Preliminary Working Draft
Please do not quote or cite without author's permission

Public reporting and consumer demand in the home health sector

Jun Li

Department of Public Administration and International Affairs
The Maxwell School of Citizenship and Public Affairs, Syracuse, NY

314 Lyman Hall, Room 320E
Syracuse, New York 13244
Phone: 315-443-8009
Email: jli208@syr.edu

Funding: This work was supported by the Agency for Healthcare Research and Quality [1R36HS026836]. I have no conflicts of interest to declare.

Abstract:

Health care report cards aim to address imperfect information and increase quality competition. I exploit a natural experiment in the home health sector to assess whether a higher rating under the federal star ratings program affects patient choice at the margin. Agencies with one more half star increased their market share by 1.4% or 0.25 (95% CI: −0.63 to 1.12) percentage points, an economically and statistically insignificant amount. I find no evidence of heterogeneous effects across the rating distribution or over time that would indicate a meaningful consumer response. I also find null effects among consumers expected to be more responsive, such as community-entry patients with more time to search for information and patients in competitive markets with more care options and star types. Suggestive evidence indicates that patient selection by agencies may have modestly impeded consumer choice. The star ratings are unlikely to improve home health quality.

Keywords: Information asymmetry, long-term care, post-acute care, quality, report cards

JEL: I10, I11, I12, I13, I18, I38

I. Introduction

"Public reporting is a key driver for improving health care quality by supporting consumer choice and incentivizing provider quality improvement," is the view that underpinned the 2010 Patient Protection and Affordable Care Act's massive expansion of public report cards in the US health care system (2015b; Reineck & Kahn, 2013). A prominent example is the Home Health Star Ratings program. Beginning in July 2015, CMS assigned each home health agency a 5-star rating based on its performance on nine quality measures relative to other agencies in the nation (Centers for Medicare & Medicaid Services, 2015a). The theoretical mechanisms that the Home Health Star ratings and other public reporting programs use to achieve better quality are simple: they rely on reallocation of consumer demand based on ratings (Bundorf, Chun, Goda, & Kessler, 2009). In one mechanism, if patients (or their agents) can identify and choose high-quality agencies, then quality improves through reallocation from lower- to higher-quality agencies. In another mechanism, if patients (or their agents) choose higher-quality agencies and agencies believe that patient demand for quality is elastic, then agencies will be incentivized to compete on quality, leading to better quality overall.

Despite its theoretical appeal, the empirical literature on public reporting in health care suggests mixed and, at best, modest success, with varied effects (Kolstad & Chernew, 2009). One reason for these lackluster results has been attributed to information complexity, which may prohibit consumers from understanding and using the information, a problem that is especially challenging for older people with physical and cognitive needs (Brook et al., 2002; Hibbard & Peters, 2003; Peters, Dieckmann, Dixon, Hibbard, & Mertz, 2007). A small body of literature suggests that the simpler, star ratings format for post-acute skilled nursing facilities and Medicare Advantage plans, may have had greater success in eliciting demand responses (Darden & McCarthy, 2015; Perraillon, Konetzka, He, & Werner, 2017; Werner, Konetzka, & Polsky, 2016). Easy-to-use formats, however, are unlikely to be a panacea when information is not actionable or is irrelevant, for example, for consumers in markets with limited provider choice (Baier et al., 2015; Grabowski & Town, 2011; Konetzka & Perraillon, 2016). Moreover, providers may actively engage in patient selection behaviors that could further erode patient choice (Dranove, Kessler, McClellan, & Satterthwaite, 2003; Ramanarayanan & Snyder, 2012).

This study asks whether providing quality information in the form of summary star ratings affects patients' choice of home health care at the margin. There are several unique features of the home health sector that makes it interesting to study. First, unlike other types of health care, there are no out-of-pocket fees or travel costs associated with Medicare home health services because patients receive care at home. Thus, Medicare home health patients have no competing cost concerns, making them the ideal subjects for studying how quality information influences consumer choice (Jung, Wu, Kim, & Polsky, 2016). Second, home health serves post-acute care patients with short-term rehabilitative needs as well as patients entering from the community with longer-term chronic care needs, presenting an opportunity to explore the effects of report cards across a diverse set of patients (Murtaugh, 2008). Finally, home health agencies and skilled nursing facilities are increasingly serving comparable populations (Werner, Coe, Qi, & Konetzka, 2019), and patients make provider choices with agents like hospitals or physicians in both contexts (Baier et al., 2015; Gadbois, Tyler, & Mor, 2017). Consequently, this study offers an indirect comparison with existing studies on star ratings for skilled nursing facilities, a setting for which star ratings appear promising.

To my knowledge, just one paper has examined consumer demand following the introduction of the Home Health Star Ratings program. Schwartz et. al (2022) used a conditional logit discrete choice model, comparing home health admissions among agencies with low, medium, and high star ratings, before and after the program was implemented. The authors estimated that higher-rated agencies gained 0.88 percentage points in market share. However, due to the pre-post design and lack of a control group, the study did not differentiate the effects of *reported* quality from other confounding factors, such as true quality differences between agencies or consumer learning. Moreover, they did not identify the marginal effects of one more half star, which is relevant information for policymakers and home health agencies assessing how to allocate resources for quality improvement.

Using national data from 2014 through 2016 with a regression discontinuity design, this study compares agencies that were virtually identical but barely on the opposite sides of an arbitrary star threshold, to identify the causal effect of an additional half star on new patient market shares. I focus on Medicare fee-for-service patients, for whom there are no provider network restrictions or cost-sharing differences across providers. This approach enables me to separate the effect of the star ratings from other sources, such as prior quality beliefs or changing quality of providers over time.

I find no evidence that higher ratings increased an agency's Medicare home health market share. One more half-star increased agencies' new patient market share by 0.25 (SE=0.45, 95% CI: −0.63 to 1.12) percentage points. The average agency accepts 18% of new patients in its market, thus even an effect of 1.12 percentage points at the upper end of the 95% confidence interval would be economically negligible. Although noisy, the results point to the same conclusion when examining across the star rating distribution. These estimates were not attenuated by low public awareness of the ratings at the beginning of the program or delays in consumer response to each release. There was also no evidence that patients who were better positioned to search for information (community-entry patients) or had more to gain from searching (patients in competitive markets with more home health options or star types) did so. These results were robust to both parametric and local randomization approaches.

I find suggestive evidence that patient selection by agencies may have modestly hindered consumers' use of higher-rated agencies. I find that one more half star increased the average market-level income of patients served by 0.6% to 0.8%. Agencies that gained one more half star increased the average income of their ZIP codes served by approximately \$257 (95% CI: −\$128, \$642) and \$314 (95% CI: \$3.31, \$625) with covariate adjustment. In labor-scarce areas, patterns suggest that agencies gaining one more half star in the middle of the distribution may have admitted fewer low-profit margin and dually eligible patients, diverting them to lower-rated agencies. These patterns in areas with low home health labor supply would be consistent with the idea that agencies cannot readily increase patient caseloads in the short run, so they capitalize on increased demand by being more selective. If these suggestive findings are real, it would indicate that the star ratings influenced consumer demand, but only slightly, and in a way that are unintended to the goals of the program.

This study contributes to the literature on interventions to resolve information asymmetry in health care. In contrast to star ratings for skilled nursing facilities (Perraillon et al., 2017; Werner et al., 2016) and managed care plans (Darden & McCarthy, 2015), it finds that after accounting for endogeneity, quality information vis-à-vis summary star ratings do not influence home health patient choice at the margin in a meaningful way. Despite an extensive investigation into potential heterogeneous effects of the star ratings by context, including those that are expected to influence patient decision-making and supply-side behaviors in various ways, this study was unable to find evidence that quality reporting in home health is likely to achieve its purported impact. The findings from this study provide useful insights into how policymakers should proceed with addressing imperfect information in health care.

Conceptual framework

The standard utility maximizing model specifies that consumers (patients and their agents) determine their best option (specific provider) using available information and their provider preferences. For Medicare fee-for-service patients there are no travel or out-of-pocket costs, and Medicare sets prices that are uniform across agencies. Thus, choosing a home health agency depends on its perceived quality and availability.

Quality varies among agencies in this market, but consumers have incomplete information at decision time. As a result, consumers may mistakenly believe that providers are of equal quality. One reason for incomplete information could be that consumers are unable to take advantage of available information. In situations where consumers must make decisions quickly, such as choosing post-acute care while being discharged from an inpatient stay, additional pressures may prevent them from searching for or synthesizing available information (Jung et al., 2016; Werner, Norton, Konetzka, & Polsky, 2012). Therefore, because of imperfect information, home health agencies are not incentivized to compete on quality due to relatively inelastic demand.

The Medicare 5-star ratings reduce cognitive barriers associated with information gathering and interpretation so patients can more easily distinguish high- from low-quality agencies (Centers for Medicare & Medicaid Services, 2015b). All else equal, if the 5-star ratings make comparing home health agencies easier, demand for higher-rated agencies should increase.

II. Institutional Background

The largest payer of home health services is Medicare (Centers for Medicare & Medicaid Services, 2021). In 2019, Medicare paid over 11,000 home health agencies $17.8 billion in 2019 for care delivered to 3.3 million US fee-for-service beneficiaries (Medicare Payment Advisory

Commission, 2021). For people eligible, the Medicare home health benefit covers patients with no out-of-pocket costs. Although formally considered a post-acute care service, there are two types of patients that receive home health care, those who need short-term, rehabilitative care (post-acute) and chronically ill people who need longer-term support (community admissions) (Murtaugh, 2008). All services are delivered to patients at their homes, and can include skilled nursing, physical therapy, speech-language pathology, occupational therapy, medical social services, home health aide services, and medical supplies. Patients are unconstrained in the number of episodes of care they can receive (Medicare Payment Advisory Commission, 2020).

Home health patients are a vulnerable group. Eligible patients must be unable to leave home or risk worsening their health by leaving. Home health patients are twice as likely to be at least 85 years old, 36% more likely to have at least three chronic conditions, 80% more likely to report fair or poor health, and 30% more likely to have incomes at or below 200% of the federal poverty level, when compared to the general Medicare population (Avalere, 2015).

To begin home health care, a physician or other qualified practitioner must first certify a patient's need and eligibility for services. Medicare fee-for-service patients are permitted to choose any Medicare-certified home health agency willing to render care under federal law. Thus, patients' choice sets are locally determined – constrained by agencies that serve their residential areas. Often, there are many agencies to choose from. More than 80 percent of patients live in ZIP codes with at least 5 agencies serving it (Medicare Payment Advisory Commission, 2021).

Patients frequently select post-acute care providers based on prior experience, either firsthand or through family and friends' recommendations (Baier et al., 2015; Gadbois et al., 2017). One or more formal caregivers, such as a hospital discharge planner, case manager, or physician, typically influence a patient's decision (Konetzka & Perraillon, 2016; Swope & Brown, 2015). For instance, a hospital discharge planner arranging a referral to home health could choose an agency for the patient or influence the patient's selection by streamlining the options offered (Baier et al., 2015; Swope & Brown, 2015). Thus, the observable demand response is a culmination of multiple agents' input.

Recognizing that many patients have difficulty selecting post-acute care, Medicare's first information disclosure initiative to support patient choice in home health began in November 2005. CMS produced a list of up to 11 quality measures for each home health agency (Jung et al., 2016). Jung et al. (2016) evaluated changes in patient choice in the three years preceding and following the release of the information. They concluded that an increase of one standard deviation in quality scores was associated with an increase of around 0.6 percentage points in market share. From a starting point of 20 percent, their findings indicate that the initiative did not meaningfully influence patient demand for higher-rated agencies (Jung et al., 2016). The modest estimates found by Jung et al. (2016) were consistent with findings in the nursing home setting, where CMS implemented a similarly structured information disclosure program (Werner et al., 2012).

Jung et al. (2016) argued that, combined with the plethora of information, the patient population's physical and cognitive limitations hampered information usage. In theory, consumers might compare agencies measure by measure. In practice, they would need to prioritize measures and reconcile discrepancies in measure-specific performance among agencies, which requires numeracy and can be cognitively taxing. Therefore, simplifying the information could be a viable improvement.

CMS updated the home health report card system with the home health star ratings more than a decade later. The new format summarized several agency-level clinical quality measures based on the agency's skill, effort, and patient characteristics. New and small agencies were ineligible (Centers for Medicare & Medicaid Services, 2015d). Since July 2015, each eligible agency has been rated between 1 and 5 stars in half-star increments every quarter, and their summary ratings have been made publicly available on the CMS Compare website. In January 2016, CMS developed a second set of star ratings to summarize patient experience in home health agencies. CMS also assigns 5-star ratings to skilled nursing facilities, hospitals, dialysis facilities, clinicians, and Medicare Advantage plans, in addition to home health care.

Studies have found larger marginal responses to the star rating format in the nursing home and Medicare Advantage markets. Perraillon et al. (2017) studied skilled nursing facilities, which serve patients similar to home health patients (Werner et al., 2019), using a regression discontinuity design. For each extra star, nursing facilities raised admissions rates by about 1 percentage point in the program's first six months. Lower-quality nursing homes saw no noteworthy effects. Darden and McCarthy (2015) found that Medicare Advantage plans increased their market share by 3 to 5 percentage points every half-star, with larger marginal effects among below-average plans.

Schwartz and colleagues (2022) used a discrete choice model, comparing home health admissions before and after the star ratings, to determine whether the star ratings affected the probability that Medicare FFS patients used a high-quality agency, i.e., 4 stars or above, or a low-quality agency, i.e., 2.5 stars or lower. They estimated that the first year of the Home Health Star Ratings program was associated with a 0.88 percentage points increase in new patient market share for high-quality agencies and a 0.68 percentage point loss for low-quality agencies. The authors also examined for heterogeneous responses by patient demographics and clinical needs and found that patients who were Black, dually enrolled in Medicaid, and with high cognitive functioning, were more likely to use agencies with 4 stars or above. Additionally, people admitted from the community were less likely to use agencies with 2.5 stars or fewer.

Like other studies using pre-post designs, however, the empirical approach used by Schwartz et al. makes it difficult to differentiate the effects of the 5-star format from the effects of other contemporaneous changes (e.g., learning, improvements in true quality). It also makes it difficult to examine the effects of a marginal increase in ratings on consumer choice for agencies in the middle of the star distribution. By construction, most home health agencies fall in the middle of the star distribution and patients usually only have choices between agencies with small differences in ratings, making this information crucial for future program design and enhancement.

III. Study Population, Data, and Outcome

A. STUDY POPULATION

This study examines Medicare fee-for-service beneficiaries served by a home health agency in one of the 50 US states and Washington, DC using individual data from January 2014 through December 2016. Patients were included if they had continuous fee-for-service coverage in the year before care began. They must have also started treatment from a star-rated agency during some point in the first six quarterly ratings.

I focus on Medicare fee-for-service patients because they can choose any Medicare-certified provider and their selection of agencies is not constrained by insurance restrictions, allowing for a more accurate measure of patient choice. Because patients are likely to choose an agency they have used before, I focus on new patients, which I define as Medicare fee-for-service patients without use of home health services in the 12 months before the start of home health care, consistent with prior approaches (Schwartz, Rahman, Thomas, Konetzka, & Mor, 2022).

B. DATA

The home health Outcome and Assessment Information Set (OASIS) comprises patient data such as service use dates, residential ZIP codes, patient characteristics, payment source, and whether the patient was admitted from an inpatient setting. The OASIS is collected by Medicare-certified home health agencies for all adult patients (Centers for Medicare & Medicaid Services, 2015c). I linked OASIS data to the Master Beneficiary Summary File to identify fee-for-service and Medicaid enrollment status.

The Home Health Compare website provides ratings and release dates. These ratings were released quarterly beginning July 16, 2015 based on care provided between October 1, 2013, and December 31, 2015, depending on the measure. I obtained the unrounded ratings underlying the publicly displayed ratings in Home Health Compare through a Freedom of Information Act data request to CMS.

I also gathered state, years in operation, and for-profit status of agencies from Home Health Compare. I used the fiscal year 2014 and 2015 Healthcare Cost Report Information System to identify agency-level affiliations with a chain organization.

C. OUTCOME: PATIENT SHARE

My primary objective is to determine whether star ratings resulted in patient demand for highly rated agencies at the margin. If patients use and respond to star ratings, then agencies with more stars should serve a greater share of patients in their market. I focus on ZIP codes since ZIP codes are the smallest geographic unit available to the public when searching for home health availability. For each quarterly release, I calculate the percent of new patients as the number of new Medicare fee-for-service home health patients treated by an agency out of all new Medicare fee-for-service home health patients in the agency's ZIP codes served during that period. Each quarterly release period includes the day after the release and up to, but not including the day of the subsequent release (Appendix Table 1). I exclude the day of release because Medicare provides the exact day but not the time that the new ratings are posted.

Agencies could change the geographic areas that they serve within a state. An agency that increases (or decreases) the number of ZIP codes served as a result of a higher star rating may appear to have lost (or gained) market shares. Therefore, if agencies are contracting or expanding their markets at the ZIP code level, this outcome measure could misrepresent consumer demand. I do not find evidence that agencies changed the number of ZIP codes they served (Appendix Table 2).

IV. Empirical Approach

To isolate the effects of one more half-star on new patient share, I use a regression discontinuity design that leverages CMS's composite star rating assignment rules. I compare agencies that are virtually identical but are barely on the opposite sides of an arbitrary star threshold (i.e., rounding of a continuous, underlying score to nearest half star). Precise knowledge of Medicare's rules provides the identifying source of variation.

Star ratings fall on a half-star scale between 1 and 5 stars. Most agencies fell in the 2.5- to 3.5-star range and the pattern was similar over time (Appendix Figure 1). To get to the final half-star rating, CMS averages across nine individual star ratings based on nine individual measures (Appendix Table 3). Each measure-specific star rating is rounded to the nearest half decimal point, which is then averaged to arrive at a composite, unrounded star rating that extends out to three decimal places. Unrounded star ratings were then rounded to the nearest half point to arrive at the final composite rating (Centers for Medicare & Medicaid Services, 2015d). For instance, an agency receives 2.5 stars if the agency's unrounded composite rating is 2.251 and receives 2 stars if its rating is 2.249. Therefore, as shown in Appendix Figure 2, the discontinuity is the rounding threshold cutoff (gray dashed line), the running score is the unrounded star ratings, and treatment is an additional half star.

I examine the effects of an additional half star on patient shares in two ways. First, I combine the thresholds and six quarterly releases into one sample, resulting in a total of 8,797 agencies. Each unrounded star rating within ±0.25 on either side of a threshold is included and centered at the rounding threshold. Combining the thresholds makes the assumption that the effects of one more half star are constant across the various thresholds (Cattaneo, Keele, Titiunik, & Vazquez-Bare, 2016). Conceptually, the effects could be argued to be homogeneous since crossing the threshold always results in the same treatment (a half star more) and choosing an agency with a higher rating is presumably always better than an agency with a lower rating, all other factors equal.

I also allow for heterogenous effects across thresholds. Thus, in the second way, I examine each threshold separately to determine whether having an additional half-star yield different effects across the distribution of the stars. Heterogeneity may be plausible if, for example, patients place more value on a change from 3.5 to 4 stars (from average to above average) than from 1 to 1.5 stars (worst to slightly less bad). For each threshold, agencies with unrounded star ratings of ±0.25 on either side of a threshold are included and centered at each rounding threshold.

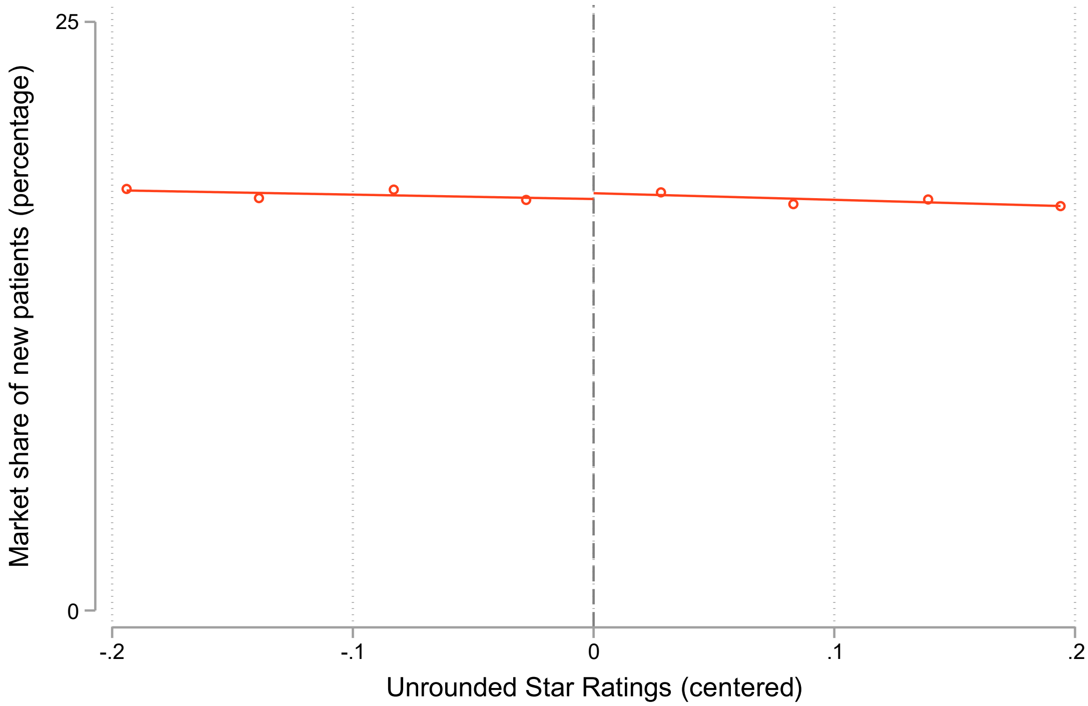
The unrounded star ratings cluster around certain values and are discrete because CMS takes the average of rounded numbers (Appendix Figure 3). An agency can only have a measure calculated if they had at least 20 patients that met the measure inclusion and exclusion criteria, and an agency can only have a star rating calculated if they had at least 5 measures (Centers for Medicare & Medicaid Services, 2015d). Even though raw unrounded star ratings extend out to three decimal places, many values are simply unattainable based on the combination of individually rounded star ratings.

The discrete nature of the unrounded star ratings has several implications, one of which is the possibility of heaping which could introduce bias if heaping is non-random (Barreca, Lindo, & Waddell, 2016). This can be an issue because there may be a disproportionate share of heaped versus non-heaped observations across the treatment and control groups. Indeed, the unrounded star ratings displayed presence of heaping, and in this case, non-heaped observations were the less common type (<1 percent).

A comparison of the characteristics of the heaped to non-heaped agencies suggest that there were systematic differences across these two groups and there is a need to adjust for compositional differences at heap points in the regression discontinuity analysis. Heaped agencies had been in operation longer (an average of 18 years versus 13 years), had more measures (9 versus 8), smaller percentage of for-profit agencies (76 versus 84), and more agencies belonging to a chain organization (19 percent versus 5 percent). Agencies in the heaped sample tended to be larger than those in the non-heaped sample. At baseline, there were almost nine times more new patients treated by agencies in the heaped sample and almost four times more Medicare patients. The non-heap points were also more common toward lower values of the unrounded star ratings (Appendix Figure 3). In other words, the data suggest that heaping, or rather, non-heaping, was not random. To account for potential bias resulting from influential non-heap observations, I follow the approach suggested by Barreca, Lindo, and Waddell (2016) and separate the non-heap points from the heap points. I focus on the estimates from the heap points as the heap sample is much larger and captures most of the home health agencies in the nation.

Because the running score is discrete, there is a finite number of values or mass points that the unrounded scores can take. This differs from the canonical continuity-based regression discontinuity design where the running score is assumed to be infinite. With a finite number of mass points, the canonical continuity-based parameter of interest, treatment effect at the cutoff, is not identifiable nonparametrically and requires parametric extrapolation if it is to be estimated (Cattaneo, Idrobo, & Titiunik, 2019).

Figure 1: Descriptive relationship between unrounded home health star ratings and share of new patients pooling all thresholds for agencies with a home health star rating from July 2015−December 2016



Notes: Pooled threshold sample includes agencies with unrounded star ratings that are centered at the rounding threshold and up to, but not including ±0.25 on either side. Each point represents an unrounded star rating value.

I use parametric extrapolation as my primary approach. Figure 1 shows a visual description of the data, which suggests that a linear model is appropriate to capture any changes in the outcome. In equation (1), my preferred specification, I use ordinary least squares regression to estimate the level shifts in the cross-sectional relationship between the share of new patients per agency per quarter and an additional half star (see Appendix I for other specifications as sensitivity tests). This specification performed the best among other alternatives in terms of Akaike information criterion, such as adjusting for quadratic functions of the unrounded star ratings and allowing for different slopes across the rounding threshold (Appendix Tables 4–6). It is also preferred due to small sample concerns and because higher-order polynomial specifications are subject to overfitting problems.

$$Pats_{jt} = \beta_0 + \beta_1 Star_{jt} + \beta_2 running_{jt} + \epsilon_{jt} \tag{1}$$

In this specification, $Pats_{jq}$ is the share of new patients per agency j in quarter $t \in \{1, 2, 3, 4, 5, 6\}$. $Star_{jt}$ is equal to 1 if the observation received a higher star rating (right of the threshold) or 0 otherwise; and $running_{jt}$ is the unrounded star rating for agency j in quarter t, centered at the rounding threshold, which is calculated from taking the difference between unrounded star rating and the threshold, where threshold can take on the values 1.25, 1.75, 2.25, 2.75, 3.25, 3.75, 4.25, and 4.75. Because I combine six quarterly releases, I have repeated observations. Therefore, to account for within-agency correlations from combining the six quarterly releases, I cluster standard errors at the home health agency level. $\widehat{\beta_1} > 0$ would be consistent with the notion that one more half star led to an increase in patient market share.

While my preferred specification does not include covariates, I include covariates as a sensitivity check for all regressions. Using pre-treatment data from January 2015 through June 2015, I calculate several agency-level characteristics. They include each agency's share of new Medicare fee-for-service patients, total new Medicare fee-for-service patients, total admissions across all payers, percent of patients who were receiving post-acute care, percent of patients by race and ethnicity, percent of patients by payer, percent of patients who were dually eligible for Medicare and Medicaid, mean patient age, and release date dummies. I also include the years that the agency had been in operation, whether the agency was a part of a chain organization, and whether it was a for-profit organization.

The canonical regression discontinuity design requires that the expected potential outcomes be smooth at the cutoff. For units immediately on either side of the threshold to be comparable, the threshold must be exogeneous to the running score and the rating assignment must be solely based on the running score with respect to the pre-determined thresholds. While this assumption is inherently untestable, it is more plausible if there are no discontinuities on variables that should not be affected by the treatment. I examine this assumption indirectly using several methods.

Table 1: Pre-treatment characteristics of the sample, January to June 2015

	Treatment (+1/2 star)		Control (-1/2 star)		Coeff.	(SE)	*p* value
Number of agencies (N)	8,090		8,015				
Number of observations (n)	20,794		20,454				
Agency characteristics							
Operational years	18.17	(13.03)	18.28	(12.96)	0.46	(0.27)	0.085
For profit	76%		76%		-0.01	(0.01)	0.391
Part of chain	29%		29%		-0.01	(0.01)	0.377
Patient characteristics							
New patient share	15.88	(21.29)	16.17	(21.53)	0.25	(0.44)	0.569
New patients, no.	37.06	(70.66)	35.74	(53.80)	0.75	(1.26)	0.549
All patients, No.	386.5	(956.96)	364.86	(671.46)	17.73	(16.21)	0.274
Percent Medicare FFS	80.02	(20.57)	80.11	(20.47)	-0.47	(0.42)	0.271
Percent Medicare Advantage	15.07	(17.39)	15.	(17.27)	0.37	(0.36)	0.303
Percent Medicaid	4.91	(9.10)	4.89	(8.95)	0.1	(0.19)	0.610
Percent dually eligible	40.36	(25.99)	40.29	(25.80)	0.28	(0.53)	0.604
Percent post acute	53.01	(23.37)	53.27	(23.11)	0.18	(0.48)	0.712
Mean age	74.85	(4.67)	74.91	(4.60)	-0.04	(0.10)	0.682
Percent female	61.73	(7.43)	61.83	(7.18)	0.1	(0.15)	0.502
Percent White	68.74	(30.13)	69.19	(29.89)	-0.42	(0.61)	0.498
Percent Black	16.17	(22.26)	15.98	(22.13)	0.22	(0.45)	0.630
Percent Hispanic	11.67	(21.89)	11.39	(21.51)	0.01	(0.45)	0.990

Notes: Regression coefficients are obtained from ordinary least squares regression assessing level shifts in the cross-sectional relationship between individual variables and an additional half star in rating, where ratings were assigned from July 2015 to December 2016. *P* values correspond to the null hypothesis that the average effect of one more half star on pre-treatment covariates at the cutoff is zero. Standard errors were clustered at the home health agency level.

First, I compare the characteristics of the sample by treatment status (Table 1). These descriptive statistics reveal that agencies with an additional half-star were comparable to those below the threshold. Both the treatment and control agencies had been in operation for nearly a decade on average, were predominantly for-profit organizations, and with about 20 percent in a chain. In the half year prior to the star ratings from January through June 2015, the average agency treated 16 percent of all new fee-for-service patients within its market. Across sex, race, and other patient characteristics, the groups were similar. These descriptive statistics support the idea that the treatment and control groups were comparable.

Second, I test for discontinuous jumps in agency-level characteristics, including the years of operation of the agency, the likelihood that the agency was for-profit, and likelihood that the agency was a part of a chain organization. I also test whether the ratings affected the types of patients treated by the agency in the half year before the start of the star ratings program. Out of the 16 placebo tests, I find one marginally significant (p = 0.085) difference of 0.46 (SE=0.27) years between the treatment and control group in terms of the length of time the agency had been in operation, but the difference was small. Overall, I do not find evidence of systematic differences between the two groups (Table 1).

Third, I examine the plausibility that home health agencies precisely manipulated their unrounded star ratings. The quarterly star ratings are computed by Medicare using historical data, and exact scores are calculated based on a national distribution. Thus, an individual agency would need to accurately predict their own score and the national distribution, well in advance, to manipulate their score. Although unlikely, I formally test for manipulation of the unrounded star ratings using the Frandsen manipulation test for discrete running score using the Stata command *rddistestk* (Frandsen, 2017). The Frandsen test uses the support points at and immediately adjacent to the threshold to test the null hypothesis that the probability mass function is smooth around the threshold.

To implement the test, one must specify the parameter $k \geq 0$ that determines the degree of deviations from linearity in the probability mass function that would lead the test to reject the null assumption of no manipulation. A small k means that small nonlinearities would result in the test rejecting the null hypothesis of no manipulation with high probability, while a large k means that the probability mass function can be highly nonlinear at the threshold before the test has sufficient power to detect manipulation. Following the approach outlined in Frandsen (2017) to estimate k, the p value is 1.000 when k=0.138. When k=0, the corresponding p value is 0.901, compatible with no manipulation (Appendix Table 4).

Lastly, while fitting a regression improves statistical precision as it allows me to incorporate information further from the cutoff, observations further from the thresholds are more likely to differ from the observations near the cutoff. Thus, I test the sensitivity of the estimates using specifications with varying flexibility (Appendix Tables 4–6) and I allow the bandwidths to range from ± 0.25, ± 0.15, and ± 0.11. However, because the running variable is discrete and it has a very coarse discretization of only eight support points (Appendix Figure 4), I am limited in how narrow the bandwidths around the threshold can be. Overall, I do not find evidence that the assumptions of the regression discontinuity estimator are violated.

V. Results

MAIN ESTIMATES

The estimates indicate that the effect of one more half star on the market share of new patients was small and non-significant (Table 2). Gaining a half star corresponded to an increase of 0.25 (SE = 0.45) percentage points, or approximately a 1-percent increase from an average of 18 percent; the corresponding 95 percent confidence interval rules out effects larger than 1.12 percentage points. When adjusted for covariates, the results are even smaller, with a point estimate of 0.03 (SE=0.19) percentage points and an upper bound of 0.40 percentage points. Although not directly comparable, these estimates are consistent with the most recent estimates from Schwartz et al. (2022), which suggested modest effects.

Table 2: Regression discontinuity estimates of the effect of having one more half star on market shares.

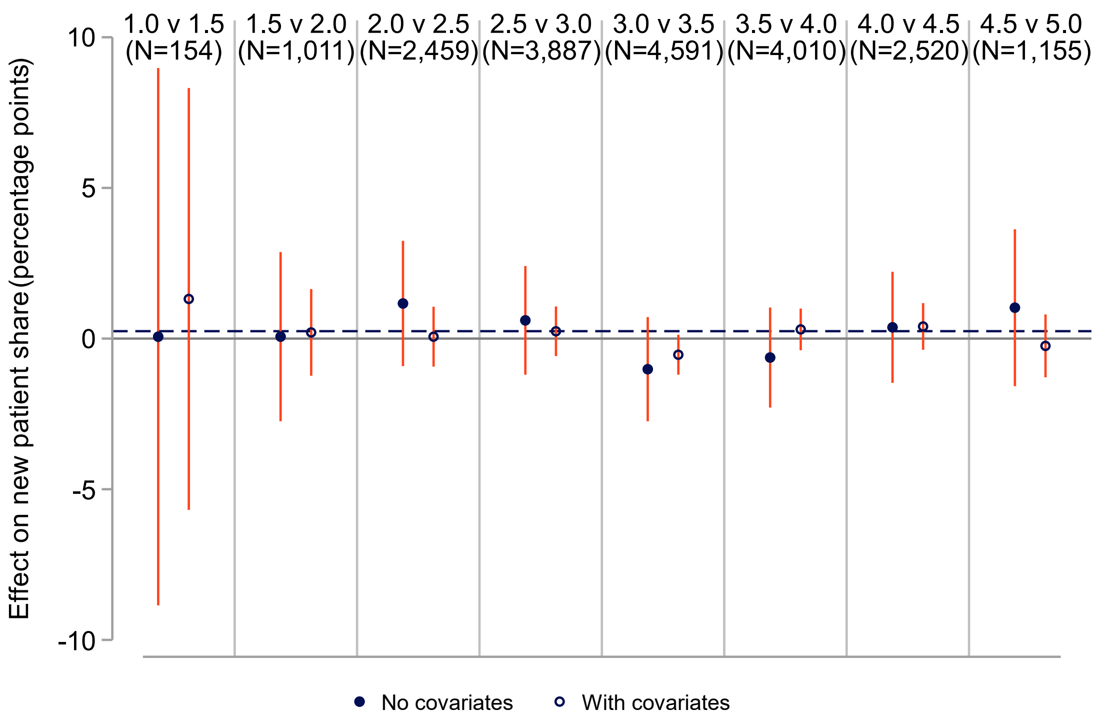
Pooled thresholds	Covariates	N agencies	Mean	(SD)	Coef.	(SE)
(-0.25, 0.25)		8,806	17.55	(21.81)		
	No				0.25	(0.45)
	Yes				0.03	(0.19)
(-0.15, 0.15)		8,633	17.55	(21.77)		
	No				0.17	(0.53)
	Yes				-0.17	(0.22)
(-0.11, 0.11)		8,112	17.58	(21.95)		
	No				0.80	(0.72)
	Yes				0.07	(0.30)

Notes: Estimates obtained from ordinary least squares regression assessing level shifts in the cross-sectional relationship between the share of new patients per agency and an additional half star in rating. Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors were clustered at the home health agency level.

The results were similarly small in magnitude and not statistically significant when the sample was restricted to narrower bandwidths around the cutoff. Estimates from the narrowest bandwidth, within ±0.11, suggest an increase of 0.80 (SE = 0.72) percentage points. Based on the 95 percent confidence interval, I am unable to rule out a larger effect of 2.20 percentage points, but a 2-percentage-point increase is still small, even though it is on the higher end of what home health report cards have shown so far.

The threshold-specific point estimates were generally similar to the point estimate from the main analysis and did not display clear patterns (Figure 2). For instance, unadjusted estimates show that 5 out of 8 point estimates were within a magnitude of 1-percentage point, including thresholds at the low (1 vs 1.5, 1.5 vs 2), middle (2.5 vs 3) and high end of the distribution (3.5 vs 4, 4 vs 4.5). While most point estimates were positive, thresholds for 3 vs 3.5 and 3.5 vs 4 were negative. Covariate adjusted estimates were similar, with 6 out of 8 within a half-percentage point in magnitude and with all estimates within 1-percentage point except for the estimate for the 1 vs 1.5 threshold.

Figure 2: Effects of one more half star on new patient market share across the star distribution



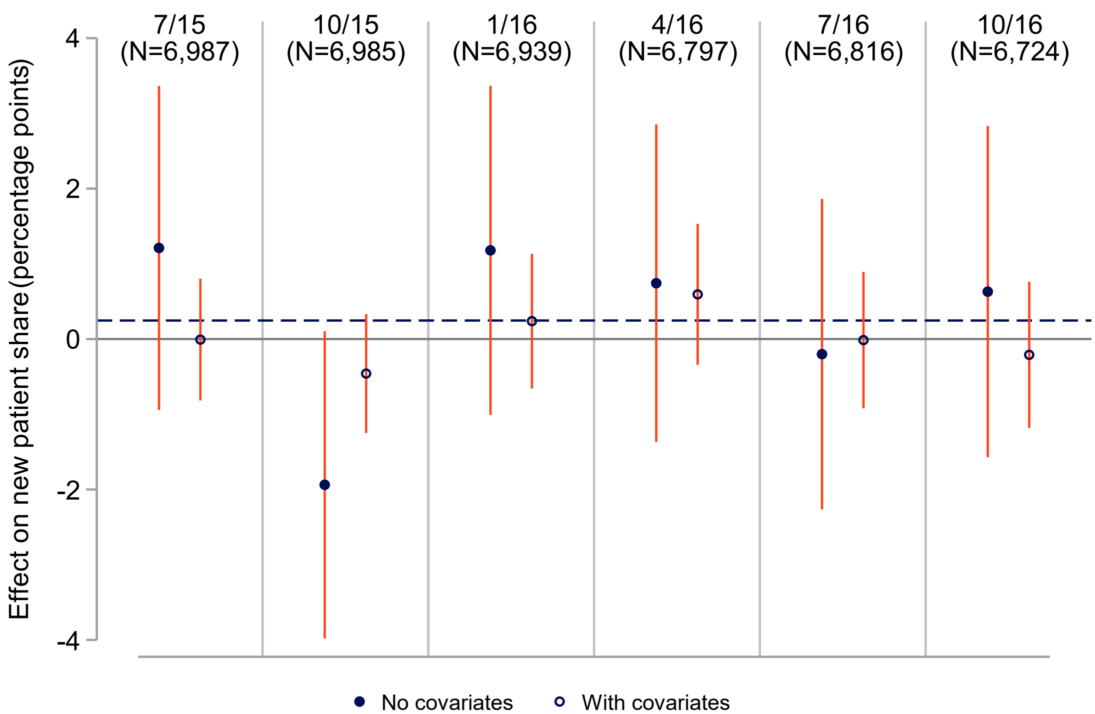
Notes: Graph displays point estimates and 95% confidence intervals. Horizontal, dashed blue line represents the pooled threshold estimate across all releases. Estimates obtained from ordinary least squares regression assessing level shifts in the cross-sectional relationship between the share of new patients per agency and an additional half star in rating. Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors were clustered at the home health agency level.

VI. Additional investigations

A. DYNAMIC EFFECTS

One potential reason for the null effects is that star ratings affect consumer demand differently over time. Patients and practitioners may need time to learn about the star ratings, leading to larger effects as the program matures. I first examine whether the effects of one more half star on patient shares increased over time (Figure 3). A positive correlation between the program age and patient shares would indicate delayed awareness of the ratings. There is no discernable pattern indicative of growing awareness of the program. In the quarter following the first release of star ratings, unadjusted point estimates indicate effects of 1.18 (SE=1.10) percentage points. By the last release in 2016, the point estimates were about 0.63 (SE=1.12) percentage points.

Figure 3: Effects of one more half star on new patient share across the first six release

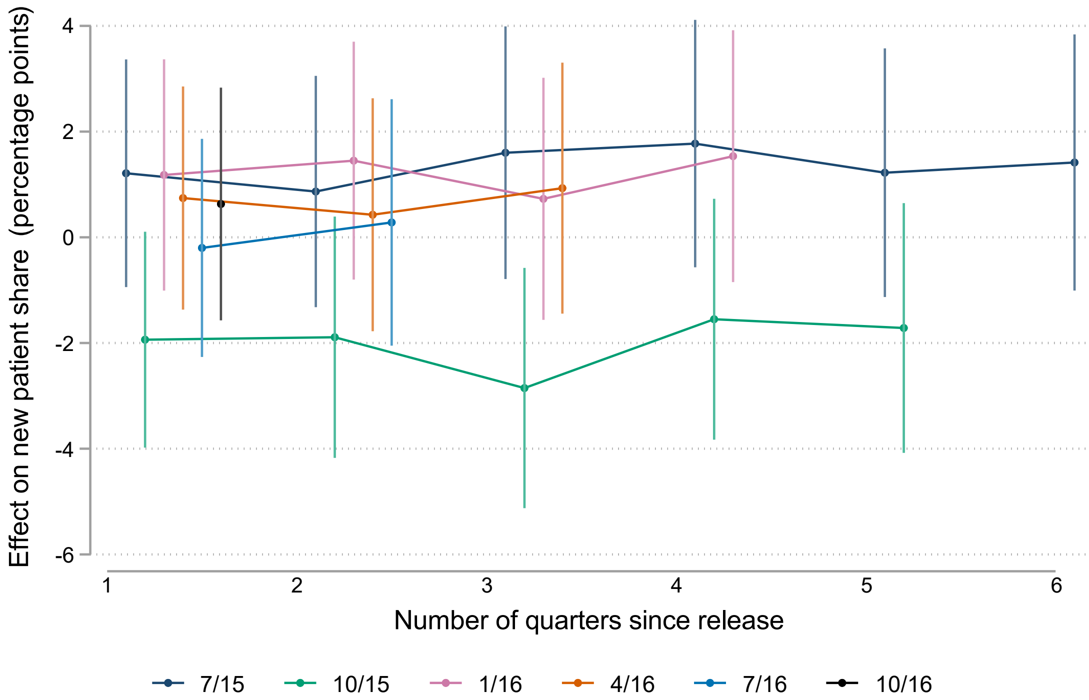


Notes: Graph displays point estimates and 95% confidence intervals. Horizontal, dashed blue line represents the pooled threshold estimate across all releases. Estimates obtained from ordinary least squares regression assessing level shifts in the cross-sectional relationship between the share of new patients per agency and an additional half star in rating. Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors were clustered at the home health agency level.

I supplement the release-specific analysis with Google Trends data to examine "home health star rating" search frequency from January 2014 through January 2020 in the US (Appendix Figure 5). More than 90 percent of adults 50 years or older own a computer or laptop and more than 65 percent of adults 70 years or older own a desktop. Smart phone use is also high, at approximately 73 percent among adults in their sixties and 55 percent among adults 70 years or older (Anderson, 2018). Since CMS's website is the public's source for star ratings, Google's search engine is likely to capture any fluctuations in interest in the ratings. These data show the relative frequency of searches over the analytic period, which provides an indication of trends. Search frequency peaked in July 2015, when the star ratings began, followed by a 50% drop. At least from Google searches, interest in home health star ratings did not increase over time.

Information delays could also affect the timing of demand responses. If a provider, such as a hospital, compiles discharge planning documents that include star ratings, patients may end up using outdated information. Information delays would delay demand, meaning increased patient shares may not be seen immediately after star ratings are released. To assess whether lags in consumer response to each release could have masked changes in patient demand, I examine up to six lags. For agencies rated in July 2015, for instance, I regard treatment status from the first release as fixed and compare outcomes for each subsequent quarter for the same cohort. This results in six cohorts of agencies. A positive correlation suggests information delays, a negative correlation indicates timely response to the ratings, and no correlation suggests no demand response. I do not find evidence of changing patient shares (Figure 4). Regardless of the lag, point estimates were flat, compatible with no demand response. Together, these results indicate that heterogeneous responses over time are unlikely to explain the null effect estimates in the first 1.5 years.

Figure 4: Lagged effects of one more half star on patient market share



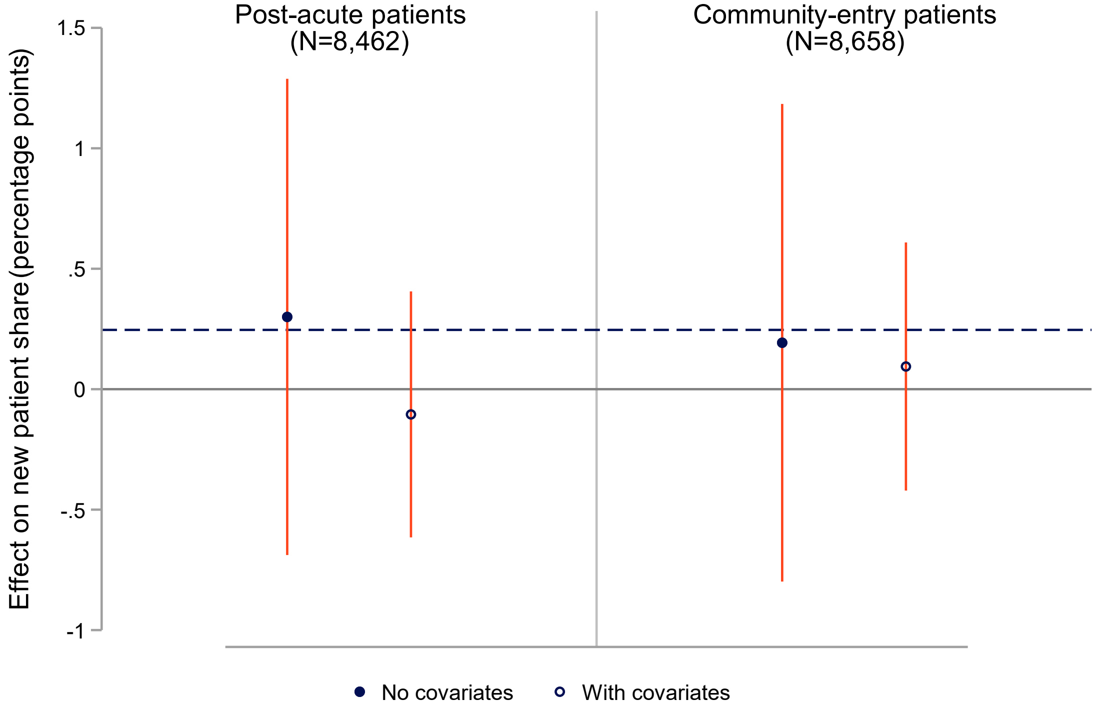
Notes: Graph displays point estimates and 95% confidence intervals. Estimates obtained from ordinary least squares regression assessing level shifts in the cross-sectional relationship between the share of new patients per agency and an additional half star in rating, where treatment status is determined by the initial release rating. Standard errors were clustered at the home health agency level.

B. SHOPPABLE SERVICES

The effects of star ratings should be more likely to affect patients who are more able to "shop" for home health services. If hospitals and/or physicians play a significant role in selecting where patients receive post-acute care and do not update their prior beliefs about home health agencies in response to a half-star rating change, star ratings would not alter patient choice. Patients entering home health from another formal care setting may not have sufficient time to search for information and therefore not be affected by the ratings. I analyze community-entry and post-acute care home health patients separately because discharging providers are less likely to influence community-entry patients and time may be less of a restriction when searching for information (Jung et al., 2016; Schwartz et al., 2022).

Figure 5 shows the results by patient type using equation 1 with pooled thresholds, with and without covariate controls. I find no evidence that an increase in star ratings at the margin produced larger effects among community-entry patients. Non-covariate adjusted regressions suggest that one more half star increased the share of new community-entry patients by 0.19 (SE=0.51) percentage points and post-acute admissions by 0.30 (SE=0.50) percentage points. Covariate adjustment attenuated the effect estimates, supporting the notion that there were no meaningful effects even among community-entry patients.

Figure 5: Effects of one more half star on new patient share by patient type



Notes: Post-acute patient refer to patients admitted to home health after discharge from an inpatient setting and community-entry patients are not. Graph displays point estimates and 95% confidence intervals. Horizontal, dashed blue line represents the pooled threshold estimate across all releases. Estimates obtained from ordinary least squares regression assessing level shifts in the cross-sectional relationship between the share of new patients per agency and an additional half star in rating. Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors were clustered at the home health agency level.

Another way that "shopping" could be constrained is if patients have few options to choose from. Options can be limited by the number of home health agencies available and by the number of star types available, both of which can be proxied by market competitiveness measured by the Herfindahl-Hirschman Index (HHI).

The first HHI measure, equation 2, is based on distinct agencies available in a ZIP code. I include all home health agencies that served any patient from any payment source in a given ZIP code from July 1, 2014, through June 30, 2015. I use ZIP codes prior to the star ratings to avoid endogeneity since home health agencies could change where they provide services based on market conditions.

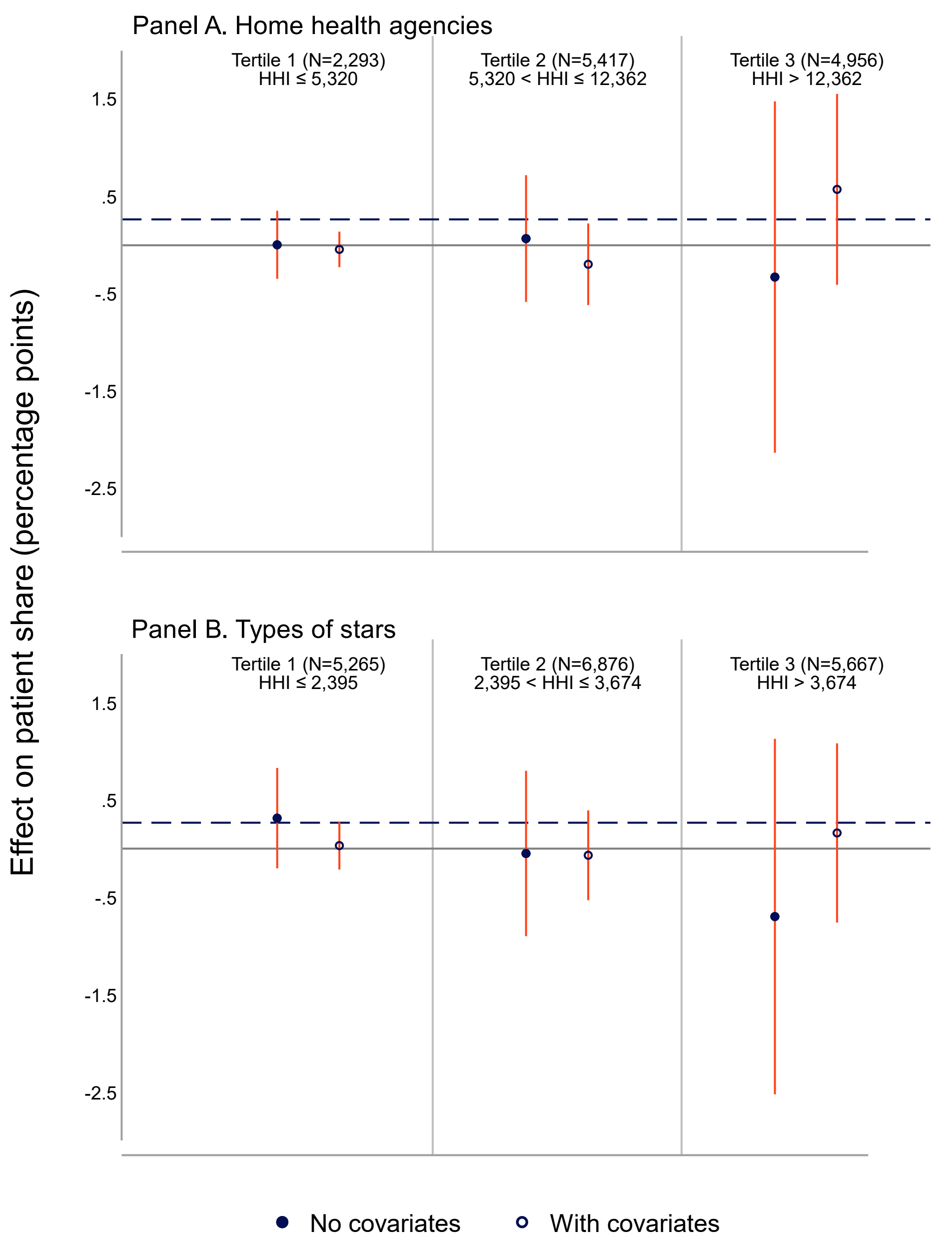
$$HHIZIP_j^{HHA} = \sum (market\ share\ of HHA\ i\ in\ ZIP\ j)^2 \quad (2)$$

Equation 3 provides the second HHI measure, which is based on distinct star types available in a ZIP code in a quarter.

$$HHIZIP_j^{Star} = \sum (market\ share\ of\ Star\ s\ in\ zip\ j)^2 \quad (3)$$

Figures 6 show the effects of one more half star across tertiles of the two HHI measures. Patterns were similar across both measures. For agencies in ZIP codes with the most competitive ZIP codes (tertile 1) the unadjusted estimates of one more half star on patient shares were within 0.5 percentage points ($HHIZIP_j^{HHA}$: 0.00 [SE =0.18]; $HHIZIP_j^{Star}$: 0.31 [SE=0.26]). Although the tertile 3 estimates were noisier, the point estimates were small and within 1 percentage points ($HHIZIP_j^{HHA}$: −0.33 [SE =0.92]; $HHIZIP_j^{Star}$: −0.70 [SE=0.93]). Thus, there was no evidence that patients in markets with more options were more likely to use the star rating information.

Figure 6: Effects of one more half star across ZIP codes with varying levels of competition measured by the number of agencies and star types



Notes: HHI = Herfindahl-Hirschman Index. Graph displays point estimates and 95% confidence intervals. Horizontal, dashed blue line represents the pooled threshold estimate across all releases. Estimates obtained from ordinary least squares regression assessing level shifts in the cross-sectional relationship between the share of new patients per agency and an additional half star in rating. Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors were clustered at the home health agency level.

C. SUPPLY-SIDE CONSIDERATIONS

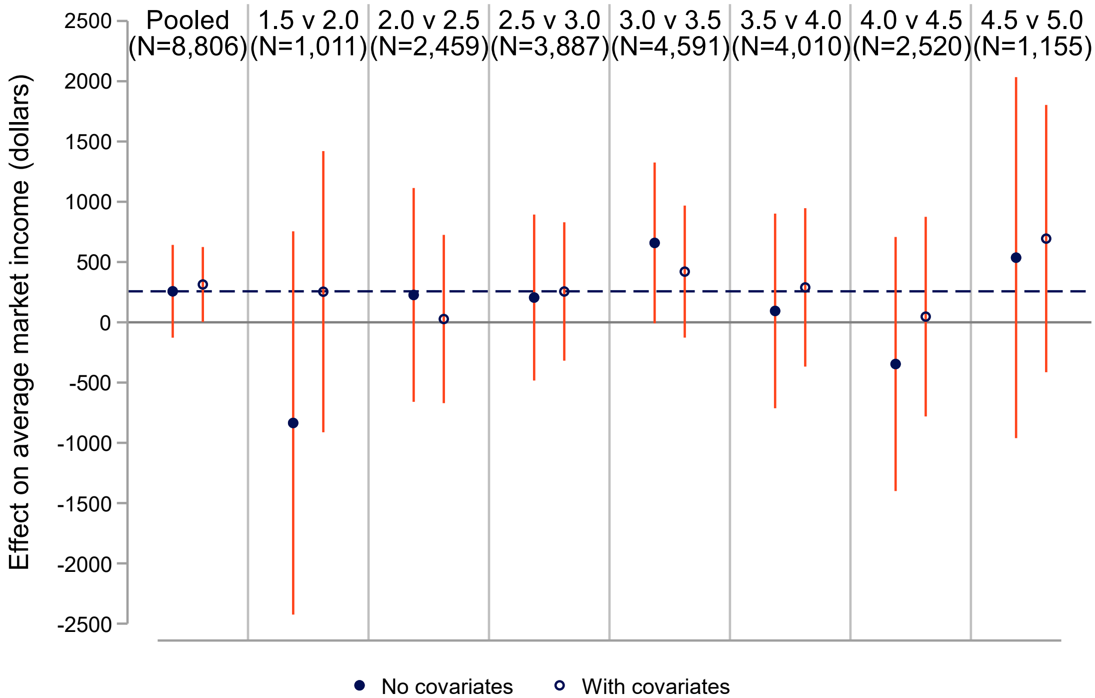
Even if the availability of star ratings increases consumer demand for higher rated agencies, capacity constraints may prevent agencies from treating more patient in the short run. Since agencies also cannot raise prices to recoup the benefits of increased demand, as prices are either administratively established by public payers or negotiated well in advance with private payers, home health agencies may resort to being more selective. Therefore, one possible effect of the star ratings is not an increase in new patient share, but greater opportunities among high-performing home health agencies to cherry-pick more desirable patients.

One measure of the desirability of a patient that is easily observed by home health agencies is the patient's local ZIP code income. Local area income is correlated with people's abilities to manage their health and crime rates (J. Chen, Mortensen, & Bloodworth, 2014; Dong, Egger, & Guo, 2020), both of which are factors cited as leading reasons for home health agencies to refuse patients (Centers for Medicare & Medicaid Services). I use median income among people 65 years or older at the ZIP code level as a proxy measure for market desirability. To estimate the average income of ZIP codes served by each agency, I use the American Community Survey to get the median income among people 65 years or older. I then calculate a weighted average of incomes among ZIP codes served by an agency, where the weight is each ZIP code's share of episodes for an agency in each period. The average income of ZIP codes served by agencies was $41,511 and was modestly higher for agencies with higher star ratings. For instance, the mean ZIP code income for agencies at the cusp of 1.5 and 2 stars was $37,091 while the mean was $41,756 for agencies at the cusp of 4.5 and 5 stars. An increase in the average income of ZIP codes served by an agency with an additional half star would imply a shift toward more desirable markets, compatible with patient selection.

The effect of one more half star on the average income of ZIP codes served by an agency was $257 (SE=196) with all thresholds pooled (*p* value = 0.190) and $314 (SE=158) (*p* value = 0.048) with covariate adjustment (Figure 7). Across the star thresholds, most of the point estimates were positive and not statistically significant. The largest point estimate was for 4.5 vs. 5 stars, $694 (SE=565), followed by 3 vs 3.5 stars, $421 (SE=279). Relative to the baseline average, all point estimates were small in magnitude, ranging from 0.0007 percent to 1.7 percent.

Thus, while there is some suggestive evidence that agencies with one more half-star treated more patients from higher-income ZIP codes in the quarter after the ratings were released, the effects were modest.

Figure 7: Effects of one more half star on the average income of ZIP codes served



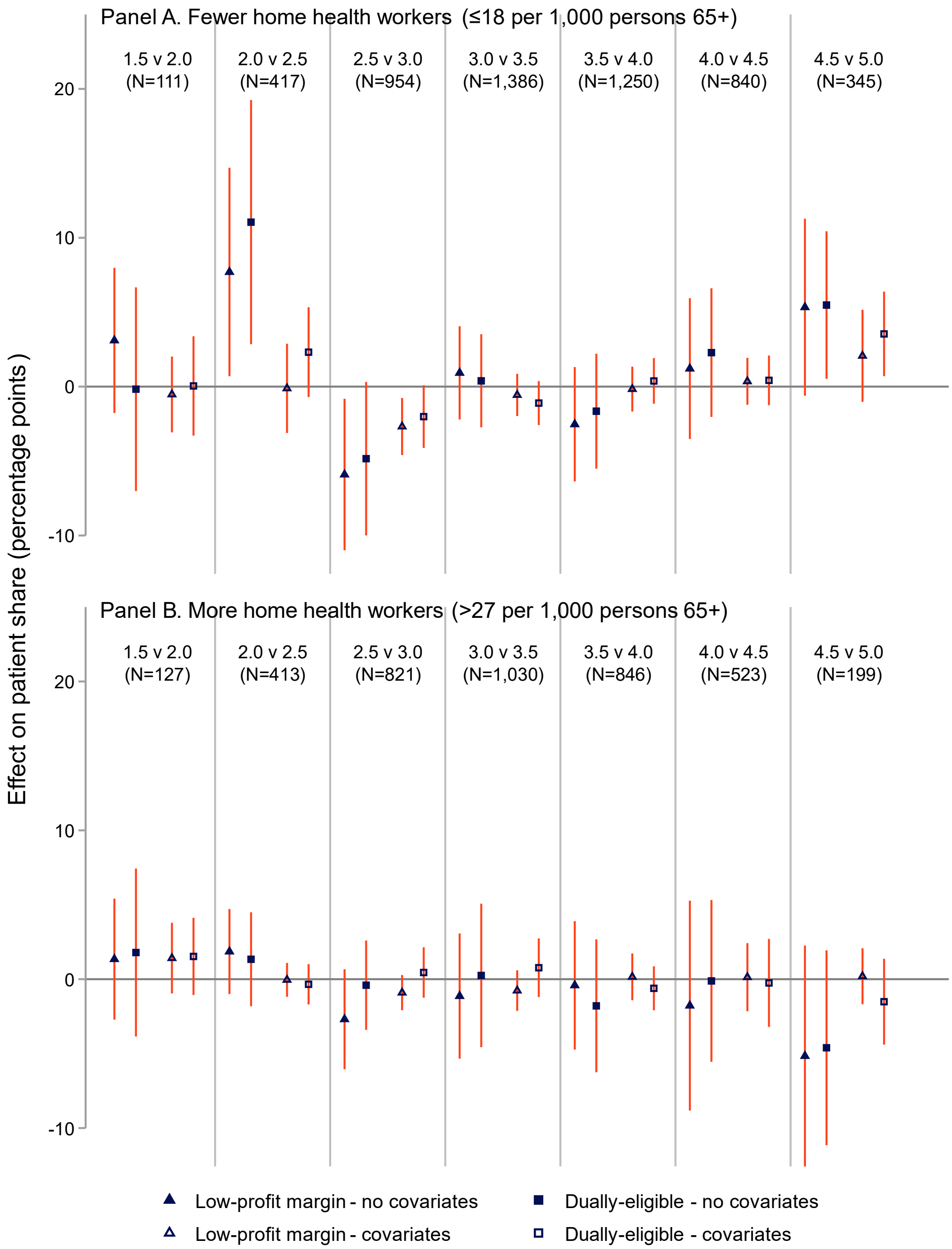
Notes: Pooled threshold sample includes agencies with unrounded star ratings that are centered at the rounding threshold and up to, but not including ±0.25 on either side. 1 vs 1.5 stars were excluded from individual threshold analyses due to small sample sizes (n=154). Average income averages across ZIP codes served by each agency and focuses on the median ZIP-code level income of persons 65 years or older. Baseline averages are: Pooled - $41,511, 1.5 v 2 - $37,800, 2 v 2.5 - $39,145, 2.5 v 3 - $40,832, 3 v 3.5 - $42,123, 3.5 v 4 - $42,779, 4 v 4.5 - $43,017, 4.5 v 5 - $41,756.

An alternative approach to proxy for patient desirability is to use patient- rather than market-level measures. I focus on the share of low-profit margin Medicare patients and the share of Medicare-Medicaid dually eligible patients (Centers for Medicare & Medicaid Services). I define low-profit margin patients as Medicare FFS enrollees with clinical characteristics associated with lower profit margins than other Medicare FFS patients on average, which includes people with poor control of clinical conditions (10 percent lower), with traumatic wounds or ulcers (20 percent lower), with significant bathing needs (20 percent lower), have overall high risk (20 percent lower), and recipients of intravenous therapy or parenteral nutrition at home (15 percent lower); Medicare-Medicaid dually-eligible patients are also associated with decreased profit margins (20 percent lower) and generally have more complex social and clinical needs than Medicare-only populations which may make them more costly for home health agencies to manage. I identify low-profit margin patients using OASIS data and dually eligible status from the MSBF. For this analysis, my sample includes all Medicare FFS patients, regardless of whether they used home health care in the past year.

Because I anticipate that decisions to admit the marginal patient are affected by capacity constraints, I proxy for capacity using county-level home health worker availability per 1,000 persons 65 years or older. Home health agencies depend heavily on labor inputs from health care workers such as nurses, therapists, and aides, and worker shortages are often touted as reasons for home health capacity constraints by industry experts (V. Chen, 2018; Galewitz, 2021). I use county-level data from the Quarterly Census of Employment and Wages to get the number of workers in home health care services (NAICS 6216). I obtain the population size of people 65 years or older from the 2014 American Community Survey. Markets with fewer workers are ZIP codes in the bottom tertile while markets with many workers are ZIP codes in the top tertile of workers available.

The patterns of the estimated effects of one more half star across the star distribution varied by area-level worker capacity (Figure 8). For agencies serving ZIP codes with fewer workers, both adjusted and unadjusted estimates suggest increased shares of less desirable patients for agencies at the tails of the star distribution and potentially decreased shares for those in the middle (Panel A). In contrast, none of the adjusted point estimates for ZIP codes with more workers exceeded 2 percentage points and there was also less variation across the distribution (Panel B). In areas with fewer workers, agencies in the middle of the distribution, which are the most prevalent types of agencies, may have admitted fewer low-profit margin and dually eligible patients after receiving a higher rating, diverting these patients to lower-rated agencies.

Figure 8: Effects of one more half star on low-profit margin and Medicare-Medicaid dually eligible patient share in ZIP codes with varying home health care worker availability



Notes: Pooled threshold sample includes agencies with unrounded star ratings that are centered at the rounding threshold and up to, but not including ±0.25 on either side. 1 vs 1.5 stars were excluded from individual threshold analyses due to small sample sizes. Panel A displays the bottom tertile of worker to population ratio and Panel B displays the top tertile. Low-profit margin patients are individuals with clinical characteristics associated with lower profit margins than other Medicare fee-for-service enrollees patients on average, which includes people with poor control of clinical conditions (10 percent lower), with traumatic wounds or ulcers (20 percent lower), with significant bathing needs (20 percent lower), have overall high risk (20 percent lower), and recipients of intravenous therapy or parenteral nutrition at home (15 percent lower).

However, if agencies in areas with fewer workers in the middle of the distribution were able to cherry-pick preferrable patients due to increased patient demand, one would expect an increase in patient shares among higher rated home health agencies where there were more workers, which was not observed. Thus, together with the findings from the market income analysis, these results suggest that the star ratings did not have a meaningful effect on how patients chose home health agencies or how home health agencies chose patients.

D. RANDOMIZATION INFERENCE

With few mass points in the running score, an alternative to imposing parametric assumptions to extrapolate at the cutoff is to use local randomization, which changes the parameter of interest from the treatment effect at the cutoff to the window near the cutoff (Cattaneo et al., 2019). In this approach, the interval between the two closest mass points on each side of the cutoff is the minimum bandwidth. So long as local randomization holds for the smallest window, Fisherian inference can be used to test the null hypothesis that one-more half star has no effect on outcomes for any unit in the sample.

To discern whether local randomization holds for the smallest window [−0.028, 0.028], I use a Binomial test to check whether treatment assignment was equally distributed across the cutoff with a probability of 0.5. The Binomial test is an alternative to continuity-based approaches such as the Frandsen density test (Cattaneo et al., 2019).

Most results indicated no statistical significance between the probability of treatment when examining the two closest mass points, −0.028 and 0.028 (Appendix Table 4). The exceptions were for the thresholds 1.5 vs 2 (*p* value = 0.008) and 4.5 vs 5 (*p* value=0.024). Threshold 2 vs 2.5 stars was marginally statistically significant (*p* value = 0.065). Although it is preferable to have balance in sizes between the treatment and control groups, as long as the observations are on average similar, size imbalances do not affect inference (Cattaneo et al., 2019).

To determine if the treatment and control groups were similar on average, I estimate the effects of one more half star on predetermined characteristics using the Stata command *rdrandinf* for the pooled (Appendix Table 8) and threshold-specific samples (Appendix Tables 8−15). For

12 out of 144 tests, I reject the Fisherian sharp null hypothesis that an additional half star had no effect. For both of the thresholds with size imbalances that were statistically significant, I failed to reject the null. Overall, the falsification analyses provided no evidence of sorting or selection that would have indicated differences between the treatment and control groups.

I examine all previously discussed outcomes and subgroups using Fisherian inference. Similar to the parametric results, I find no evidence that a marginal increase in star ratings led to increased patients shares (Appendix Table 17). The results were also similar when considering variation by release (Appendix Table 18), and by patients who differed in their likelihood of shopping for home health services (Appendix Table 19). I also fail to reject the null for post-acute and community-entry patients and in competitive ZIP codes with more home health agencies and star types (Appendix Table 19).

Results focusing on supply-side selection behaviors were also generally similar between the two approaches. The pooled threshold estimate suggests an increase of $147 ($p$ value = 0.426) from a baseline of $41,530 per year although not statistically significant at the $p<0.05$ level (Appendix Table 20). Moreover, similar to the parametric analyses, the randomization inference results indicate that shares of less desirable patients for agencies at the tails of the star distribution increased and potentially decreased for those in the middle (Appendix Table 21). In other words, both analytic approaches provide suggestive evidence of small cherry-picking behaviors by home health agencies.

VII. Conclusion

This study examines an intervention to mitigate imperfect information in the home health care industry. I use national data to understand how the home health 5-star ratings affect patient choice at the margin in the first year and a half of the program. The central finding is that the 5-star ratings have no discernible effect on consumer demand. To establish this, I show that one more half-star (i) does not increase market shares for home health agencies, (ii) does not result in heterogeneous effects over the star rating distribution, over time, or for patients with more opportunities to search for information or with more home health options to consider, (iii) may lead to patient selection behaviors by home health agencies, but only modestly.

One may question why the star ratings did not have a larger effect on consumer demand particularly since costs are not competing factors and larger effects have been observed for skilled nursing facilities. General awareness of home health report cards may be low (Baier et al., 2015). Low responsiveness could also be due to preconceived notions about low variability in home health quality. In interviews with hospital discharge planners from over 30 hospitals in Michigan conducted from June 2020 through September 2020, many discharge planners believed that home health agencies were similar in quality, unlike nursing facilities. Consequently, a lack of awareness combined with pre-existing beliefs of patients and their agents could deter the influence of the information.

Another possibility is that the star ratings information does not correlate with dimensions of quality important to consumers. Home health agencies self-reported the assessment data for eight out of the nine quality measures used to determine star ratings. As a result, the star ratings may be susceptible to gaming, and patients and their agents may not find the information useful in gauging true quality. Thus, the utility of the star ratings information may be low, leading to no discernable effect on patient demand.

Ultimately, the goal of information disclosure policies is to improve patient choice and increase provider accountability. Rather than targeting fee-for-service patients, CMS might be better served if the ratings were designed for Medicare Advantage plans or referring providers instead. If private insurers respond to ratings or if referring providers are incentivized to send patients to highly rated care, then star ratings could add value even if there is minimal use from fee-for-service Medicare patients. Even if CMS adapts the star ratings and elicits a larger demand response in the future, such as by targeting health plans or providers, any designs must guard against unintended side effects such as for providers to cherry-pick patients, manipulate performance measures, or neglect unmeasured tasks (Dranove et al., 2003; Eggleston, 2005). It remains to be seen whether report cards can be an effective tool for improving the health care system and patient welfare.

References

Anderson, G. O. (2018). Getting Connected: Older Americans Embrace Technology to Enhance Their Lives. *AARP Research.* Retrieved from https://www.aarp.org/research/topics/technology/info-2018/technology-use-attitudes.html

Avalere. (2015). *Home Health Chartbook 2015: Prepared for the Alliance for Home Health Quality and Innovation*. Retrieved from http://ahhqi.org/images/uploads/AHHQI_2015_Chartbook_FINAL_October.pdf

Baier, R. R., Wysocki, A., Gravenstein, S., Cooper, E., Mor, V., & Clark, M. (2015). A Qualitative Study of Choosing Home Health Care After Hospitalization: The Unintended Consequences of 'Patient Choice' Requirements. *Journal of General Internal Medicine, 30*(5), 634-640. doi:10.1007/s11606-014-3164-7

Barreca, A. I., Lindo, J. M., & Waddell, G. R. (2016). Heaping-Induced Bias in Regression-Discontiniuty Designs. *Economic Inquiry, 54*(1), 268-293. doi:https://doi.org/10.1111/ecin.12225

Brook, R. H., McGlynn, E. A., Shekelle, P. G., Marshall, M., Leatherman, S., Adams, J. L., . . . Klein, D. J. (2002). *Report Cards for Health Care: Is Anyone Checking Them?* Santa Monica, CA: RAND Corporation.

Bundorf, M. K., Chun, N., Goda, G. S., & Kessler, D. P. (2009). Do markets respond to quality information? The case of fertility clinics. *Journal of Health Economics, 28*(3), 718-727. doi:https://doi.org/10.1016/j.jhealeco.2009.01.001

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019). A Practical Introduction to Regression Discontinuity Designs: Foundations. *Cambridge Elements: Quantitative and Computational Methods for Social Science, 2*. doi:10.1017/9781108684606

Cattaneo, M. D., Keele, L., Titiunik, R., & Vazquez-Bare, G. (2016). Interpreting Regression Discontinuity Designs with Multiple Cutoffs. *The Journal of Politics, 78*(4), 1229-1248. doi:10.1086/686802

Centers for Medicare & Medicaid Services. *Report to Congress Medicare Home Health Study: An Investigation on Access to Care and Payment for Vulnerable Patient Populations*. Retrieved from https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HomeHealthPPS/Downloads/HH-Report-to-Congress.pdf

Centers for Medicare & Medicaid Services. (2015a). *Frequently Asked Questions (FAQs) about the Quality of Patient Care Star Ratings*. Retrieved from https://www.cms.gov/files/document/quality-patient-care-star-rating-faqsupdated-5-20-15-sxfpdf

Centers for Medicare & Medicaid Services. (2015b). Home Health Compare Quality of Patient Care Star Ratings [Press release]. Retrieved from https://www.cms.gov/newsroom/fact-sheets/home-health-compare-quality-patient-care-star-ratings

Centers for Medicare & Medicaid Services. (2015c). *OASIS-C1/ICD-10 Guidance Manual*. Retrieved from https://www.cms.gov/files/document/cy2015-home-health-archives.pdf

Centers for Medicare & Medicaid Services. (2015d). *Quality of Patient Care Star Ratings Methodology*. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-

Assessment-Instruments/HomeHealthQualityInits/Downloads/Quality-of-Patient-Care-Star-Ratings-Methodology-Report-updated-5-11-15.pdf

Centers for Medicare & Medicaid Services. (2021). *National Health Expenditures: Table 14 Home Health Care Expenditures*. Retrieved from: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical

Chen, J., Mortensen, K., & Bloodworth, R. (2014). Exploring Contextual Factors and Patient Activation: Evidence From a Nationally Representative Sample of Patients With Depression. *Health Education & Behavior, 41*(6), 614-624. doi:10.1177/1090198114531781

Chen, V. (2018). *Paying to stay: effects of vertical integration and competition in a post-acute care setting.*

Darden, M., & McCarthy, I. M. (2015). The Star Treatment: Estimating the Impact of Star Ratings on Medicare Advantage Enrollments. *Journal of Human Resources, 50*(4), 980-1008. doi:10.3368/jhr.50.4.980

Dong, B., Egger, P. H., & Guo, Y. (2020). Is poverty the mother of crime? Evidence from homicide rates in China. *PloS one, 15*(5), e0233034-e0233034. doi:10.1371/journal.pone.0233034

Dranove, D., Kessler, D., McClellan, M., & Satterthwaite, M. (2003). Is More Information Better? The Effects of "Report Cards" on Health Care Providers. *Journal of Political Economy, 111*(3), 555-588. doi:10.1086/374180

Eggleston, K. (2005). Multitasking and mixed systems for provider payment. *Journal of Health Economics, 24*(1), 211-223. doi:10.1016/j.jhealeco.2004.09.001

Frandsen, B. (2017). Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete. In (pp. 281-315): Emerald Publishing Limited.

Gadbois, E. A., Tyler, D. A., & Mor, V. (2017). Selecting a Skilled Nursing Facility for Postacute Care: Individual and Family Perspectives. *Journal of the American Geriatrics Society*. doi:10.1111/jgs.14988

Galewitz, P. (2021, 06//). With Workers in Short Supply, Seniors Often Wait Months for Home Health Care. *Kaiser Health News*.

Grabowski, D. C., & Town, R. J. (2011). Does Information Matter? Competition, Quality, and the Impact of Nursing Home Report Cards. *Health Services Research, 46*(6pt1), 1698-1719. doi:10.1111/j.1475-6773.2011.01298.x

Hibbard, J. H., & Peters, E. (2003). Supporting Informed Consumer Health Care Decisions: Data Presentation Approaches that Facilitate the Use of Information in Choice. *Annual Review of Public Health, 24*(1), 413-433. doi:10.1146/annurev.publhealth.24.100901.141005

Jung, J. K., Wu, B., Kim, H., & Polsky, D. (2016). The Effect of Publicized Quality Information on Home Health Agency Choice. *Medical Care Research and Review, 73*(6), 703-723. doi:10.1177/1077558715623718

Kolstad, J. T., & Chernew, M. E. (2009). Quality and Consumer Decision Making in the Market for Health Insurance and Health Care Services. *Medical Care Research and Review, 66*(1_suppl), 28S-52S. doi:10.1177/1077558708325887

Konetzka, R. T., & Perraillon, M. C. (2016). Use Of Nursing Home Compare Website Appears Limited By Lack Of Awareness And Initial Mistrust Of The Data. *Health Affairs, 35*(4), 706-713. doi:10.1377/hlthaff.2015.1377

Medicare Payment Advisory Commission. (2020). *Chapter 9: Home health care services (March 2020 report)*. Retrieved from Washington, DC: https://www.medpac.gov/document/http-www-medpac-gov-docs-default-source-reports-mar20_medpac_ch9_sec-pdf/

Medicare Payment Advisory Commission. (2021). *Chapter 8: Home Health Care Services*. Retrieved from Washington, DC: https://www.medpac.gov/wp-content/uploads/2021/10/mar21_medpac_report_ch8_sec.pdf

Murtaugh, C. M., Timothy R. Peng, Stanley Moore, and Gill A. Maduro. (2008). *Assessing Home Health Care Quality for Post-Acute and Chronically Ill Patients: Final Report*. Retrieved from https://aspe.hhs.gov/reports/assessing-home-health-care-quality-post-acute-chronically-ill-patients-final-report-1

Perraillon, M. C., Konetzka, R. T., He, D., & Werner, R. M. (2017). Consumer Response to Composite Ratings of Nursing Home Quality. *American Journal of Health Economics*, 1-36. doi:10.1162/ajhe_a_00115

Peters, E., Dieckmann, N., Dixon, A., Hibbard, J. H., & Mertz, C. K. (2007). Less Is More in Presenting Quality Information to Consumers. *Medical Care Research and Review, 64*(2), 169-190. doi:10.1177/1077558707064002030l

Ramanarayanan, S., & Snyder, J. (2012). Information Disclosure and Firm Performance: Evidence from the Dialysis Industry. *SSRN*. doi:http://dx.doi.org/10.2139/ssrn.2152796

Reineck, L. A., & Kahn, J. M. (2013). Quality Measurement in the Affordable Care Act. A Reaffirmed Commitment to Value in Health Care. *American Journal of Respiratory and Critical Care Medicine*, *187*(10), 1038-1039. doi:10.1164/rccm.201302-0404ED

Schwartz, M. L., Rahman, M., Thomas, K. S., Konetzka, R. T., & Mor, V. (2022). Consumer selection and home health agency quality and patient experience stars. *Health Services Research, 57*(1), 113-124. doi:10.1111/1475-6773.13867

Swope, C., & Brown, H. (2015). *Inside the mind of the hospital discharge planner*. Retrieved from https://www.advisory.com

Werner, R. M., Coe, N. B., Qi, M., & Konetzka, R. T. (2019). Patient Outcomes After Hospital Discharge to Home with Home Health Care vs to a Skilled Nursing Facility. *JAMA Internal Medicine*, *179*, 617-623. doi:doi.org/10.1001/jamainternmed.2018.7998

Werner, R. M., Konetzka, R. T., & Polsky, D. (2016). Changes in Consumer Demand Following Public Reporting of Summary Quality Ratings: An Evaluation in Nursing Homes. *Health Services Research, 51*, 1291-1309. doi:10.1111/1475-6773.12459

Werner, R. M., Norton, E. C., Konetzka, R. T., & Polsky, D. (2012). Do consumers respond to publicly reported quality information? Evidence from nursing homes. *Journal of Health Economics, 31*(1), 50-61. doi:10.1016/j.jhealeco.2012.01.001

Appendices

I. Additional Specifications

1. Linear (preferred)	$Pats_{jt} = \beta_0 + \beta_1 Star_{jt} + \beta_2 running_{jt} + X\beta_{jt} + \epsilon_{jt}$
2. Linear interaction	$Pats_{jt} = \beta_0 + \beta_1 Star_{jt} + \beta_2 running_{jt} + \beta_3 running_{jt} \times Star_{jt} + X\beta_{jt} + \epsilon_{jt}$
3. Quadratic	$Pats_{jt} = \beta_0 + \beta_1 Star_{jt} + \beta_2 running_{jt} + \beta_3 running_{jt}^2 + X\beta_{jt} + \epsilon_{jt}$
4. Quadratic interaction	$Pats_{jt} = \beta_0 + \beta_1 Star_{jt} + \beta_2 running_{jt} + \beta_3 running_{jt}^2 + \beta_4 running_{jt} \times Star_{jt} + \beta_5 running_{jt}^2 \times Star_{jt} + X\beta_{jt} + \epsilon_{jt}$

Where:

$Pats_{jt}$ is the number of new patients per agency j in quarter t; β_0 is the mean number of new patients after conditioning on the unrounded star ratings and covariates; $Star_{jt}$ is equal to 1 if the observation received a higher star rating (right of the threshold) or 0 otherwise; and $running_{jt}$ is the unrounded star rating for agency j in quarter t, centered at the rounding threshold; $X\beta_{jt}$ is a vector of agency-level covariates and dummy variables for star rating releases. Standard errors were clustered at the home health agency level. In all specifications, β_1 provides the marginal effect of having one more half-star at the rounding threshold.

II. Appendix Tables

Appendix Table 1: Star rating release dates and days included in analysis.

Releases examined in study	Days in post-release period
Quarter 1 (July 16, 2015)	83
Quarter 2 (October 8, 2015)	111
Quarter 3 (January 28, 2016)	82
Quarter 4 (April 20, 2016)	83
Quarter 5 (July 13, 2016)	97
Quarter 6 (October 19, 2016)	73

Appendix Table 2. Effects of one more half star on the number of ZIP codes served.

			Unadjusted			Adjusted		
	Mean	(SD)	Coef.	(SE)	*p* value	Coef.	(SE)	*p* value
Pooled thresholds	35.81	(31.64)	0.87	(0.67)	0.197	0.52	(0.52)	0.312
By threshold								
1 v 1.5	29.13	(22.85)	-13.88	(18.98)	0.466	-14.55	(10.72)	0.177
1.5 v 2	32.31	(26.40)	-2.41	(2.37)	0.308	-2.04	(1.62)	0.207
2 v 2.5	33.24	(30.46)	2.89	(1.53)	0.059	1.27	(1.07)	0.234
2.5 v 3	34.07	(29.65)	1.70	(1.17)	0.146	0.78	(0.93)	0.398
3 v 3.5	36.63	(32.62)	1.32	(1.27)	0.297	0.26	(0.93)	0.776
3.5 v 4	37.25	(34.36)	0.75	(1.46)	0.605	0.01	(1.01)	0.995
4 v 4.5	36.48	(29.64)	-1.48	(1.53)	0.334	-1.08	(1.13)	0.339
4.5 v 5	32.27	(24.33)	-1.57	(1.70)	0.355	-0.35	(1.40)	0.802

Notes: Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors are clustered at the home health agency level.

Appendix Table 3: Measures included in the first six quarters of the home health star ratings program.

Measure Type	Measure and brief description
Process of Care	***Timely initiation of care*** Percentage of home health quality episodes in which the start or resumption of care date was either on the physician specified date or within 2 days of the referral date or inpatient discharge date, whichever is later.
Process of Care	***Drug education on all medications provided to patient/caregiver*** Percentage of home health quality episodes during which patient/caregiver was instructed on how to monitor the effectiveness of drug therapy, how to recognize potential adverse effects, and how and when to report problems.
Process of Care	***Influenza immunization received for the current flu season*** Percentage of home health quality episodes during which patients were offered and refused influenza immunization for the current flu season.
Health Outcome	***Improvement in ambulation*** Percentage of home health quality episodes during which the patient improved in ability to ambulate.
Health Outcome	***Improvement in bed transferring*** Percentage of home health quality episodes during which the patient improved in ability to get in and out of bed.
Health Outcome	***Improvement in bathing*** Percentage of home health quality episodes during which the patient got better at bathing self.
Health Outcome	***Improvement in pain*** Percentage of home health quality episodes during which the patient's frequency of pain when moving around improved.
Health Outcome	***Improvement in shortness of breath*** Percentage of home health quality episodes during which the patient became less short of breath or dyspneic.
Health Outcome	***Acute care hospitalization*** Percentage of home health stays in which patients were admitted to an acute care hospital during the 60 days following the start of the home health stay.

Appendix Table 4. Density tests using the Frandsen manipulation and binomial tests.

	Frandsen density test				Binomial test		
	Support points, No.	Delta	*k*	*p* value	n < 0	n ≥ 0	*p* value
Pooled thresholds	4	0.50	0.138	1.000	5,184	5,217	0.754
By threshold							
1 v 1.5	4	0.50	0.138	0.370	7	10	0.629
1.5 v 2	6	0.33	0.058	0.851	177	232	0.008
2 v 2.5	6	0.33	0.058	0.583	570	635	0.065
2.5 v 3	6	0.33	0.058	0.986	1,028	1,048	0.677
3 v 3.5	6	0.33	0.058	0.826	1,271	1,314	0.409
3.5 v 4	6	0.33	0.058	0.729	1,095	1,048	0.320
4 v 4.5	6	0.33	0.058	0.874	701	652	0.192
4.5 v 5	6	0.33	0.058	0.715	335	278	0.024

Notes: Binomial tests are conducted on the two consecutive mass points closest to the cutoff [−0.028, 0.028]

Appendix Table 5: Functional form assessment of regression discontinuity design for all thresholds, 1 vs. 1.5 stars, and 1.5 vs. 2 stars.

	Covariates	Coeff. (SE)		AIC
Pooled Thresholds (N = 8,806)				
Linear	No	0.25	(0.45)	371,340
	Yes	0.03	(0.19)	298,516
Linear Interaction	No	0.25	(0.45)	371,341
	Yes	0.03	(0.19)	298,518
Quadratic	No	0.25	(0.45)	371,341
	Yes	0.03	(0.19)	298,518
Quadratic Interaction	No	0.44	(0.78)	371,345
	Yes	-0.25	(0.32)	298,520
1 v 1.5 stars (N = 154)				
Linear	No	0.06	(4.51)	2,053
	Yes	1.32	(3.54)	1,857
Linear Interaction	No	7.31	(5.47)	2,047
	Yes	3.64	(3.31)	1,857
Quadratic	No	8.37	(6.10)	2,048
	Yes	4.01	(3.62)	1,857
Quadratic Interaction	No	-1.70	(5.04)	2,050
	Yes	-4.19	(5.43)	1,859
1.5 v 2 stars (N = 1,011)				
Linear	No	0.06	(1.43)	19,876
	Yes	0.20	(0.73)	16,884
Linear Interaction	No	0.03	(1.38)	19,878
	Yes	0.27	(0.73)	16,886
Quadratic	No	0.10	(1.40)	19,878
	Yes	0.26	(0.74)	16,886
Quadratic Interaction	No	2.40	(2.08)	19,880
	Yes	1.55	(1.17)	16,888

Notes: Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors are clustered at the home health agency level.

Appendix Table 6: Functional form assessment of regression discontinuity design 2 vs. 2.5 stars, 2.5 vs. 3 stars, and 3 vs. 3.5 stars.

	Covariates	Coeff.	(SE)	AIC
2 v 2.5 stars (N = 2,459)				
Linear	No	1.16	(1.06)	56,255
	Yes	0.06	(0.51)	46,642
Linear Interaction	No	1.29	(1.06)	56,256
	Yes	0.06	(0.51)	46,644
Quadratic	No	1.28	(1.06)	56,256
	Yes	0.07	(0.51)	46,644
Quadratic Interaction	No	1.02	(1.71)	56,259
	Yes	-0.17	(0.82)	46,647
2.5 v 3 stars (N = 3,887)				
Linear	No	0.61	(0.92)	93,891
	Yes	0.24	(0.42)	75,950
Linear Interaction	No	0.53	(0.92)	93,891
	Yes	0.24	(0.42)	75,952
Quadratic	No	0.50	(0.92)	93,890
	Yes	0.24	(0.42)	75,952
Quadratic Interaction	No	-0.18	(1.59)	93,891
	Yes	-0.42	(0.71)	75,954
3 v 3.5 stars (N = 4,591)				
Linear	No	-1.02	(0.88)	115,889
	Yes	-0.54	(0.34)	92,508
Linear Interaction	No	-1.03	(0.88)	115,887
	Yes	-0.54	(0.34)	92,510
Quadratic	No	-1.03	(0.88)	115,887
	Yes	-0.54	(0.34)	92,509
Quadratic Interaction	No	0.82	(1.45)	115,887
	Yes	-0.33	(0.53)	92,512

Notes: Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors are clustered at the home health agency level.

Appendix Table 7: Functional form assessment of regression discontinuity design 3.5 vs. 4 stars, 4 vs. 4.5 stars, and 4.5 vs. 5 stars.

	Covariates	Coeff.	(SE)	AIC
3.5 v 4 (N = 4,010)				
Linear	No	-0.63	(0.85)	96,453
	Yes	0.30	(0.35)	76,960
Linear Interaction	No	-0.60	(0.85)	96,455
	Yes	0.31	(0.35)	76,962
Quadratic	No	-0.60	(0.85)	96,455
	Yes	0.30	(0.35)	76,962
Quadratic Interaction	No	0.15	(1.40)	96,459
	Yes	0.48	(0.59)	76,965
4 v 4.5 stars (N = 2,520)				
Linear	No	0.37	(0.94)	59,768
	Yes	0.40	(0.39)	47,231
Linear Interaction	No	0.45	(0.96)	59,770
	Yes	0.40	(0.40)	47,233
Quadratic	No	0.43	(0.96)	59,770
	Yes	0.40	(0.40)	47,233
Quadratic Interaction	No	0.65	(1.57)	59,773
	Yes	-0.20	(0.61)	47,235
4.5 v 5 stars (N = 1,155)				
Linear	No	1.03	(1.33)	27,127
	Yes	-0.24	(0.53)	21,663
Linear Interaction	No	1.43	(1.39)	27,127
	Yes	-0.09	(0.57)	21,664
Quadratic	No	1.47	(1.40)	27,127
	Yes	-0.09	(0.57)	21,664
Quadratic Interaction	No	1.71	(2.02)	27,131
	Yes	-0.42	(0.78)	21,667

Notes: Covariates include agency organizational characteristics (agency age, chain affiliation, for-profit status), pre-star ratings Medicare patient characteristics from 1/2015 to 6/2015 (total patient count, percent discharged from an inpatient institution, percent female, percent white, percent black, percent Hispanic, average age, percent that were fee-for-service enrollees, percent that were Medicare Advantage enrollees, percent that were Medicaid enrollees, percent that were dually enrolled in Medicaid and Medicare, number of new patients, share of new patients), and star rating release dummy variables. Standard errors are clustered at the home health agency level.

Appendix Table 8: Local randomization analysis using pre-treatment characteristics on the pooled threshold sample

	Mean (SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	17.96	(12.69)	5,184	5,217	0.48	0.062
For profit	0.77	(0.42)	5,184	5,217	-0.01	0.182
Part of chain	0.30	(0.46)	5,065	5,103	-0.01	0.119
Patient characteristics						
New patient share	15.96	(21.37)	5,149	5,187	0.35	0.391
New patients, no.	35.48	(50.91)	5,149	5,187	1.20	0.272
All patients, No.	359.14	(625.93)	5,149	5,187	28.17	0.045
Percent Medicare FFS	80.00	(20.41)	5,149	5,187	-0.31	0.467
Percent Medicare Advantage	15.05	(17.21)	5,149	5,187	0.16	0.660
Percent Medicaid	4.95	(9.03)	5,149	5,187	0.15	0.394
Percent dually eligible	40.02	(25.46)	5,149	5,187	0.38	0.447
Percent post acute	53.39	(23.16)	5,149	5,187	-0.25	0.580
Mean age	74.92	(4.54)	5,149	5,187	-0.03	0.717
Percent female	61.77	(7.07)	5,149	5,187	0.13	0.350
Percent White	69.06	(29.79)	5,149	5,187	-0.43	0.466
Percent Black	15.38	(21.39)	5,149	5,187	0.57	0.180
Percent Hispanic	11.91	(22.00)	5,149	5,187	-0.34	0.427

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 9: Local randomization analysis using pre-treatment characteristics on the 1 v 1.5 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	8.54	(2.76)	7	10	-1.64	0.293
For profit	1.00	(0.00)	7	10	0.00	1.000
Part of chain	0.00	(0.00)	6	10	0.00	1.000
Patient characteristics						
New patient share	4.12	(3.41)	7	10	-1.52	0.254
New patients, no.	16.43	(14.37)	7	10	-1.73	0.789
All patients, No.	108.71	(83.65)	7	10	-12.51	0.764
Percent Medicare FFS	70.45	(26.04)	7	10	11.91	0.332
Percent Medicare Advantage	14.85	(8.39)	7	10	-8.47	0.086
Percent Medicaid	14.70	(25.19)	7	10	-3.44	0.802
Percent dually eligible	75.85	(17.74)	7	10	-17.93	0.233
Percent post acute	41.89	(22.88)	7	10	-11.74	0.258
Mean age	70.06	(6.44)	7	10	1.56	0.534
Percent female	60.46	(6.57)	7	10	-1.38	0.687
Percent White	20.04	(12.70)	7	10	13.86	0.180
Percent Black	30.39	(31.47)	7	10	-4.04	0.766
Percent Hispanic	45.77	(42.86)	7	10	-10.53	0.602

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 10: Local randomization analysis of pre-treatment characteristics on the 1.5 v 2 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	13.05	(10.19)	177	232	-0.42	0.656
For profit	0.91	(0.29)	177	232	0.00	1.000
Part of chain	0.11	(0.31)	171	227	-0.02	0.546
Patient characteristics						
New patient share	7.42	(12.78)	174	232	1.20	0.422
New patients, no.	17.89	(18.26)	174	232	0.26	0.914
All patients, No.	144.90	(404.98)	174	232	-5.19	0.886
Percent Medicare FFS	81.48	(21.23)	174	232	0.22	0.922
Percent Medicare Advantage	13.41	(16.05)	174	232	-1.50	0.356
Percent Medicaid	5.11	(11.22)	174	232	1.28	0.303
Percent dually eligible	50.73	(25.43)	174	232	2.71	0.336
Percent post acute	40.28	(19.50)	174	232	-2.62	0.227
Mean age	73.01	(4.65)	174	232	-0.46	0.349
Percent female	61.25	(8.79)	174	232	0.38	0.622
Percent White	48.49	(32.90)	174	232	-3.95	0.256
Percent Black	20.73	(27.03)	174	232	2.56	0.355
Percent Hispanic	28.39	(34.59)	174	232	1.24	0.731

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 11: Local randomization analysis of pre-treatment characteristics on the 2 v 2.5 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	14.89	(11.16)	570	635	1.59	0.019
For profit	0.84	(0.37)	570	635	-0.02	0.446
Part of chain	0.17	(0.38)	555	615	0.00	0.893
Patient characteristics						
New patient share	10.64	(16.17)	561	631	1.87	0.079
New patients, no.	22.34	(28.92)	561	631	0.33	0.842
All patients, No.	198.90	(366.29)	561	631	37.79	0.131
Percent Medicare FFS	78.91	(22.69)	561	631	-2.33	0.096
Percent Medicare Advantage	15.09	(18.17)	561	631	0.59	0.596
Percent Medicaid	6.00	(11.40)	561	631	1.73	0.018
Percent dually eligible	46.30	(27.06)	561	631	0.70	0.633
Percent post acute	45.61	(22.79)	561	631	2.47	0.051
Mean age	73.41	(4.98)	561	631	-0.06	0.826
Percent female	61.66	(7.57)	561	631	0.14	0.750
Percent White	55.45	(32.21)	561	631	4.07	0.030
Percent Black	20.95	(25.15)	561	631	-0.64	0.637
Percent Hispanic	20.51	(29.32)	561	631	-4.29	0.008

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 12: Local randomization analysis of pre-treatment characteristics on the 2.5 v 3 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	18.29	(12.98)	1,028	1,048	1.39	0.016
For profit	0.75	(0.43)	1,028	1,048	-0.01	0.730
Part of chain	0.25	(0.43)	996	1,019	0.02	0.365
Patient characteristics						
New patient share	16.45	(22.07)	1,020	1,038	0.71	0.476
New patients, no.	29.24	(42.80)	1,020	1,038	1.12	0.539
All patients, No.	286.67	(519.68)	1,020	1,038	47.76	0.062
Percent Medicare FFS	79.38	(20.76)	1,020	1,038	-0.80	0.379
Percent Medicare Advantage	13.80	(16.14)	1,020	1,038	1.44	0.055
Percent Medicaid	6.82	(11.44)	1,020	1,038	-0.64	0.192
Percent dually eligible	42.64	(24.83)	1,020	1,038	-2.07	0.057
Percent post acute	52.47	(22.60)	1,020	1,038	1.96	0.046
Mean age	74.23	(4.90)	1,020	1,038	0.37	0.090
Percent female	61.73	(7.51)	1,020	1,038	0.11	0.766
Percent White	67.56	(29.49)	1,020	1,038	1.80	0.157
Percent Black	17.44	(22.93)	1,020	1,038	-1.01	0.302
Percent Hispanic	11.38	(20.42)	1,020	1,038	-0.49	0.568

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 13: Local randomization analysis of pre-treatment characteristics on the 3 v 3.5 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	19.28	(12.55)	1,271	1,314	0.64	0.204
For profit	0.75	(0.43)	1,271	1,314	-0.03	0.070
Part of chain	0.31	(0.46)	1,250	1,294	0.02	0.315
Patient characteristics						
New patient share	18.95	(24.13)	1,263	1,307	0.42	0.672
New patients, no.	36.94	(55.40)	1,263	1,307	2.44	0.290
All patients, No.	385.71	(601.75)	1,263	1,307	62.36	0.030
Percent Medicare FFS	78.21	(20.36)	1,263	1,307	-0.10	0.892
Percent Medicare Advantage	16.91	(17.94)	1,263	1,307	0.01	0.984
Percent Medicaid	4.88	(8.38)	1,263	1,307	0.09	0.798
Percent dually eligible	37.91	(23.71)	1,263	1,307	-0.93	0.304
Percent post acute	56.48	(20.93)	1,263	1,307	1.08	0.198
Mean age	75.21	(4.12)	1,263	1,307	0.17	0.273
Percent female	61.71	(7.09)	1,263	1,307	0.63	0.022
Percent White	72.99	(26.91)	1,263	1,307	0.94	0.359
Percent Black	14.18	(19.34)	1,263	1,307	0.54	0.480
Percent Hispanic	9.15	(17.89)	1,263	1,307	-1.10	0.115

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 14: Local randomization analysis of pre-treatment characteristics on the 3.5 v 4 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	19.56	(13.46)	1,095	1,048	-0.06	0.916
For profit	0.72	(0.45)	1,095	1,048	0.00	0.869
Part of chain	0.39	(0.49)	1,066	1,022	-0.06	0.006
Patient characteristics						
New patient share	17.98	(21.56)	1,092	1,042	-0.76	0.412
New patients, no.	45.12	(64.93)	1,092	1,042	2.94	0.381
All patients, No.	502.58	(886.10)	1,092	1,042	7.50	0.846
Percent Medicare FFS	79.98	(19.75)	1,092	1,042	-0.42	0.621
Percent Medicare Advantage	15.70	(16.93)	1,092	1,042	0.71	0.343
Percent Medicaid	4.31	(7.12)	1,092	1,042	-0.29	0.339
Percent dually eligible	35.65	(23.98)	1,092	1,042	1.35	0.205
Percent post acute	57.86	(21.94)	1,092	1,042	-1.18	0.224
Mean age	75.55	(4.18)	1,092	1,042	-0.25	0.188
Percent female	62.07	(6.09)	1,092	1,042	-0.24	0.398
Percent White	74.74	(26.21)	1,092	1,042	-0.72	0.549
Percent Black	12.93	(18.15)	1,092	1,042	0.48	0.536
Percent Hispanic	8.53	(17.27)	1,092	1,042	0.20	0.798

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 15: Local randomization analysis of pre-treatment characteristics on the 4 v 4.5 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	17.82	(12.90)	701	652	-0.71	0.306
For profit	0.75	(0.44)	701	652	0.02	0.516
Part of chain	0.38	(0.49)	690	641	-0.05	0.048
Patient characteristics						
New patient share	14.94	(19.75)	698	650	0.22	0.844
New patients, no.	42.12	(48.83)	698	650	2.40	0.450
All patients, No.	408.36	(528.04)	698	650	45.31	0.267
Percent Medicare FFS	82.36	(19.17)	698	650	1.50	0.149
Percent Medicare Advantage	14.26	(17.50)	698	650	-1.48	0.113
Percent Medicaid	3.38	(5.89)	698	650	-0.02	0.969
Percent dually eligible	36.55	(25.62)	698	650	3.72	0.009
Percent post acute	54.99	(25.21)	698	650	-3.94	0.004
Mean age	75.71	(4.18)	698	650	0.07	0.754
Percent female	61.44	(6.63)	698	650	0.05	0.884
Percent White	73.85	(28.79)	698	650	-4.58	0.004
Percent Black	12.64	(19.72)	698	650	1.40	0.204
Percent Hispanic	10.36	(20.74)	698	650	1.87	0.112

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 16 Local randomization analysis of pre-treatment characteristics on the 4.5 v 5 threshold sample

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Agency characteristics						
Operational years	15.00	(10.85)	335	278	0.52	0.557
For profit	0.86	(0.35)	335	278	-0.04	0.184
Part of chain	0.26	(0.44)	331	275	0.03	0.438
Patient characteristics						
New patient share	12.27	(18.59)	334	277	1.32	0.399
New patients, no.	35.30	(38.25)	334	277	-0.60	0.822
All patients, No.	294.12	(476.47)	334	277	-51.96	0.095
Percent Medicare FFS	85.08	(18.34)	334	277	2.28	0.116
Percent Medicare Advantage	12.18	(16.13)	334	277	-1.90	0.137
Percent Medicaid	2.74	(5.70)	334	277	-0.38	0.386
Percent dually eligible	44.64	(29.49)	334	277	-2.26	0.346
Percent post acute	46.73	(26.99)	334	277	-2.25	0.321
Mean age	75.87	(4.68)	334	277	-0.27	0.500
Percent female	62.35	(7.66)	334	277	-0.46	0.476
Percent White	64.90	(34.60)	334	277	-3.43	0.243
Percent Black	14.92	(23.85)	334	277	2.64	0.200
Percent Hispanic	14.48	(25.44)	334	277	-1.22	0.541

Notes: Pre-treatment characteristics were from January to June 2015. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on pre-treatment covariates. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 17: Local randomization analysis of one more half star on new patient shares

	Mean (SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Pooled thresholds	17.44	(21.83)	5,184	5,217	0.32	0.453
By threshold						
1 v 1.5	6.35	(4.79)	7	10	-1.94	0.266
1.5 v 2	8.85	(12.59)	177	232	1.82	0.242
2 v 2.5	12.54	(18.53)	570	635	1.89	0.096
2.5 v 3	18.33	(23.29)	1,028	1,048	0.55	0.587
3 v 3.5	20.23	(24.05)	1,271	1,314	0.08	0.913
3.5 v 4	19.34	(22.02)	1,095	1,048	-0.48	0.616
4 v 4.5	16.15	(19.41)	701	652	-0.04	0.955
4.5 v 5	13.66	(18.42)	335	278	1.51	0.340

Notes: Analysis of new Medicare fee-for-service market shares among agencies that received a star rating from July 2015−December 2016. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on outcomes. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 18: Local randomization analysis of one more half star on new patient shares over time

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
By release						
7/2015	17.25	(21.34)	785	899	2.41	0.026
10/2015	17.76	(20.92)	897	818	0.89	0.391
1/2016	17.34	(22.59)	920	866	0.51	0.632
4/2016	16.76	(20.86)	830	857	-0.17	0.849
7/2016	17.04	(21.96)	882	925	-1.78	0.071
10/2016	18.43	(23.14)	870	852	0.27	0.792

Notes: Analysis of new Medicare fee-for-service market shares among agencies that received a star rating from July 2015−December 2016. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on outcomes. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 19: Local randomization analysis of one more half star on new patient shares among patients with varying likelihood to shop for home health services

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Patient Source						
Post acute	23.80	(23.28)	4,662	4,702	0.44	0.371
Community entry	22.61	(24.70)	4,940	4,958	0.07	0.890
ZIP HHI – Firms						
Tertile 1	6.57	(7.52)	19,734	18,892	-0.11	0.164
Tertile 2	16.70	(16.58)	17,943	18,703	-0.11	0.502
Tertile 3	41.52	(32.02)	17,443	19,021	0.35	0.285
ZIP HHI – Star types						
Tertile 1	8.23	(11.23)	19,296	18,777	0.18	0.113
Tertile 2	18.24	(19.82)	18,472	18,875	0.05	0.818
Tertile 3	37.91	(32.02)	17,352	18,964	0.36	0.278

Notes: HHI = Herfindahl-Hirschman Index. Tertile 1 = low HHI. Analysis of new Medicare fee-for-service market shares among agencies that received a star rating from July 2015−December 2016. *p* values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on outcomes. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

Appendix Table 20: Local randomization analysis of one more half star on mean income among ZIP codes served

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	p value
Market income (dollars)						
Pooled	41,530	(9,604)	5,184	5,217	147	0.426
1 v 1.5	32,849	(6,049)	7	10	3,131	0.382
1.5 v 2	37,763	(8,738)	177	232	-268	0.754
2 v 2.5	39,020	(8,229)	570	635	464	0.336
2.5 v 3	40,230	(8,436)	1,028	1,048	870	0.020
3 v 3.5	41,928	(9,178)	1,271	1,314	202	0.571
3.5 v 4	43,257	(10,757)	1,095	1,048	-74	0.867
4 v 4.5	43,214	(10,364)	701	652	-277	0.623
4.5 v 5	41,278	(9,514)	335	278	502	0.522

Notes: Mean income averages across ZIP codes served by each agency and focuses on the median ZIP-code level income of persons 65 years or older. Analysis of agencies that received a star rating from July 2015−December 2016. p values correspond to the Fisherian sharp null hypothesis that one more half star has no effect on outcomes. All inferences use the two mass points closest to the cutoff [−0.028, 0.028].

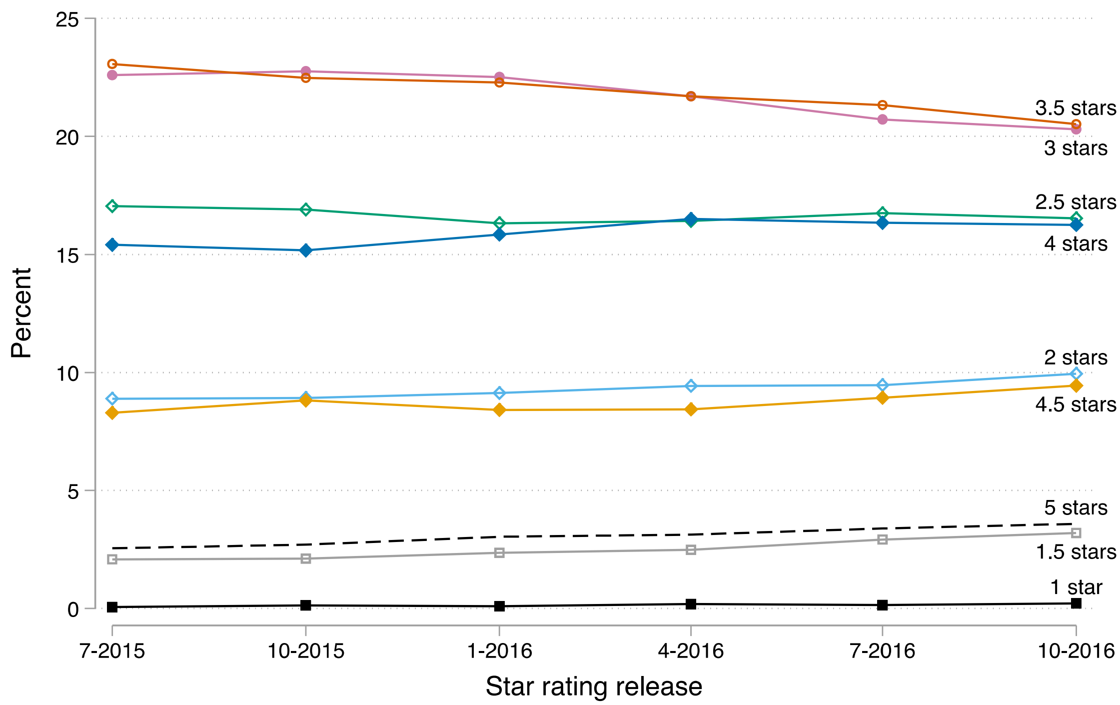
Appendix Table 21: Local randomization analysis of one more half star on low-profit margin and Medicare-Medicaid dually eligible patient share in ZIP codes with varying home health care worker availability

	Mean(SD), control		n < 0	n ≥ 0	Difference in means	*p* value
Fewer home health workers (≤18 per 1,000 persons 65+)						
1.5 v 2						
Low profit	5.20	(5.21)	108	198	2.26	0.039
Duals	10.55	(13.69)	110	198	-0.36	0.822
2 v 2.5						
Low profit	10.44	(13.54)	360	441	7.72	0.000
Duals	13.58	(17.23)	360	447	10.58	0.000
2.5 v 3						
Low profit	22.66	(24.72)	1,277	1,314	-4.71	0.000
Duals	26.11	(24.92)	1,279	1,319	-4.36	0.000
3 v 3.5						
Low profit	19.46	(21.51)	2,665	2,365	1.51	0.020
Duals	21.93	(22.42)	2,671	2,374	1.29	0.058
3.5 v 4						
Low profit	21.62	(23.06)	2,311	2,345	-0.19	0.773
Duals	22.41	(23.63)	2,313	2,347	-0.21	0.760
4 v 4.5						
Low profit	20.22	(21.63)	1,497	1,513	3.25	0.000
Duals	18.85	(20.76)	1,498	1,520	4.67	0.000
4.5 v 5						
Low profit	17.92	(21.37)	478	378	3.52	0.031
Duals	15.75	(16.93)	479	378	5.75	0.000
More home health workers (>27 per 1,000 persons 65+)						
1.5 v 2						
Low profit	4.26	(6.48)	121	172	1.34	0.213
Duals	7.77	(8.80)	121	173	0.41	0.780
2 v 2.5						
Low profit	6.16	(11.24)	855	885	1.66	0.006
Duals	7.54	(11.16)	855	885	1.55	0.014
2.5 v 3						
Low profit	12.99	(17.91)	1,803	1,944	-2.08	0.000
Duals	14.22	(18.28)	1,807	1,948	-0.56	0.328
3 v 3.5						
Low profit	17.11	(20.74)	2,016	2,966	-0.69	0.227
Duals	17.04	(20.46)	2,016	2,969	0.47	0.438
3.5 v 4						
Low profit	20.23	(21.94)	2,187	2,522	-0.17	0.792
Duals	18.04	(22.10)	2,188	2,522	-0.58	0.352
4 v 4.5						
Low profit	14.65	(18.40)	1,087	1,312	-1.28	0.073
Duals	12.18	(16.97)	1,087	1,312	0.65	0.316
4.5 v 5						
Low profit	13.72	(20.69)	369	258	-2.83	0.088
Duals	12.10	(19.49)	369	259	-2.98	0.037

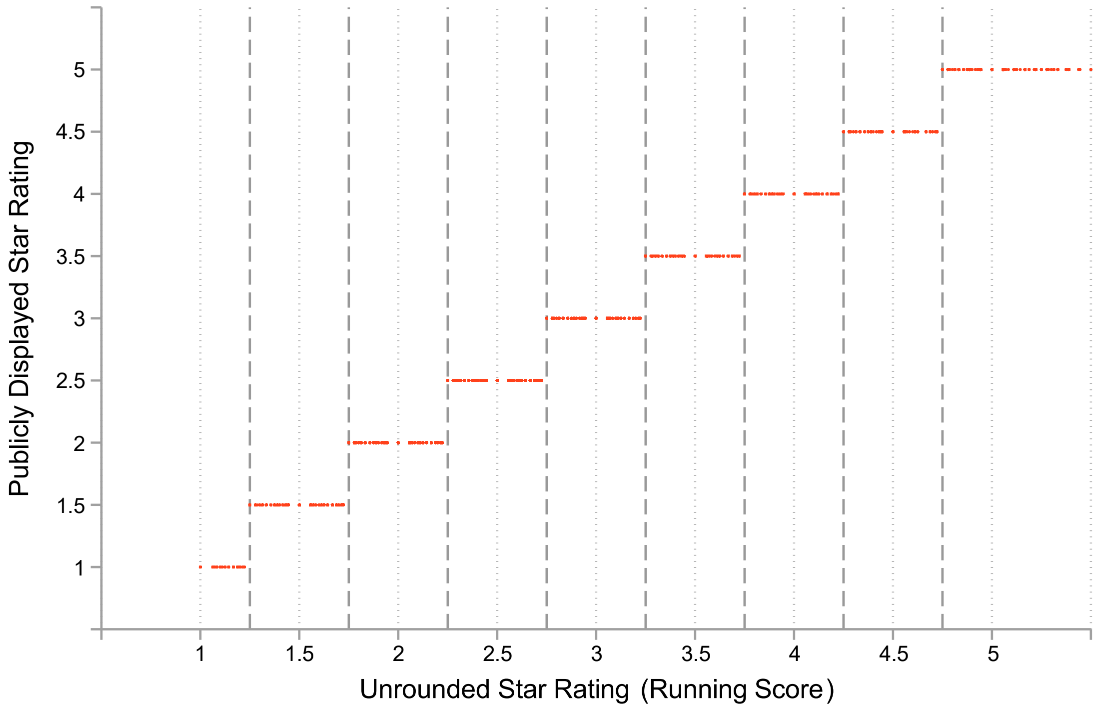
Notes: 1 vs 1.5 stars were excluded from individual threshold analyses due to small sample sizes. Fewer home health workers correspond to the bottom tertile of worker to population ratio and More home health workers correspond to the top tertile. Low-profit margin patients are individuals with clinical characteristics associated with lower profit margins than other Medicare fee-for-service enrollees patients on average, which includes people with poor control of clinical conditions (10 percent lower), with traumatic wounds or ulcers (20 percent lower), with significant bathing needs (20 percent lower), have overall high risk (20 percent lower), and recipients of intravenous therapy or parenteral nutrition at home (15 percent lower).

III. Appendix Figures

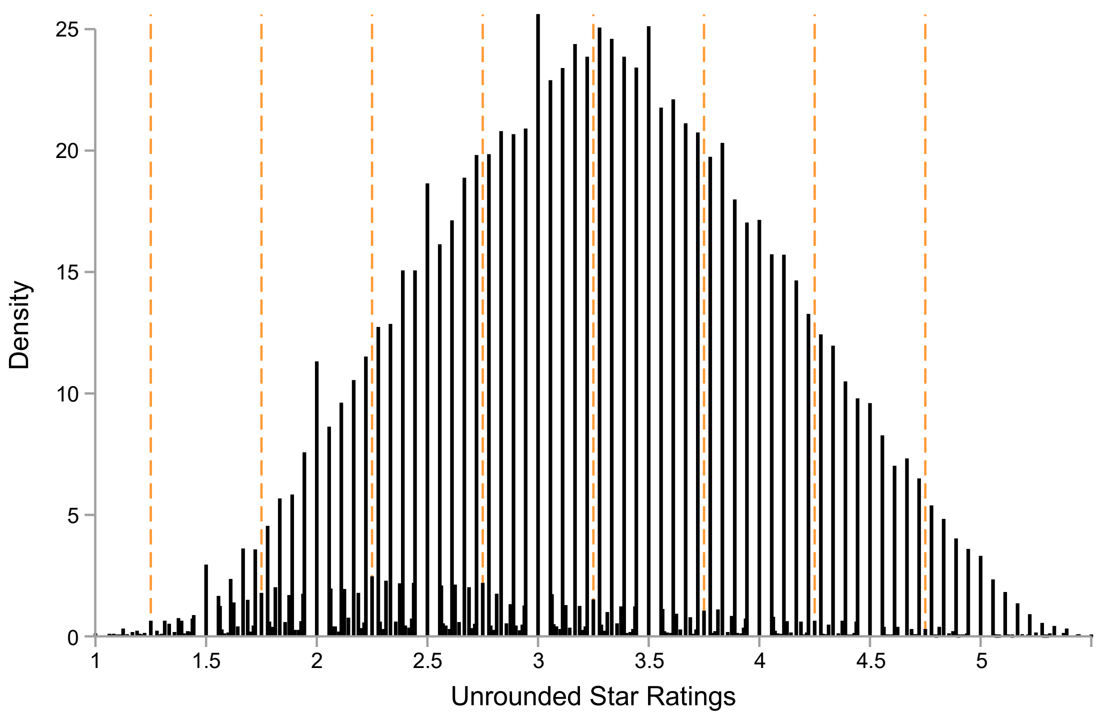
Appendix Figure 1: Shares of agencies by star rating over time.



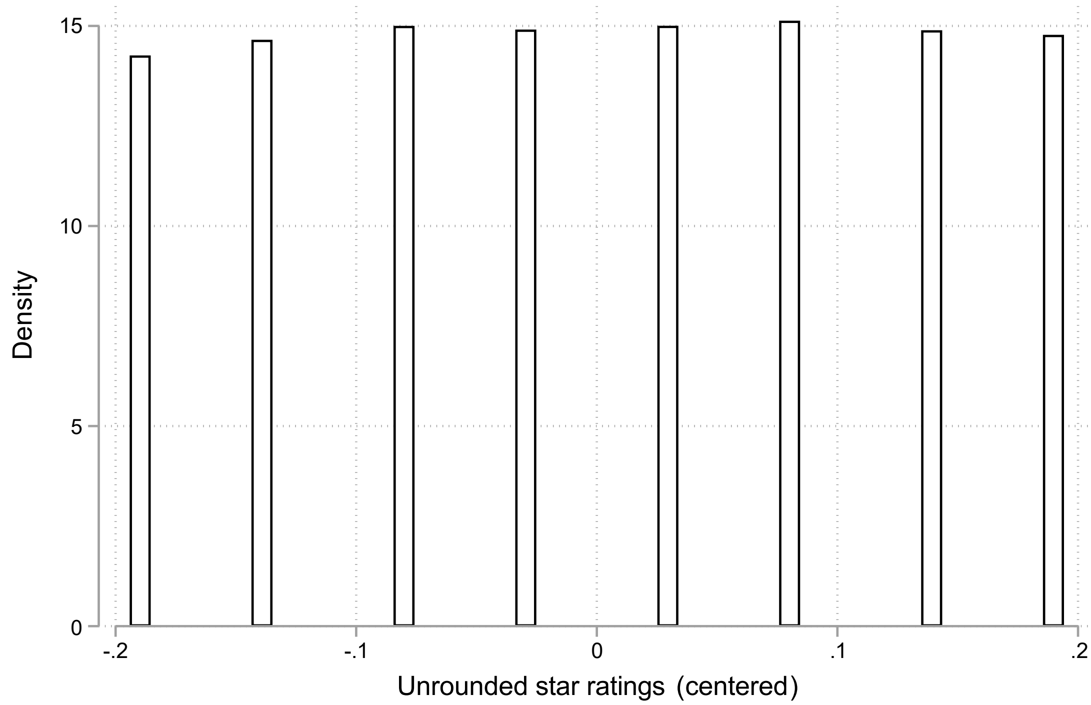
Appendix Figure 2: Relationship between unrounded star ratings and star ratings in the first six quarters of the Home Health Star Ratings program.



Appendix Figure 3: Distribution of unrounded star ratings in the first six quarters of the home health star ratings program.



Appendix Figure 4: Histogram of running variable, unrounded star ratings.



Note: In pooled threshold analyses, observations were removed if they had running variables equal to ± 0.25

Appendix Figure 5: Google trends data for searches on "home health star rating" from January 2014 to January 2020.