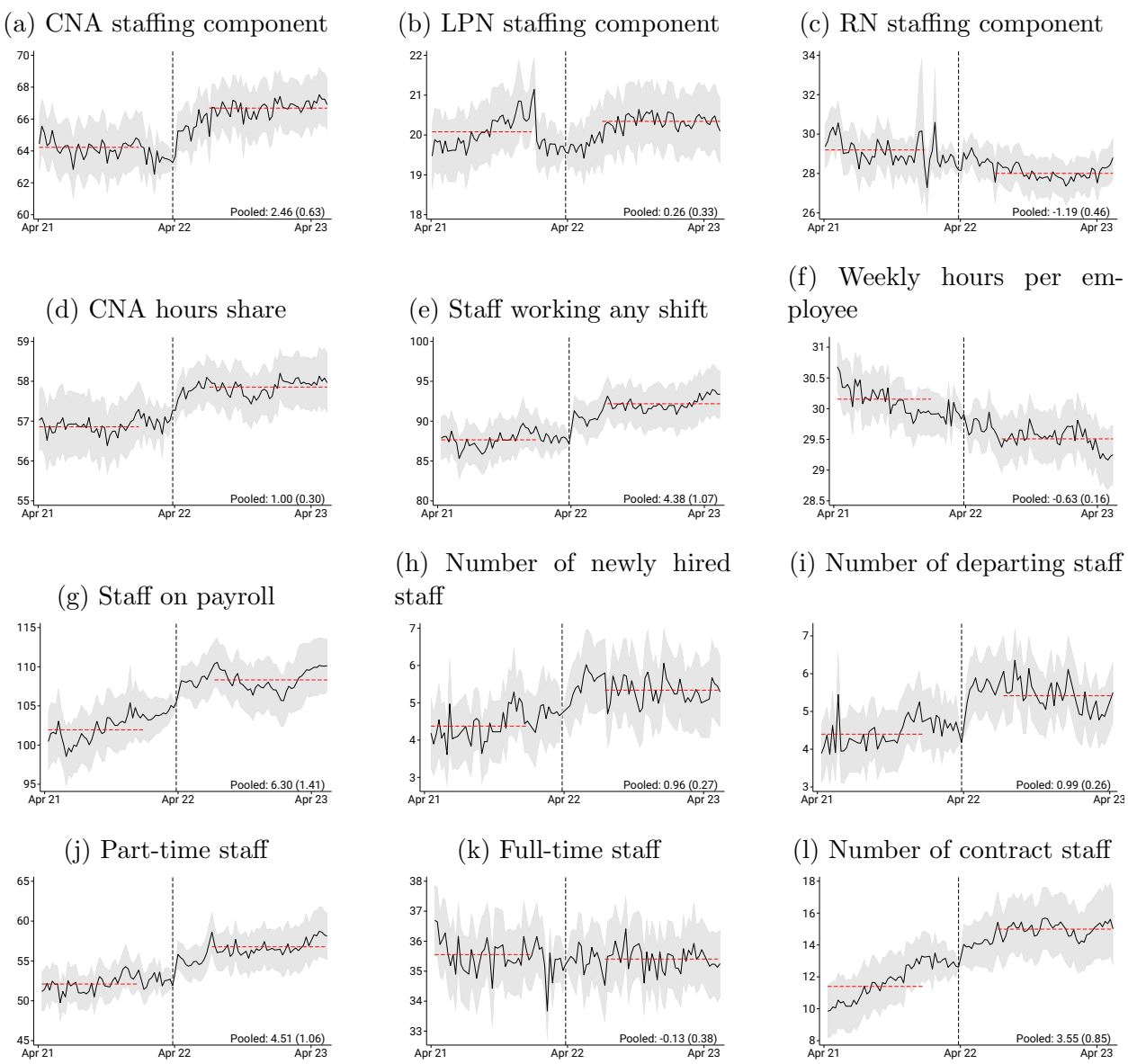


Figure E.2: Event study of different margins of staffing adjustment, low-Medicaid facilities

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Low Medicaid facilities defined as whether the facility had below the median share of Medicaid residents in 2019 (58.3%). Employee counts are scaled by the facility's average daily resident census (per 100 residents). Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

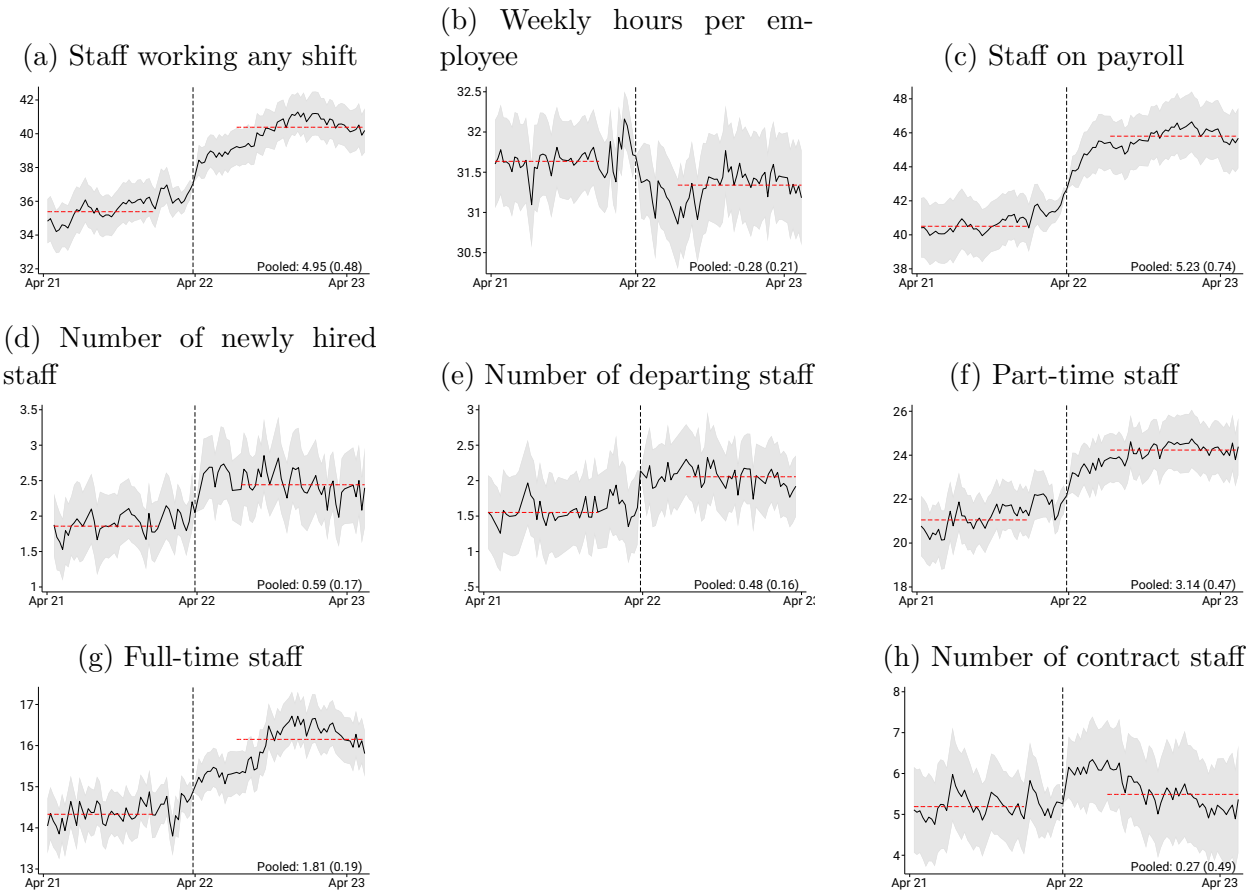


Figure E.3: Event study of different margins of staffing adjustment, CNA staff only, high-Medicaid facilities

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Sample limited to facilities with above the median share of Medicaid residents in 2019 (58.3%). Employee counts are scaled by the facility's average daily resident census (per 100 residents). Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

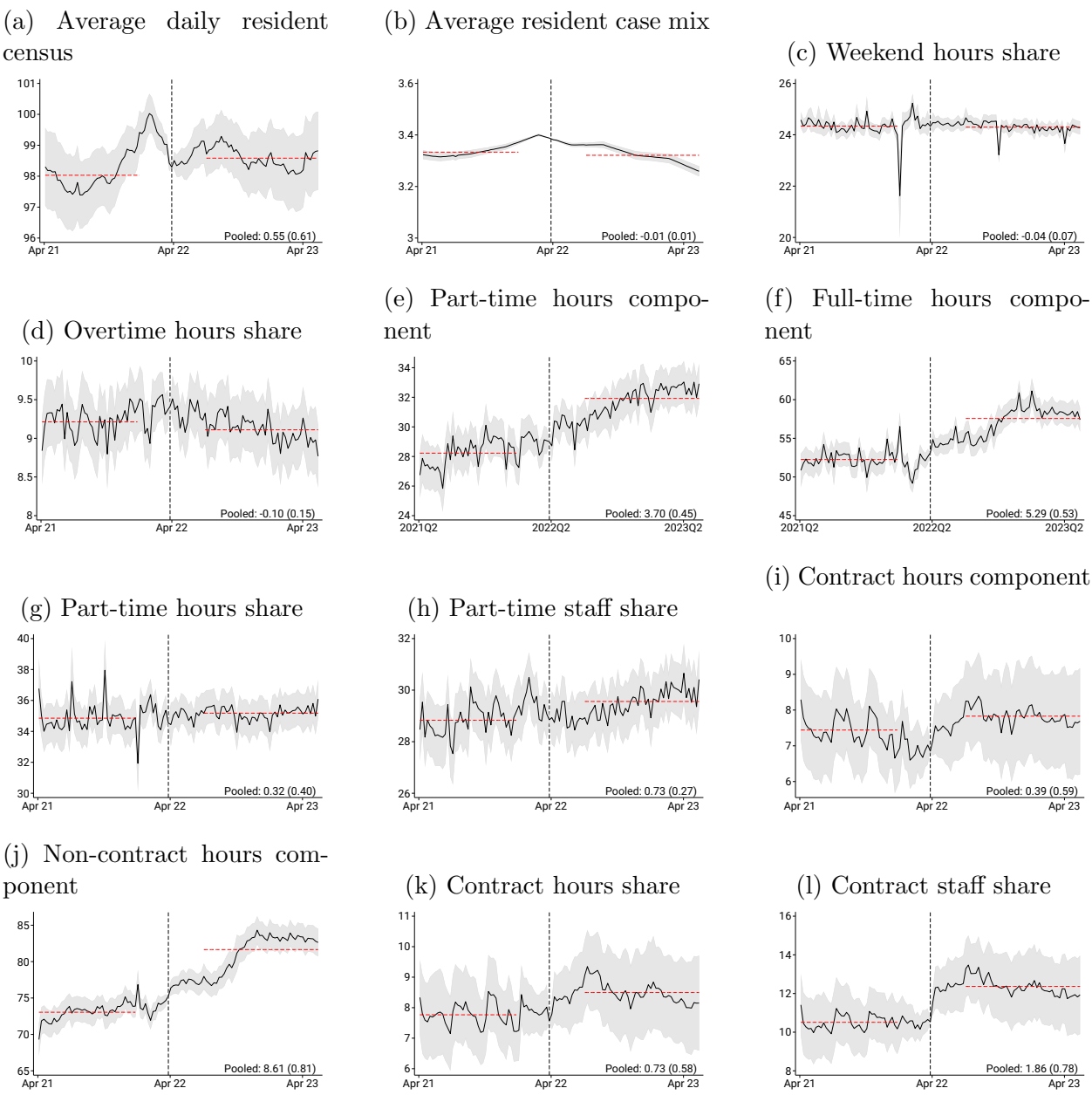


Figure E.4: Event study of additional dependent variables, high-Medicaid facilities

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Sample limited to facilities with above the median share of Medicaid residents in 2019 (58.3%). Employee counts are scaled by the facility's average daily resident census (per 100 residents). Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

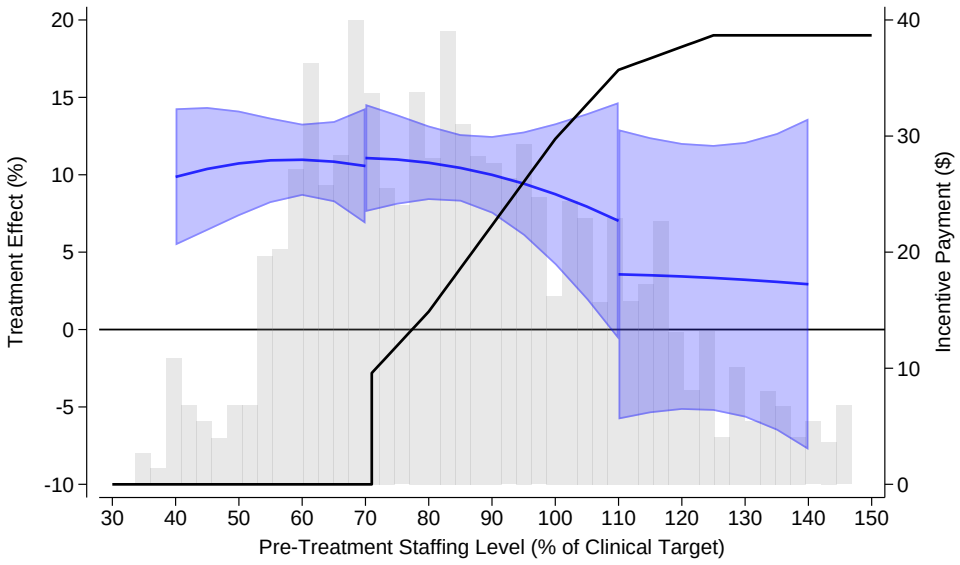


Figure E.5: Heterogeneity by pre-treatment target staffing level: Piecewise

Notes: Figure presents treatment effects across the pre-treatment staffing level distribution. Treatment effects are calculated using piecewise quadratic polynomials in pre-treatment staffing separately across different regions of the payment schedule. Treatment is assumed to scale linearly in Medicaid share. Shaded area denotes 95% confidence interval with robust standard errors clustered by facility. Baseline staffing distribution in gray histogram; black line denotes the incentive payment schedule.

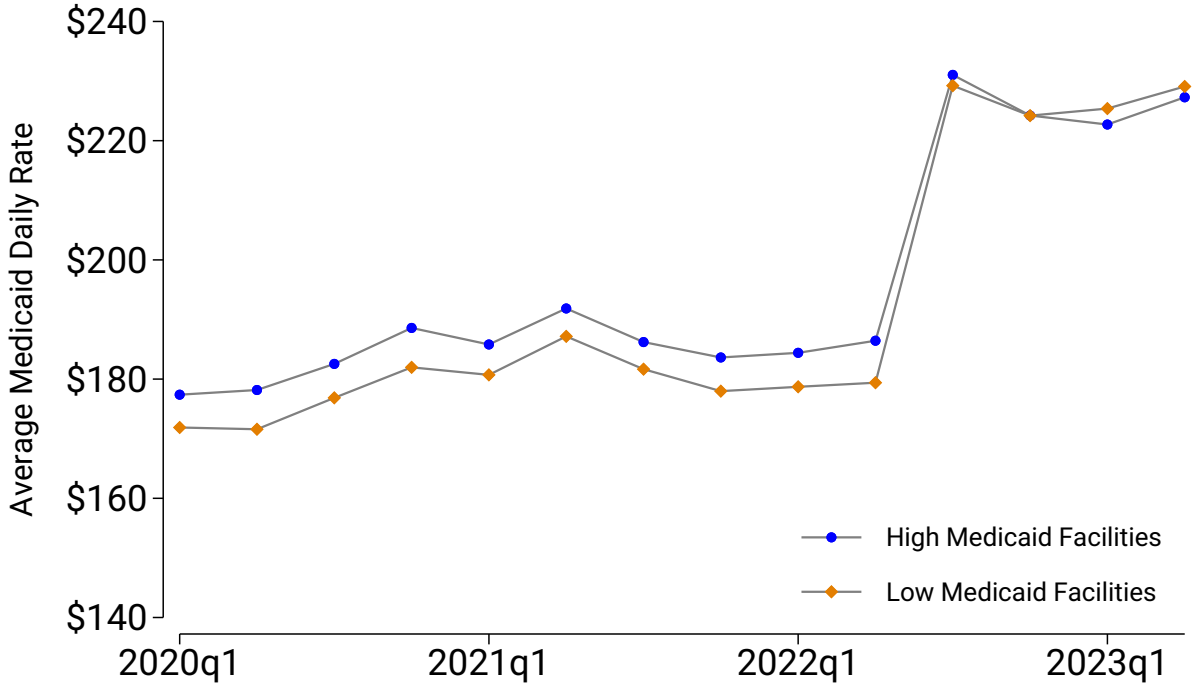


Figure E.6: Impact of reform on daily Medicaid rates

Notes: Average total Medicaid rate by quarter. Facilities are classified based on the median share of Medicaid residents in 2019 (58.3%)

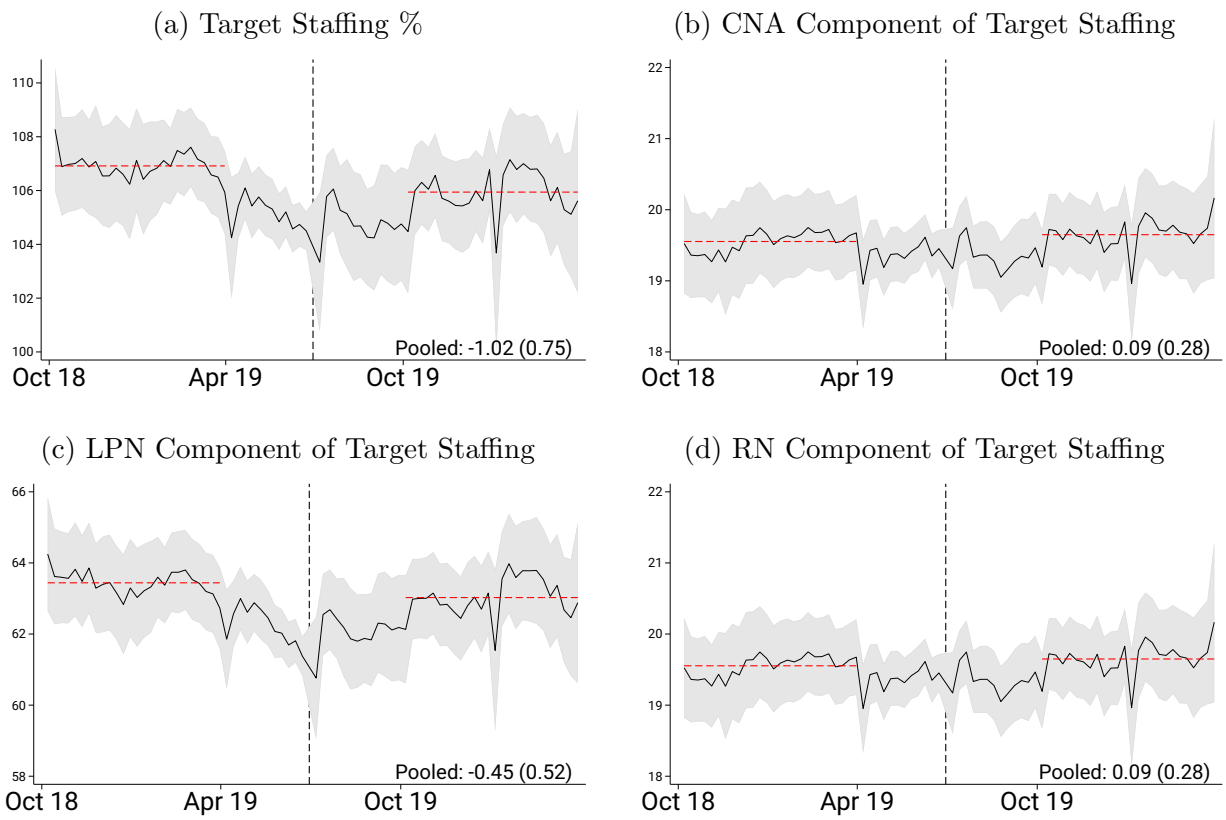
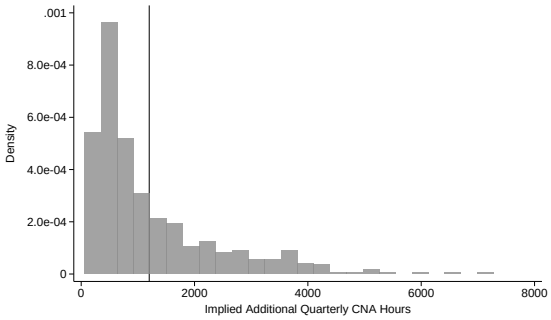


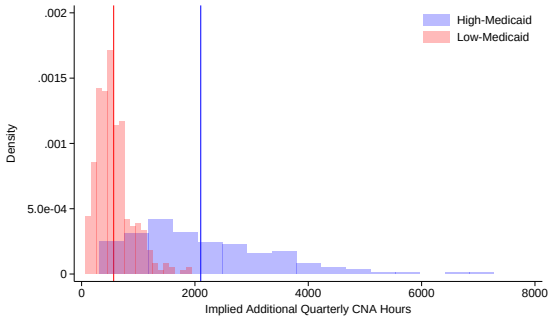
Figure E.7: Event study of staffing levels before and after 2019 reform, continuous treatment

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Dependent variable is total nurse (RN, LPN, and CNA) staffing hours, expressed as a percent of the STRIVE target staffing level. Independent variables are week indicators, interacted with each facility's "exposure" to the reform, defined as the change in reimbursement per resident day. Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

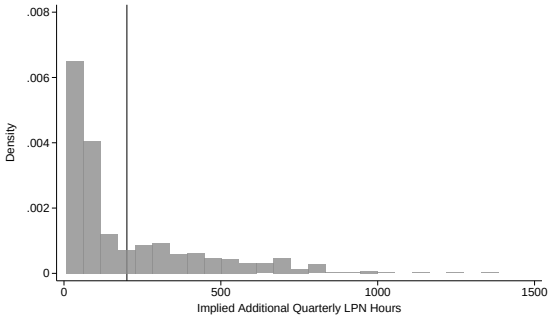
(a) Quarterly CNA hours: All Facilities



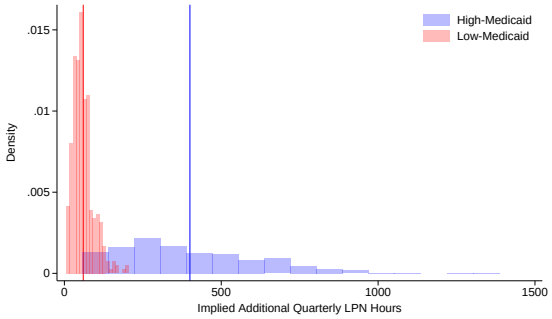
(b) Quarterly CNA hours: By Medicaid



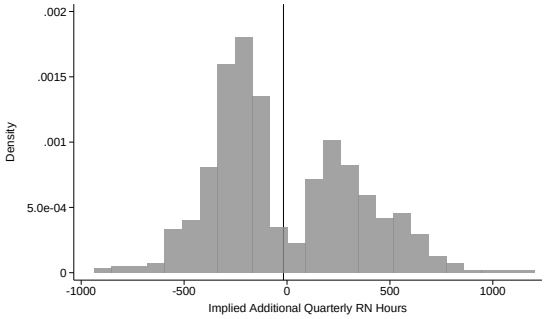
(c) Quarterly LPN hours: All Facilities



(d) Quarterly LPN hours: By Medicaid



(e) Quarterly RN hours: All Facilities



(f) Quarterly RN hours: By Medicaid

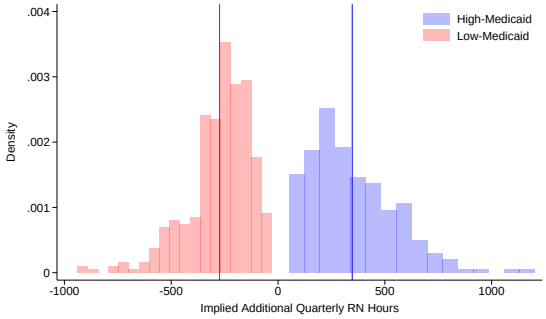


Figure E.8: Distributions of additional hours implied by estimates

Notes: Figure presents histograms of implied additional hours of nursing induced by the reform. These distributions, implied by the treatment effect estimates in Section 4.1, are used to determine the cost efficacy of the reform in Section 4.3.

	All Facilities (1)	High-Medicaid (2)	Low-Medicaid (3)
Additional Staffing Expenditure ($)			
CNA	28,083.14	49,189.38	13,238.61
LPN	7,722.45	15,412.13	2,314.10
RN	-777.53	15,696.08	-12,363.84
Total Value of Additional Staffing ($)	36,447.68	80,497.07	3,319.62
Average Facility Incentive Payment ($)	93,338.37	93,593.71	93,146.34
Implied Passthrough (%)	39.05	86.01	3.56
Medicaid Utilization (%)	62.07	78.47	49.75
Implied Medicaid-Only Passthrough (%)	30.90	69.26	1.91

Table E.2: Assessment of Cost of Reform

Notes: Table provides a cost effectiveness calculation from the incentive component of the reform. The top panel contains the additional staffing expenditure in 2023Q1 implied by the point estimates. The market price of the marginal hours is calculated using wages and benefits from Medicare cost reports. Hourly costs for CNAs, LPNs, and RNs are $23.40, $38.52 and $45.19, respectively. The average facility incentive payment reflects the payments received in 2023Q1. The implied passthrough is the ratio of the value of the marginal staffing expenditure over the additional quarterly payment under the reform. The Medicaid-only passthrough considers only the benefit of additional staffing accrued to Medicaid patients. Separate estimates are given for all facilities (column (1)), high-Medicaid facilities (column (2)), and low-Medicaid facilities (column (3)).

F Alternate sample definitions and matched control analysis

F.1 Alternative Sample Definitions

Extended pre-period: Our main sample period begins in 2021Q2, which allows us to observe facilities for a full year prior to the Illinois rate reform. In Figure F.1, we replicate the analysis of Figure 2, but extend the pre-period by a year to begin at the beginning of 2020Q2. This allows for an additional year of pre-period data while also continuing to exclude the first quarter of 2020Q1, which marked the start of the COVID-19 pandemic in the US and in which PBJ reporting was not mandatory. Pre-trends remain fairly stable with this additional year of pre-period data, although with generally more movement in 2020 than the year immediately prior to the reform. This volatility may reflect waves of the pandemic, which may have affected Illinois facilities differently than non-Illinois facilities.

Balanced panel: The main results use data for all facility-weeks that appear in the PBJ data, allowing facilities to appear for different numbers of weeks. To instead construct a balanced sample of facility-weeks, we retain the last period of reporting after any software change or gap in reporting, and require that facilities are observed for the entire sample period (11,287 facilities (73.6%) of the 15,346 facilities are included). Figure F.2 shows the results of our primary analysis in this fully balanced sample of facilities. The results are quantitatively similar.

F.2 Matched control analysis

In this section, we consider a matched control approach wherein we match each Illinois facility to a a narrower group of control facilities in other states that are most observationally similar to Illinois firms.

The matched control groups are constructed from *pre-pandemic* data – the last values observed for each matching variable in 2019. The matching variables include: ownership status, county population density, NHC Overall rating, share of patients on Medicaid, average daily number of residents, share of hours worked by new employees, and overall STRIVE ratio, as well as the STRIVE contributions coming from CNA, LPN, and RN workers separately. Given the overall STRIVE ratio is our primary outcome, we apply calipers of 15 percent to ensure similar matches on this variable. For each Illinois facility, there are 4-5 matched controls; any facility with fewer than 4 matches is excluded from the analysis.

Accordingly, our matched specification is given by:

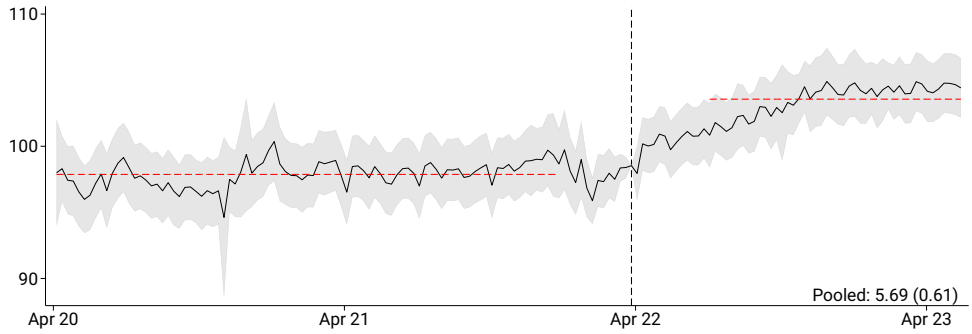
$$y_{ict} = \sum_{\tau \neq -1} \beta^{\tau}(IL_i \times d_t^{\tau}) + \alpha_{ic} + \alpha_{ct} + \varepsilon_{ict} \tag{3}$$

where i indexes facility, c indexes match cohort, and t indexes calendar week. Notice that our data vary at the facility-cohort level, as some non-Illinois facilities may be matched controls for multiple treated facilities. Accordingly we include both facility-cohort fixed effects, α_{ic}, to account for any residual facility-level differences within each cohort that persist after the matching exercise, as well as cohort-by calendar-week fixed effects, α_{ct}, which allow us to control for differential time trends across match cohorts. This specification is considerably more flexible than standard calendar year fixed effects, which impose the same time trend for all facilities.

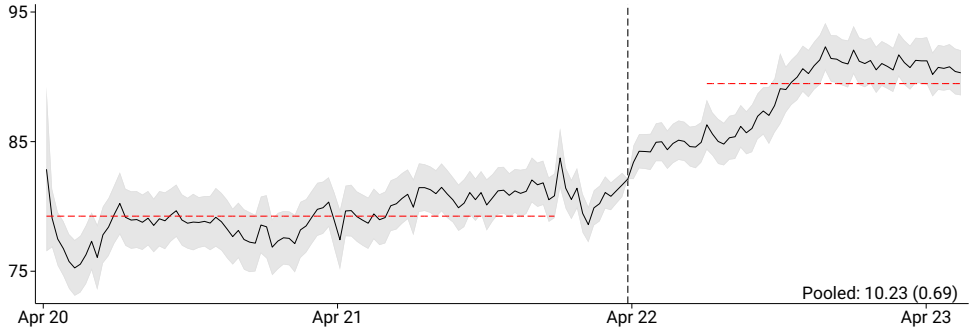
As in Equation 1, the remaining terms identify the treatment effect, under the identifying assumptions. The d_t^{τ} terms denote calendar-week dummies, and IL_i is an indicator for Illinois facilities. The coefficients of interest are the β^{τ} terms that capture the residual differences in y_{ict} of the Illinois facilities relative to their matched counterparts. The key identification assumption is that of parallel trends: *within* a match cohort, any differential patterns in y_{ict} are attributable to the Illinois reform.

The results of the matching process are shown in Appendix Table F.1. Relative to all non-Illinois facilities, the matched sample are substantially more similar on observable characteristics. The results from estimating equation (3) for our primary analysis are reported in Appendix Figure F.3a. The estimates are nearly identical to those recovered from the unmatched control group.

(a) Target Staffing: All Facilities



(b) Target Staffing: High-Medicaid Facilities



(c) Target Staffing: Low-Medicaid Facilities

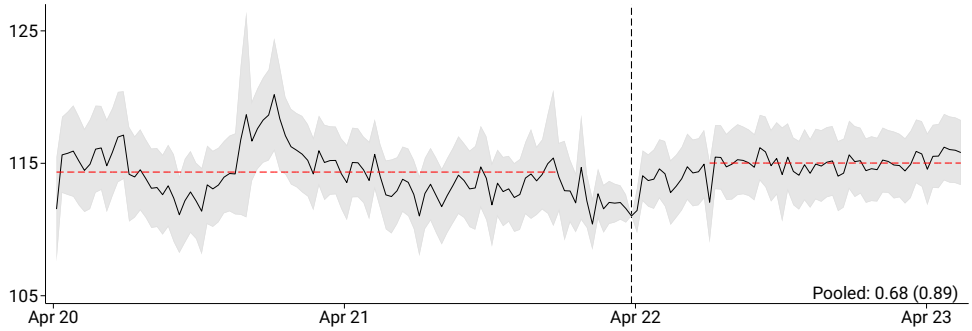
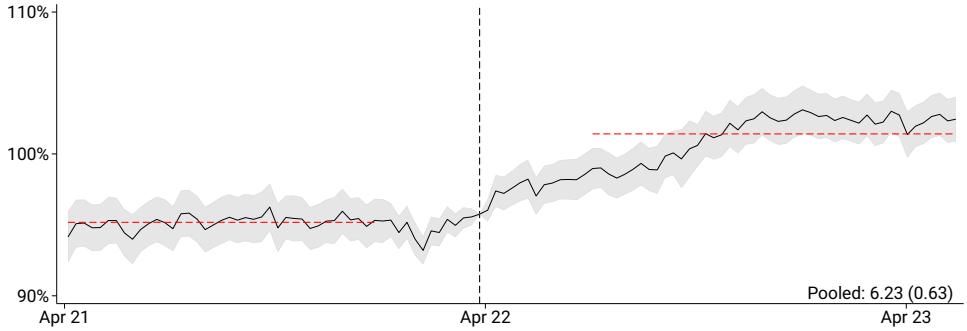


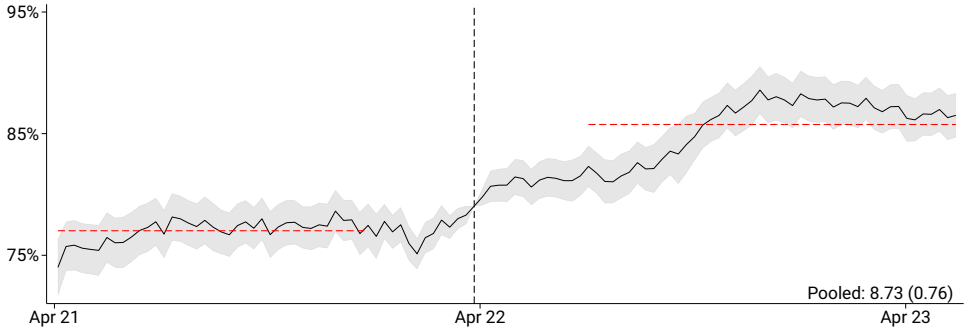
Figure F.1: Event study of staffing levels, by Medicaid payer share, extended pre-period

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

(a) Target Staffing: All Facilities



(b) Target Staffing: High-Medicaid Facilities



(c) Target Staffing: Low-Medicaid Facilities

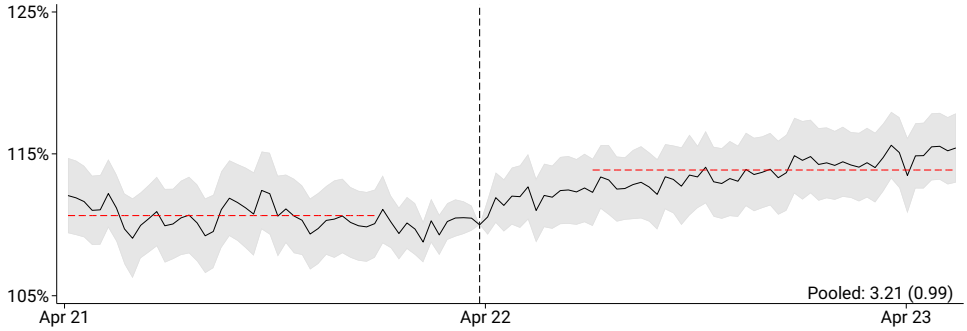


Figure F.2: Event study of staffing levels, by Medicaid payer share, balanced panel

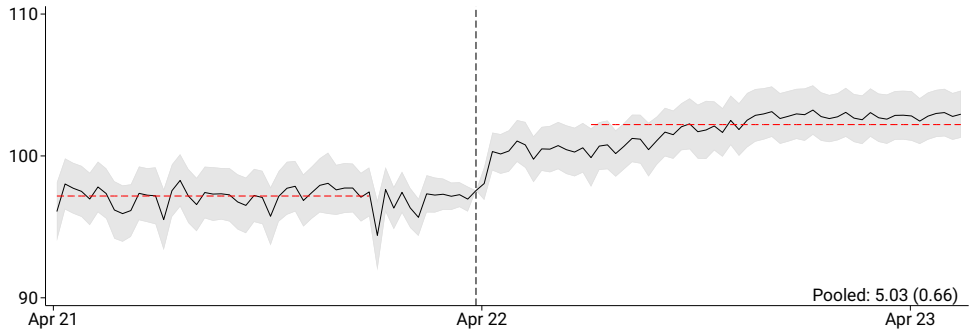
Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Gray area denotes 95% confidence interval with robust standard errors clustered by facility. The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

	Illinois (1)	Non-Illinois (2)	Matched Controls (3)	P-value (4)
Matching Variables				
STRIVE Ratio	1.050	1.227	1.069	0.054
RN STRIVE	0.231	0.208	0.204	< 0.001
LPN STRIVE	0.195	0.281	0.219	< 0.001
CNA STRIVE	0.624	0.739	0.646	< 0.001
Share of Hours by New Employees	0.132	0.125	0.127	0.002
NHC Overall Rating	2.89	3.04	2.80	0.141
Medicaid Share	0.535	0.620	0.566	0.002
For-profit	0.730	0.695	0.730	0.989
Government	0.039	0.066	0.039	0.977
Non-profit	0.231	0.231	0.230	0.977
Large Central Metro	0.271	0.206	0.264	0.700
Large Fringe Metro	0.252	0.187	0.253	0.947
Medium Metro	0.097	0.214	0.098	0.922
Small Metro	0.078	0.109	0.079	0.900
Micropolitan	0.165	0.130	0.166	0.925
Noncore	0.137	0.147	0.139	0.901
Alzheimer's Unit	0.168	0.136	0.146	0.159
Non-matching Variables				
Total Beds	123.5	107.5	106.6	< 0.001
Occupancy Rate	0.720	0.806	0.811	< 0.001
Age	79.9	78.5	79.4	0.095
Female	0.573	0.577	0.572	0.739
Black	0.171	0.159	0.151	0.061
N	619	12,006	3,024	

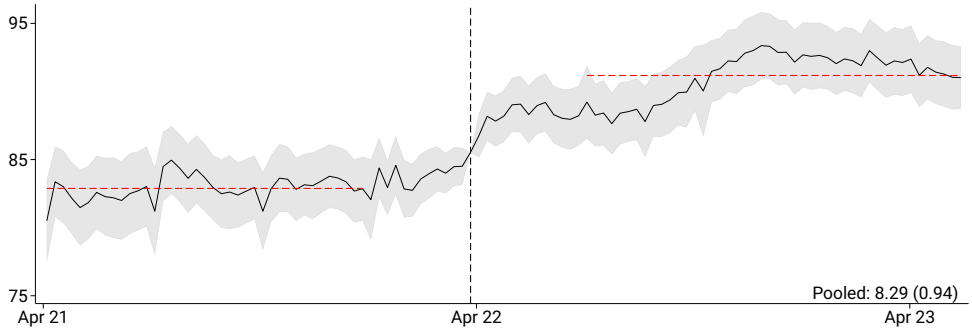
Table F.1: Matching Approach Summary Statistics

Notes: Table presents statistics from the Illinois sample in column (1), all non-Illinois facilities in column (2), and the matched control sample in column (3). P-values from comparison of means between columns (1) and (3) are presented in column (4).

(a) Target Staffing: All Facilities



(b) Target Staffing: High-Medicaid Facilities



(c) Target Staffing: Low-Medicaid Facilities

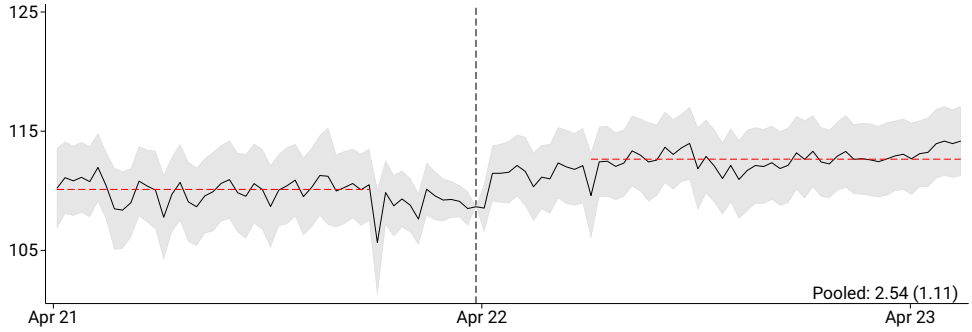


Figure F.3: Event study of staffing levels, by Medicaid payer share, matched control analysis

Notes: Figure presents results from difference-in-difference event study regressions with each outcome centered around the mean value in Illinois during the week before the effective date. Gray area denotes 95% confidence interval with robust standard errors clustered by facility. Sample limited to Illinois facilities and a matched control group for other states (Appendix F.2). The vertical line indicates the effective date of the reform: April 1, 2022. The red horizontal lines indicate pre-treatment and post-treatment averages, excluding the quarter prior and the quarter after the reform. The pooled estimate in the lower-right corner provides the difference between the post- and pre-treatment average coefficients. The standard error of this difference is reported in parentheses.

G Robustness to alternative inference approaches

In this section, we discuss our approach to statistical inference and describe the robustness of our findings to alternative approaches. While our main results are readily apparent in aggregate time series data (Appendix Figure G.1), one may nonetheless be concerned that our policy analysis relies on the comparison of outcomes for firms in one treated state against all other states. Accordingly, we conduct a battery of tests to assess the robustness of our findings, rather than rely on only one approach.

To assess the robustness of our results, we consider a slightly modified version of our main difference-in-differences regression (1). Specifically, we estimate the model:

$$y_{it} = \sum_{\tau \neq -1} \beta^{\tau}(IL_i \times q_t^{\tau}) + \alpha_i + \alpha_t + \varepsilon_{it} \tag{4}$$

in which the original calendar-week terms d_t^{τ}, meant to capture differential trends for facilities in Illinois relative to other states, are replaced with calendar-quarter terms q_t^{τ}. This substitution is made only for brevity in comparing standard errors; there is no meaningful change to the inference from this grouping. Notice that our calendar week fixed effects capturing aggregate trends α_t remain unchanged.

In our main specification, we cluster our standard errors at the facility level. Our reasoning for this is that there may be autocorrelation in the error term at the facility-level. Indeed, heteroskedastic-robust standard errors that do not account for this clustering are likely to be underestimated (Appendix Table G.1, panel A). Moreover, as our panel data are analyzed at the facility-week level, this is the natural unit of clustering. Reassuringly, we find no meaningful change in our inference when we shift to the quarterly effects model (Appendix Table G.1, panel B).

Because our treatment is defined at the state level, one may instead prefer to cluster at the level of treatment-assignment (i.e., the state level). Indeed, when we cluster at this level, we find that our standard errors decline slightly from the facility-level clustering approach (Appendix Table G.1 panel C). Cameron et al. (2008) point out that when conducting inference with small numbers of clusters (in our case, states), cluster-robust standard errors may be biased downwards, which can potentially explain why we find smaller standard errors when we cluster at the state rather than facility level. Accordingly, we follow and assess the sensitivity of our inference to non-parametric approaches. We implement a block bootstrapping procedure, in which we resample states with replacement. We conduct 2,000 bootstrap replications. This procedure generates standard errors that closely mirror those from the state-cluster robust approach (Appendix Table G.1 panel D). Next, we implement the wild-

cluster bootstrap-t procedure recommended by Cameron et al. (2008). We use Rademacher weights in our wild bootstrap procedure, using states as clusters. The resulting p-values from the bootstrapped distribution of t-statistics, in panel E, show no meaningful divergence from the prior approaches. Finally, we consider a fully non-parametric permutation test approach, in which for each of the other 49 untreated states, we assign a dummy treatment status, and re-estimate equation (4) using this permuted treatment variable. We plot each of the corresponding β^{τ} estimates in Appendix Figure G.2, and report the resulting p-values (calculated from Illinois's rank in the empirical distribution) in Appendix Table G.1 panel F. We find inference results that are consistent with each of the preceding models.

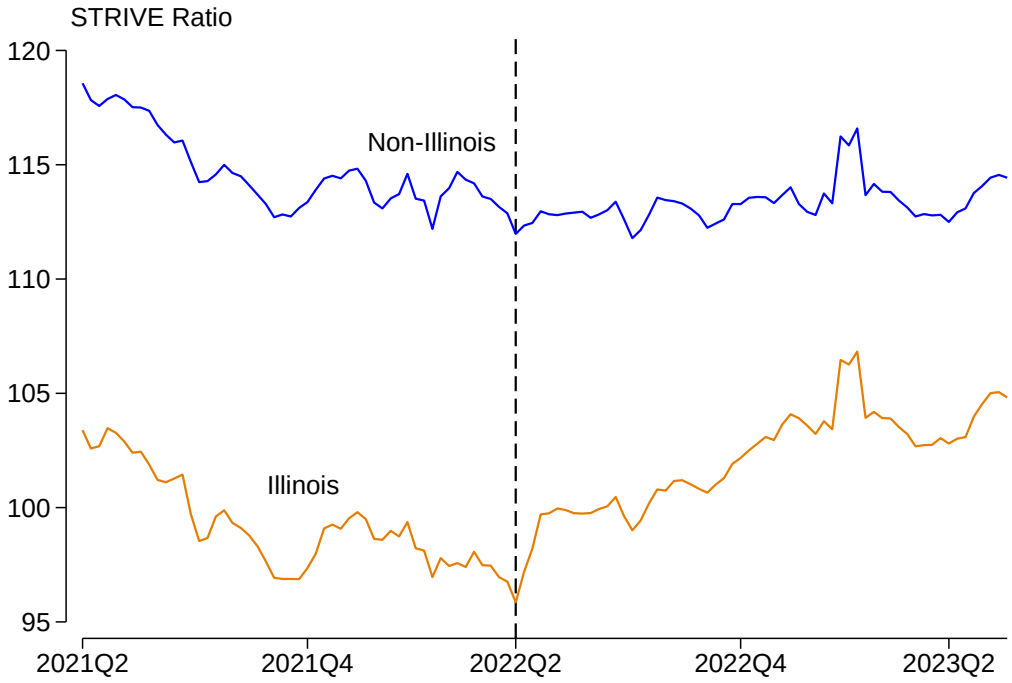


Figure G.1: Evolution of target staffing over time

Notes: Figure plots 3-week rolling average of target staffing in Illinois facilities compared to facilities in all other states.

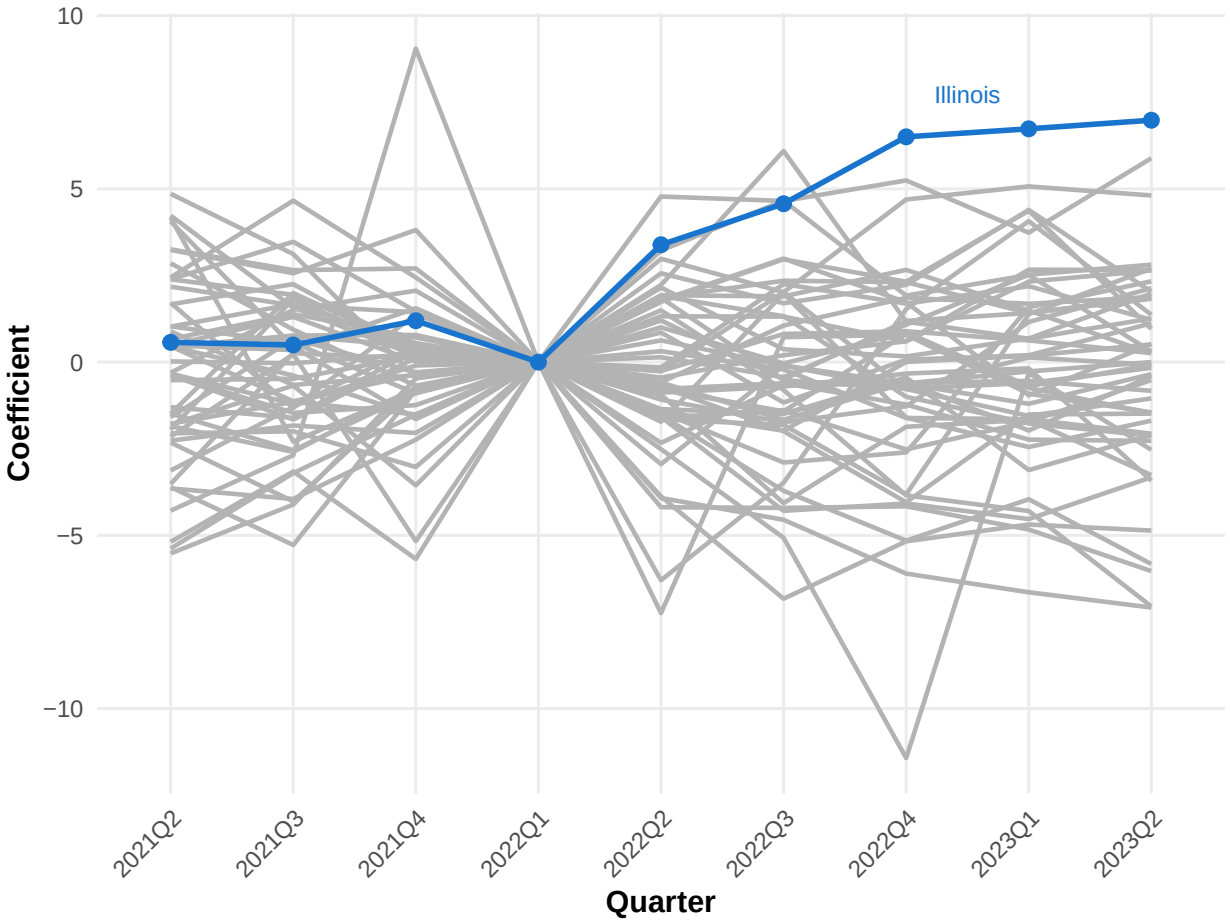


Figure G.2: Permutation test

Notes: Results from permutation test. Each state is assigned a treatment status, and we estimate the main event study regression (1) using this new treatment dummy for each state. Illinois is plotted in blue; all other states are in gray. Quarter effects replace d_t^τ in the treatment interaction term for visual clarity. Week fixed effects denoted by α_t persist.

	2021Q2	2021Q3	2021Q4	2022Q2	2022Q3	2022Q4	2023Q1	2023Q2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Heteroskedasticity-Robust Standard Errors								
Coefficient	0.434	0.202	0.623	2.654	4.015	5.941	6.486	6.885
Standard Error	(0.267)	(0.261)	(0.286)	(0.247)	(0.258)	(0.276)	(0.253)	(0.379)
P-Value	[0.104]	[0.438]	[0.029]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Panel B: Facility-Cluster-Robust Standard Errors								
Coefficient	0.434	0.202	0.623	2.654	4.015	5.941	6.486	6.885
Standard Error	(0.679)	(0.683)	(0.595)	(0.535)	(0.631)	(0.720)	(0.674)	(0.732)
P-Value	[0.523]	[0.767]	[0.295]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Panel C: State-Cluster-Robust Standard Errors								
Coefficient	0.434	0.202	0.623	2.654	4.015	5.941	6.486	6.885
Standard Error	(0.291)	(0.254)	(0.226)	(0.280)	(0.339)	(0.397)	(0.458)	(0.481)
P-Value	[0.142]	[0.430]	[0.008]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Panel D: Block-Bootstrap Standard Errors								
Coefficient	0.434	0.202	0.623	2.654	4.015	5.941	6.486	6.885
Standard Error	(0.315)	(0.287)	(0.298)	(0.341)	(0.361)	(0.403)	(0.507)	(0.500)
Panel E: Wild Cluster Bootstrap-t								
Coefficient	0.434	0.202	0.623	2.654	4.015	5.941	6.486	6.885
P-Value	[0.117]	[0.436]	[0.014]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Panel F: Permutation Test								
Coefficient	0.434	0.202	0.623	2.654	4.015	5.941	6.486	6.885
P-Value	[0.340]	[0.400]	[0.180]	[0.020]	[0.060]	[0.000]	[0.000]	[0.000]

Table G.1: Alternative inference approaches

Notes: Table provides alternative approaches to inference. For clarity, the model described in equation (1) is modified to contain quarter, rather than week, effects in the main treatment interaction term (the d_t^{τ} terms). Calendar week fixed effects (α_t) capturing aggregate trends remain the same. Panel A reports the results from heteroskedasticity-robust standard errors. Panel B corresponds to our primary analysis, and provides results from cluster-robust standard errors, where the clustering is at the facility level. Panel C presents results from state-level cluster-robust standard errors. Panel D presents the results from a block bootstrap procedure using states as blocks (Bertrand et al., 2004). Panel E presents results from a wild cluster bootstrap-t procedure (Cameron et al., 2008). Panel F presents the results from a permutation test in which each other state is assigned treatment status.