Systematic Misreporting in Nursing Home Quality Measures

Andrew R. Olenski, PhD[1, 2]

Ryan Zhao, BA[3]

Ashvin Gandhi, PhD[2, 3]

[1] Department of Health Policy and Management, Yale School of Public Health, New Haven, CT

[2] National Bureau of Economic Research, Cambridge, MA

[3] Anderson School of Management, University of California, Los Angeles, Los Angeles, California and National Bureau of Economic Research, Cambridge, MA

Word Count: 2654 words, 4 figures and 1 table

Date: July 31, 2026

Corresponding author to which reprints should be addressed:

Andrew R. Olenski, PhD
Department of Health Policy and Management
Yale School of Public Health
60 College St
New Haven, CT 06510
Email: andrew.olenski@yale.edu

Systematic Misreporting in Nursing Home Quality Measures

Abstract

Background Public quality reporting in health care often relies on providers to measure themselves on dimensions of care that outsiders cannot easily observe. This dependence on self-reported data creates scope for strategic misreporting. In January 2026, the Centers for Medicare & Medicaid Services began validating one self-reported nursing home quality measure — inappropriate antipsychotic prescribing for long-stay residents — against externally recorded Medicare claims, revealing the probable extent of prior misreporting across facilities.

Methods We analyzed federal data on nursing home quality measures and characteristics before and after the claims-based validation. For each facility, we calculated the change in the reported antipsychotic prescribing rate from the pre-validation period to the post-validation period, as a facility-level marker of probable misreporting. We examined the association of this change with other self-reported quality measures, independently measured quality measures, and facility characteristics. Using these estimates, we computed the implications of misreporting for consumer-facing quality-based star ratings.

Results Mean antipsychotic prescribing rates jumped substantially following claims-based validation. Large increases suggestive that a facility was previously underreporting antipsychotic use was associated with better scores on most self-reported measures—lower rates of infection and catheter use, fewer falls, and lower rates of reported decline in the activities of daily living—but worse scores on all independently-measured outcomes—more inspection citations, higher fines, lower staffing levels, and higher hospitalization and emergency department visit rates. This systematic divergence between independently measured and self-reported quality was not explained by geographic confounders and was concentrated among for-profit and chain-affiliated facilities. Computing star ratings adjusted for this behavior reveals two distortions: favorable misreporting directly inflated facilities' own ratings while harming other facilities by raising benchmarks.

Conclusions Misreporting of self-reported quality measures appears to be systematic: facilities revealed as likely to be inflating their self-reported performance on one quality measure appear to perform systematically better on virtually all self-reported quality measures but worse on all independently assessed quality measures. This raises concerns about the fidelity of quality measurement systems that rely on self-reporting.

1 Introduction

Many dimensions of healthcare quality are difficult for anyone except the provider to observe. Therefore, providers are often required to measure and report data about themselves that is used by regulators to construct quality ratings and even to determine payments. This is true for physicians, hospitals, nursing homes, home health agencies, and dialysis clinics, among others, and the usefulness of these measures depend crucially on the accuracy of providers' self-reporting.

Self-reported measures play a particularly large role in the nursing home industry, where key indicators of quality care cannot be discerned from administrative data such as claims alone. The Centers for Medicare & Medicaid Services (CMS) collects and publishes both self-reported quality measures and quality measures that are independently assessed from administrative data. CMS also aggregates these measures into a Five-Star Quality Rating System that patients, regulators, payers, and researchers rely on to distinguish better- from worse-performing facilities.[1,2]

Many of these nursing home quality measures are derived from the Minimum Dataset (MDS), a standardized clinical assessment tool completed by nursing home staff.[3,4] Facilities can therefore improve their quality ratings by coding MDS resident assessments more favorably. Prior work has documented underreporting of falls,[5] pressure ulcers,[6] and inappropriate antipsychotic prescribing[7] in the MDS relative to Medicare claims, as well as discrepancies between self-reported and independently measured staffing.[8]

These findings raise a broader question: how systematic is quality misreporting? Favorable coding on any isolated measure may have little effect on the composite ratings most consumers and regulators rely on. But if the same providers are misreporting across many measures, then the aggregate distortion could severely undermine the rating system. The problem compounds further if misreporting is negatively correlated with true quality: inaccuracies would then be concentrated among the lowest-performing providers, leaving the ratings least reliable precisely where poor care most needs to be identified.

To investigate this question, we studied a January 2026 policy change in which CMS began validating a prominent self-reported quality measure—the inappropriate use of antipsychotics to

sedate long-stay nursing home residents—against external Medicare and Medicaid claims.[9] Prior to the policy change, the measure was derived from self-reports alone: facility staff recorded in the MDS whether a resident was recently given antipsychotics and whether the resident had a legitimate clinical need, such as a diagnosis of schizophrenia. Facilities have been shown to both under-report usage of antipsychotics and over-report diagnoses indicating clinical need in order to appear more favorable on this measure.[7,10,11] The policy change addressed this misreporting by measuring both antipsychotic use and schizophrenia diagnoses using claims data rather than self-reports.

We investigate the extent to which new claims-based measures of inappropriate antipsychotic use exceed the immediately preceding self-reported measures. We interpret a facility having much higher rates under the claims-based measure than the preceding self-reported measure to be a marker of likely prior misreporting. We estimated the extent to which this marker of misreporting is associated with performance on other self-reported and independently-measured dimensions of quality. Finally, we assessed the implications of systematic misreporting of quality measures for the composite five-star rating for nursing home quality.

2 Methods

2.1 Data Sources and Study Population

The core of our analysis used historical releases of the CMS Care Compare files. From these releases we constructed a facility-level antipsychotic prescribing time series spanning 2021Q1 through 2026Q1. The study population consisted of all Medicare- or Medicaid-certified nursing homes with non-missing antipsychotic prescribing rates in both the last available pre-validation period and the first post-validation period, yielding 14,002 facilities.

We linked these data to a facility cross-section assembled from the 2026Q1 Care Compare files, which provided both self-reported and independent quality measures as well as facility characteristics. We supplemented with the most recent data from LTCFocus, providing

information on facility residents, and included local area economic and demographic variables from the Census Bureau's Small Area Income and Poverty Estimates.

2.2 Measuring Prior Misreporting

For each facility, we computed the change in antipsychotic prescribing (ΔAP) as the first post-validation antipsychotic prescribing rate (measure code 481) minus the last pre-validation rate (measure code 419). Large positive values for ΔAP indicate that a facility performed much better when antipsychotic prescribing was self-reported than when it was measured independently from claims. Such cases suggest that a facility was under-reporting antipsychotic use or over-reporting schizophrenia diagnoses to inflate their performance. We therefore interpret ΔAP as a marker of how heavily the facility had misreported this particular quality measure.

Facilities were assigned to quintiles of ΔAP for some comparisons, but all regression analyses used continuous ΔAP.

2.3 Outcomes

We examined two primary sets of outcomes. Self-reported quality outcomes included the six MDS measures that enter the long-stay quality-measure star rating, excluding the antipsychotic outcome: urinary tract infections (measure code 407), catheter use (406), falls with major injury (410), walking ability decline (451), activities of daily living (ADL) decline (401), and rates of pressure ulcers (479).

Independently-assessed quality measures included health deficiency citations and total fines from state inspections, registered nurse and total care hours per resident day from Payroll-Based Journal data, and claims-based hospitalization and emergency department visit rates per 1,000 residents.

Finally, we investigated how prior misreporting varied with facility and local-area characteristics included for-profit ownership, chain affiliation, certified beds, resident racial composition, payer mix, county poverty, population density, and median household income, to identify the characteristics of providers most engaged in misreporting.

2.4 Statistical Analysis

For each quality measure, we estimated ordinary least-squares regressions of the measure on continuous ΔAP. We considered both unadjusted models and adjusted models that included controls for ownership status, chain affiliation, number of beds, and state fixed effects. For ease of interpretation, estimates are reported as percentage points of each outcome's standard deviation associated with a 1-percentage-point increase in ΔAP. To ensure that larger and more positive coefficients consistently indicate better quality, coefficients were multiplied by -1 when higher values of the raw denote worse quality.

Supplementary analyses estimated the quality-measure regressions separately by ownership type and assessed whether ΔAP was correlated with the pre-validation antipsychotic prescribing rate.

All analyses were performed using R, version 4.4. The study used publicly available facility-level data and contained no individual patient information. The 95% confidence intervals around reported estimates reflect 0.025 in each tail.

2.5 Star-Rating Adjustment

To gauge how favorable coding might affect the composite long-stay quality-measure Five-Star Rating, we computed each facility's rating under two scenarios and compared them. The baseline rating used facilities' self-reported MDS values scored with real benchmarks recovered from historical data. The adjusted rating removed from each of the six MDS-derived measures the average amount of favorable coding implied by each facility's ΔAP, as captured by the linear relationship with ΔAP estimated separately for each measure. We then re-computed the benchmarks using adjusted data, and re-scored facilities accordingly.

Correcting for favorable coding alters ratings through two channels. First, it deflates the ratings of facilities that misreported (i.e., for which $\Delta AP > 0$). Second, by deflating the ratings of misreporting facilities, it improves the relative (i.e., percentile) performance of facilities that reported honestly. Our adjusted rating incorporates both. Appendix A.3 provides additional details, and Appendix C examines the robustness of the exercise to measurement error in ΔAP.

3 Results

3.1 Changes in Antipsychotic Prescribing

Mean antipsychotic prescribing rates were stable at approximately 14.2%–14.7% from 2021 through December 2025 (Figure 1, panel a). After claims validation began in January 2026, the mean rate rose to 16.7%. The mean within-facility change (ΔAP) was 2.2 percentage points (pp), with an interquartile range of −0.1 to 4.1 pp and a 90th percentile of 6.8 pp (Figure 1, panel b). Reported rates declined for 3,582 facilities (25.6%) consistent with the absence of prior misreporting at these facilities. A placebo analysis comparing changes from 2025Q3 to 2025Q4 shows that declines of this magnitude reflect ordinary period-to-period variation unrelated to the policy (Appendix C.1).

Facilities that increased most differed systematically from those that increased least (Table 1). Mean ΔAP ranged from −2.9 pp in the lowest quintile to +8.0 pp in the highest, and the post-validation prescribing rate in the highest quintile (25.0%) was nearly double that in the lowest (13.2%). Relative to the lowest quintile, facilities in the highest quintile were more likely to be for-profit (81.6% vs. 63.8%), chain-affiliated (72.9% vs. 62.5%), and larger (101 vs. 92 certified beds; all $P < 0.001$).

3.2 Self-Reported vs. Independent Quality Measures

The association between ΔAP and measured quality differed according to the source of the measure. On self-reported MDS measures, facilities with larger increases in antipsychotic prescribing reported better quality. As compared with facilities in the lowest quintile of ΔAP, those in the highest quintile reported lower rates of urinary tract infection (1.58% vs. 2.24%), catheter use (0.99% vs. 1.28%), falls with major injury (3.12% vs. 3.52%), decline in walking ability (15.2% vs. 18.5%), and decline in activities of daily living (14.8% vs. 16.2%) (P<0.001 for all comparisons; Table 1). Only the rate of pressure ulcers did not differ significantly between the highest and lowest quintiles (5.00% vs. 5.04%; P=0.70).

In contrast, facilities with larger ΔAP performed worse on every independently measured outcome. As compared with facilities in the lowest quintile, those in the highest quintile had more health deficiency citations (10.8 vs. 8.2), higher fines ($39,652 vs. $25,829), lower registered-nurse staffing (0.61 vs. 0.77 hours per resident-day), lower total direct-care staffing (3.85 vs. 4.15 hours per resident-day), a lower registered-nurse share of licensed hours (40.3% vs. 46.8%), and higher rates of hospitalization (2.06 vs. 1.76 per 1,000 resident-days) and emergency department visits (2.05 vs. 1.76 per 1,000 resident-days) ($P<0.001$ for all comparisons; Table 1).

In regressions of each measure on continuous ΔAP, we found that five of the six self-reported MDS measures had a positive association with misreporting (Figure 2 panel a). Conversely, all seven independently measured outcomes reflected negative associations (Figure 2 panel b). These patterns held both with and without adjustment for state fixed effects, for-profit ownership, chain affiliation, and certified beds.

Binned scatter plots of the quality measures against ΔAP showed approximately monotonic relationships across the distribution of ΔAP (Supplementary Figure S1). Regressions estimated separately by ownership type, reveals similar patterns within both for-profit and non-profit/public facilities (Supplementary Figure S2).

3.3 Facility Characteristics and Geography

Figure 3 panel (a) shows standardized associations between ΔAP and facility and local-area characteristics for consistency across outcomes. For interpretability, here we report associations per 1-SD increase in ΔAP (4.2 pp) here. A 1-SD increase in ΔAP was associated with a higher probability of for-profit ownership (6.02 pp; 95% CI, 5.32 to 6.73) and chain affiliation (3.17 pp; 95% CI, 2.36 to 3.99), and with larger facility size (1.88 beds; 95% CI, 1.04 to 2.73). It was also associated with higher shares of Black residents (3.75 pp; 95% CI, 3.16 to 4.33) and Medicaid-enrolled residents (3.85 pp; 95% CI, 3.24 to 4.46), and with lower shares of White residents (−4.08 pp; 95% CI, −4.52 to −3.64) and Medicare residents (−0.59 pp; 95% CI, −0.91 to −0.27). Associations with local-area characteristics were weaker: a 1-SD increase in ΔAP was associated

with a county poverty rate that was 0.35 pp higher (95% CI, 0.28 to 0.42), population density that was 83.2 persons per square mile higher (95% CI, 41.4 to 125.1), and median household income that was $411 lower (95% CI, −718 to −103).

Increases in reported antipsychotic prescribing varied across geographies but were observed throughout the United States (Figure 3 panel b). Above-median increases were observed in several regions, including parts of southern California, Louisiana, and south Texas.

3.4 Quality-Measure Star Ratings

We next assessed whether this favorable coding affected the composite long-stay quality-measure Five-Star Rating, comparing each facility's baseline five-star quality rating with the adjusted rating that corrects the six MDS-derived measures for the expected favorable coding implied by ΔAP (Appendix A.3).

Our simulation of correcting misreporting suggests that this would not change the five-star quality rating for most facilities (94.1%). However, the changes that would occur are strongly graded by ΔAP (Figure 4). Facilities with the largest ΔAP were the most likely to be downgraded: in the highest ΔAP quintile, 14.2% received a lower rating (mean change, −0.15 stars), reflecting the removal of favorable coding from their own measures (Figure 4, panel b). In contrast, 5.9% of facilities with no indication of misreporting ($\Delta AP \leq 0$) saw their quality five-star rating increase. While the raw scores for these facilities remained unchanged, correcting misreporting improved their percentile ranking used by the five-star system. These findings highlight two channels through which misreporting affects ratings: by directly inflating the ratings of facilities that misreported and by indirectly depressing the ratings of facilities that reported accurately.

These findings are not an artifact of ordinary measurement error in ΔAP. In a placebo window spanning no policy change, period-to-period variation did not reproduce the right-tail increases that drive the adjustment. We show this and other robustness in Appendix C.

4 Discussion

Many determinants of healthcare quality are difficult for anyone but the provider to observe. Accordingly, providers are often asked to self-report on a number of dimensions. These self-reported measures then underly quality scores that patients, regulators, and payers rely upon. This system can work well if providers report faithfully but is particularly concerning if misreporting is substantial and systematic.

The January 2026 claims-validation of the antipsychotic prescribing measure allows us to assess the extent of misreporting in nursing home quality measures. Many facilities' independently-assessed performance was substantially worse than their self-reported performance. These same facilities also performed better on most self-reported quality measures but worse on every independently-assessed measure. This pattern is difficult to rationalize except by these facilities engaging in systematically favorable misreporting while actually providing lower-quality care. The positive associations between post-validation increases and for-profit ownership and chain-affiliation are consistent with the large prior evidence linking these characteristics to reduced nursing home quality.[12–14]

While others have documented misreporting,[7,10,11] to our knowledge, ours is the first study to show systematic correlation in misreporting and assess its implications for the quality rating system. Our findings have important implications for public quality reporting, where composite ratings are especially vulnerable to misreporting that is systematically correlated across measures. A provider that misreports one measure favorably modestly improves its profile; a provider that misreports many measures can increase its star rating and make the rating system meaningfully less reliable.

In the nursing home setting, 14.2% of facilities in the highest quintile of misreporting had higher star ratings due to misreporting. This inflation in ratings is particularly harmful because these facilities tend to be lower-quality based on independently-assessed quality measures. Faithfully reporting facilities are also harmed: 5.9% of these facilities received lower star ratings because their percentile was deflated due to others' misreporting.

Our study has several limitations. First, our primary measure is only a proxy for misreporting, not a direct resident-level measure of unsupported diagnoses or prescribing. Some of the change may reflect resident turnover, legitimate coding corrections, or other contemporaneous changes. However, the sharp timing of the increase and the stability of rates before validation support misreporting being the dominant driver. Second, the analyses are cross-sectional, and we cannot fully exclude unobserved facility characteristics driving the observed quality patterns. However, the consistent positive correlation with self-reported measures, consistent negative correlation with independently-assessed measures, and the robustness of all findings to adjustment for state and facility characteristics make alternative interpretations less plausible. Third, we observe facility-level measures rather than individual assessment records, and so cannot identify the specific coding decisions behind the facility-level patterns.

Our study raises the question of whether quality ratings should rely on self-reporting when providers have both the discretion and the incentive to code in their own favor. The transition we study shows that sometimes self-reported measures can be directly replaced with independently-assessed measures. Where claims-based alternatives don't exist, regulators might look to measures reported by third-parties such as patients or inspectors. Alternatively, regulators could regularly audit self-reported measures or place lower weight on them in composite ratings, especially when a provider perform much better on self-reported measures than independently-assessed ones.

References

[1] Rachel M Werner, R Tamara Konetzka, and Daniel Polsky. Changes in consumer demand following public reporting of summary quality ratings: an evaluation in nursing homes. *Health services research*, 51:1291–1309, 2016.

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

[3] Centers for Medicare & Medicaid Services. *Design for Care Compare Nursing Home Five-Star Quality Rating System: Technical Users' Guide*, April 2026. URL `https://www.cms.gov/medicare/provider-enrollment-and-certification/certificationandcomplianc/downloads/usersguide.pdf`.

[4] Debra Saliba and Joan Buchanan. Development and validation of a revised nursing home assessment tool: Mds 3.0. Technical report, RAND Health, 2008. URL https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/nursinghomequalityinits/downloads/mds30finalreport.pdf.

[5] Prachi Sanghavi, Shengyuan Pan, and Daryl Caudry. Assessment of nursing home reporting of major injury falls for quality measurement on nursing home compare. *Health services research*, 55(2):201–210, 2020.

[6] Prachi Sanghavi and Zihan Chen. Underreporting of quality measures and associated facility characteristics and racial disparities in us nursing home ratings. *JAMA network open*, 6(5):e2314822, 2023.

[7] Becky A Briesacher, Brianne Mui, John W Devlin, and Benjamin Koethe. Nursing homes underreport antipsychotic prescribing. *Aging & mental health*, 24(4):668–672, 2020.

[8] Fangli Geng, David G Stevenson, and David C Grabowski. Daily nursing home staffing levels highly variable, often below cms expectations. *Health Affairs*, 38(7):1095–1100, 2019.

[9] Centers for Medicare & Medicaid Services. QSO-25-20-NH REVISED: Updates to nursing home care compare, 2025. URL https://www.cms.gov/files/document/qso-25-20-nh-revised-2025-09-10.pdf. Memorandum to State Survey Agency Directors; revised September 10, 2025.

[10] Amanda C Chen and David C Grabowski. Nursing homes underreport antipsychotic use but overreport diagnoses qualifying for appropriate use. *Health affairs scholar*, 1(2):qxad022, 2023.

[11] Office of Inspector General, U.S. Department of Health and Human Services. Nursing homes inappropriately diagnosed residents with schizophrenia to mask the misuse of antipsychotic drugs. Technical report, U.S. Department of Health and Human Services, 2026. URL https://oig.hhs.gov/reports/all/2026/nursing-homes-inappropriately-diagnosed-residents-with-schizophrenia-to-mask-the-misuse-of-antipsychotic-drugs/.

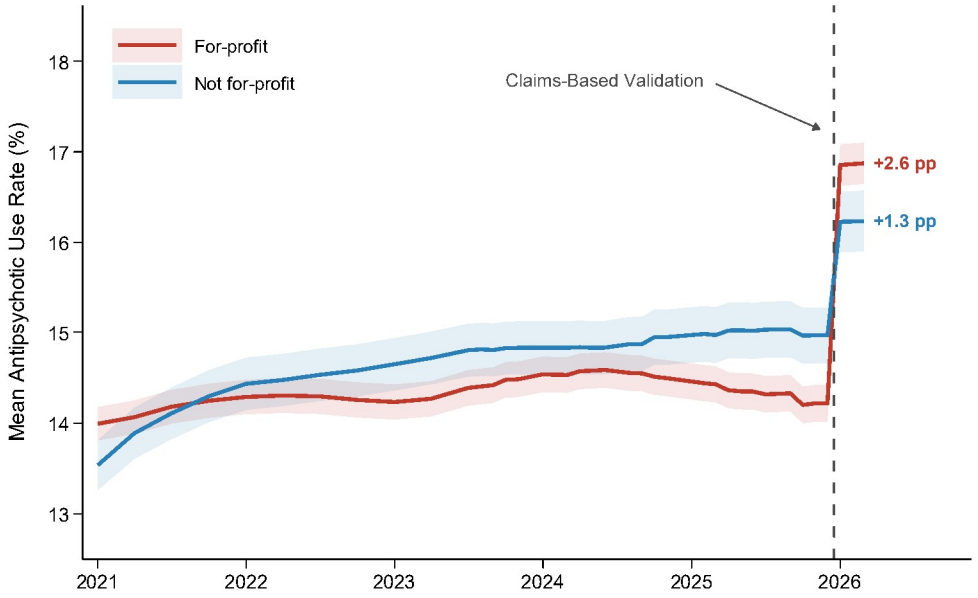
[12] Vikram R Comondore, Philip James Devereaux, Qi Zhou, Samuel B Stone, Jason W Busse, Nikila C Ravindran, Karen E Burns, Ted Haines, Bernadette Stringer, Deborah J Cook, et al. Quality of care in for-profit and not-for-profit nursing homes: systematic review and meta-analysis. *Bmj*, 339, 2009.

[13] David C Grabowski and David G Stevenson. Ownership conversions and nursing home performance. *Health services research*, 43(4):1184–1203, 2008.

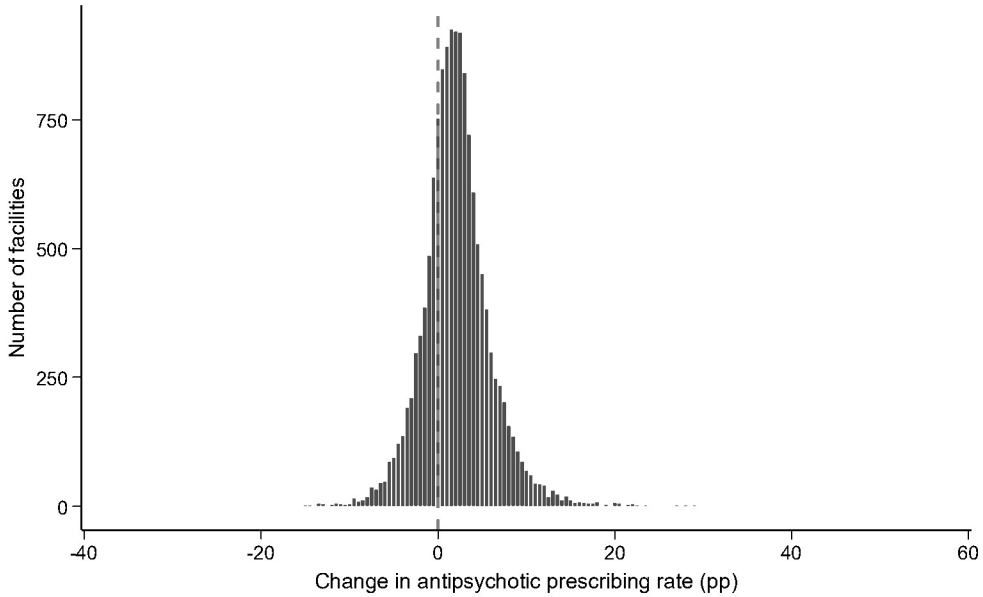
[14] Liran Einav, Amy Finkelstein, and Neale Mahoney. Producing health: measuring value added of nursing homes. *Econometrica*, 93(4):1225–1264, 2025.

Figure 1: Antipsychotic prescribing rates before and after claims-based validation

(a) Mean antipsychotic prescribing rates over time, by ownership type

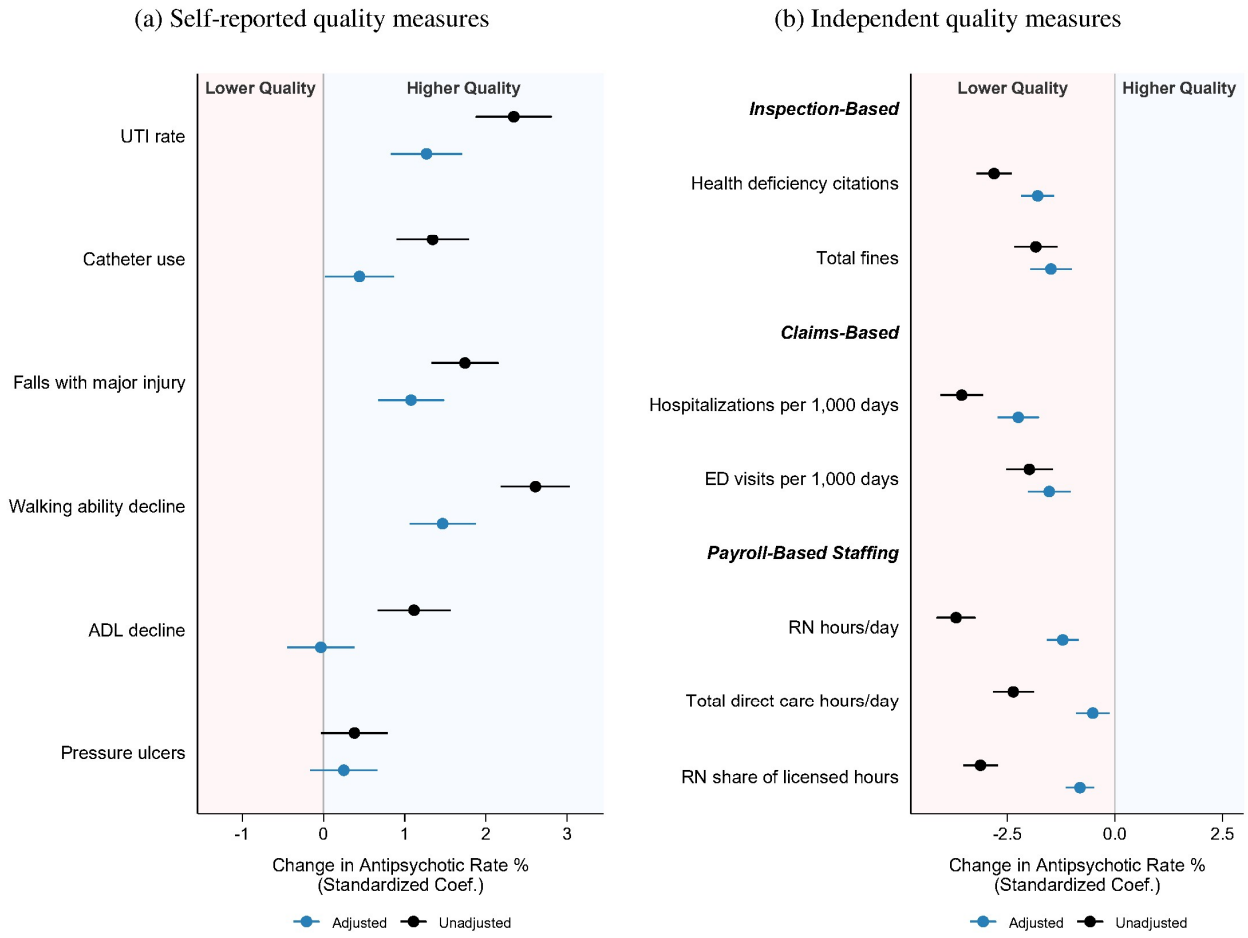


(b) Distribution of within-facility changes in antipsychotic prescribing rate



Notes: Panel (a) shows mean antipsychotic prescribing rates by quarter from 2021Q1 through 2023Q2 and by month from July 2023 through March 2026, separately for for-profit and non-profit or government-owned facilities. The vertical dashed line marks the January 2026 respecification, when CMS began validating schizophrenia exclusion diagnoses against claims. Shaded bands denote 95% confidence intervals. Panel (b) shows ΔAP, the within-facility change from the last pre-validation antipsychotic prescribing rate to the first post-validation rate. The vertical dashed line denotes zero change. $N = 14{,}002$ facilities.

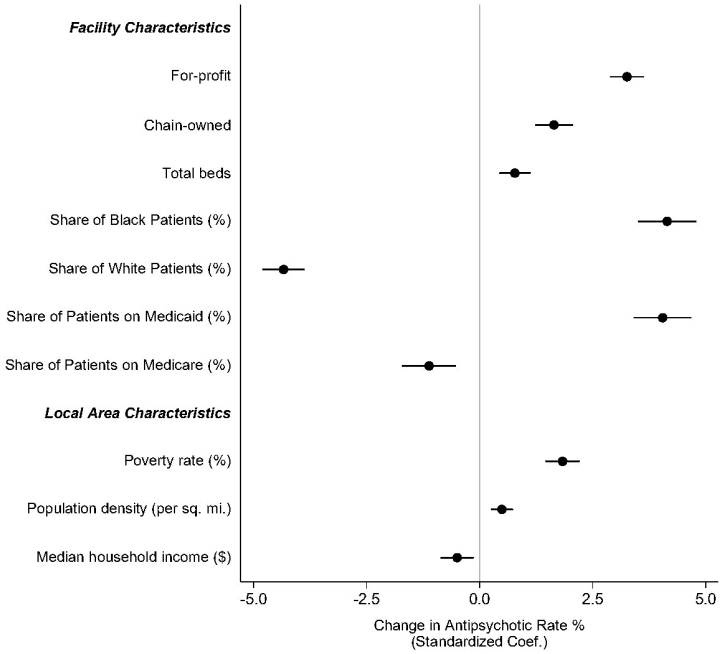
Figure 2: Association between changes in antipsychotic prescribing rate (ΔAP) and quality measures



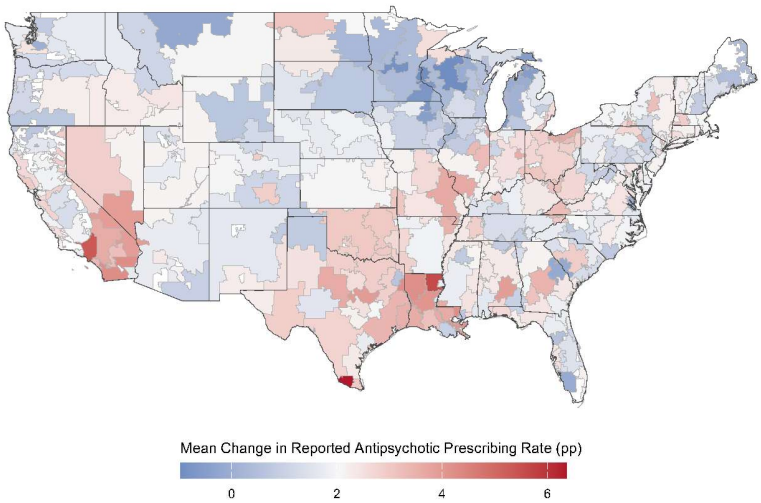
Notes: Each point is the standardized coefficient from a regression of the quality measure on within-facility ΔAP. Unadjusted estimates are bivariate. Adjusted estimates control for state fixed effects, for-profit ownership, chain affiliation, and certified beds. Coefficients are divided by the standard deviation of the dependent variable and multiplied by 100; values therefore represent percentage points of an outcome standard deviation associated with a 1-percentage-point increase in ΔAP. Coefficients are oriented so that positive values indicate better quality. Horizontal lines denote 95% confidence intervals. Independent measures are from inspections, Medicare claims, or payroll data; self-reported measures are from the MDS. $N = 14{,}002$ facilities.

Figure 3: Facility characteristics and geographic variation in antipsychotic prescribing rate changes

(a) Facility and local-area characteristics



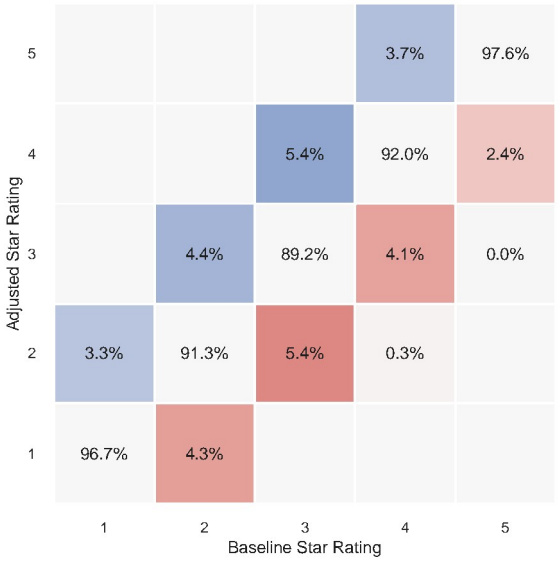
(b) Mean ΔAP across Hospital Referral Regions



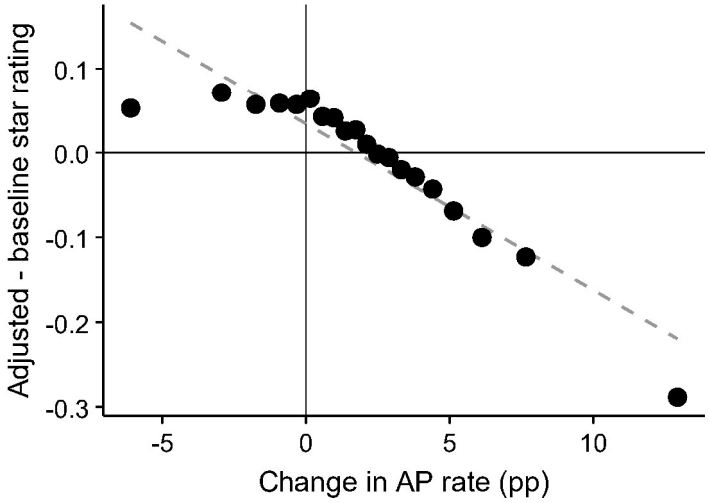
Notes: Panel (a) shows bivariate standardized coefficients from regressions of each facility or local-area characteristic on ΔAP. No covariates or fixed effects are included in this descriptive panel; horizontal lines denote 95% confidence intervals. Panel (b) displays mean ΔAP by Hospital Referral Region (HRR) for the continental United States. Facilities are assigned to HRRs using ZIP codes and the Dartmouth Atlas crosswalk. The map color scale is centered at the median HRR-level ΔAP; red regions had above-median increases and blue regions had below-median increases.

Figure 4: Adjusted quality-measure star ratings under historically deflated cutpoints

(a) Transition matrix for adjusted star ratings



(b) Change in adjusted star ratings by ΔAP



Notes: Panel (a) compares the reconstructed official quality-measure star rating with the adjusted rating. The official baseline is computed from observed current values under recovered, unadjusted historical cutpoints. The adjusted rating is computed after deflating historical MDS values, recomputing the cutpoints, deflating current MDS values, and applying the deflated cutpoints. Columns sum to 100%. Panel (b) plots 20 equal-sized bins of the facility-level change in star rating (adjusted minus official) against ΔAP. Negative values indicate that the adjusted rating is lower than the official baseline.

Table 1: Facility characteristics and quality measures by quintile of antipsychotic prescribing rate change

	Q1 (n=2801)	Q2 (n=2800)	Q3 (n=2800)	Q4 (n=2800)	Q5 (n=2801)	*P*
Mean AP rate change, pp	-2.9	0.3	1.9	3.6	8.0	<.001
Post-validation AP rate, %	13.2	13.2	14.7	17.0	25.0	<.001
Facility characteristics						
For-profit, %	63.8	71.4	75.1	78.3	81.6	<.001
Chain-affiliated, %	62.5	69.8	70.2	74.7	72.9	<.001
Certified beds	92.4	112.6	119.2	114.0	101.2	<.001
Independently measured outcomes						
Health deficiency citations	8.2	9.1	9.5	10.1	10.8	<.001
Total fines, $	25829	29198	30991	35721	39652	<.001
RN hours/resident/day	0.77	0.67	0.64	0.60	0.61	<.001
Total direct care hours/res/day	4.15	3.93	3.87	3.80	3.85	<.001
RN share of licensed hours, %	46.8	42.9	42.1	40.9	40.3	<.001
Hospitalizations per 1,000 days	1.76	1.79	1.87	1.93	2.06	<.001
ED visits per 1,000 days	1.76	1.66	1.72	1.84	2.05	<.001
Self-reported MDS measures						
UTI rate, %	2.24	1.68	1.55	1.56	1.58	<.001
Catheter use, %	1.28	0.90	0.89	0.87	0.99	<.001
Falls with major injury, %	3.52	3.31	3.26	3.17	3.12	<.001
Walking ability decline, %	18.51	16.57	16.08	15.74	15.25	<.001
ADL decline, %	16.15	14.78	14.57	14.48	14.76	<.001
Pressure ulcers, %	5.04	4.98	5.12	5.12	5.00	0.696

Notes: Facilities are assigned to quintiles based on ΔAP, the within-facility change in antipsychotic prescribing rate from the last pre-validation period (measure code 419) to the first post-validation period (measure code 481). Independent outcomes are derived from state inspections, Payroll-Based Journal staffing data, and Medicare claims. Self-reported measures are derived from the Minimum Data Set. *P* values are from two-sided *t*-tests comparing quintile 5 with quintile 1. RN = registered nurse; HPRD = hours per resident per day; ED = emergency department; UTI = urinary tract infection; ADL = activities of daily living.