Evaluating Nursing Home Quality Measures: Evidence from Nursing Home Closures

Amitabh Chandra, Timothy J. Layton, Connie Xu

Abstract: Poor quality of care in nursing homes has been a longstanding public policy concern. Moreover, the industry has faced major criticism regarding the reliability of reported quality measures that often guide policies and are also used by patients to select nursing homes. A main concern with these observational quality estimates is whether they provide an unbiased and informative estimate of nursing homes' causal impacts on patient health. In this project, we test for bias in these measures using a quasi-experimental design based on the reallocation of patients to alternative facilities following a nursing home closure. The exogenous variation in market share resulting from a closure allows us to validate quality measures by comparing how health outcomes change differentially in markets where a facility with high measured quality closes versus markets where a facility with low measured quality closes. Rich, longitudinal patient health data allows us to determine not only whether quality measures change differentially in these markets but also whether various dimensions of patient health, including physical health, mental health, cognitive capacity, and activities of daily living, change differentially when high- versus low-quality nursing homes close. Beyond validating quality measures, we can use this variation to estimate the causal nursing home effects on quality. This allows us to further investigate the extent to which racial health disparities among nursing home residents are driven by patients from different racial and ethnic groups utilizing nursing homes that differ in their causal effects on patient health. Ultimately, these results will provide high-quality actionable estimates of nursing home effects on patient health, the extent to which those effects are correlated with quality measures, the extent to which those effects are correlated with other observable nursing home characteristics, and the extent to which heterogeneity in nursing home effects on patient health are responsible for racial health disparities.

Significance: Quality measures are used in many healthcare settings and play a central role in policymaking. These measures have far-reaching consequences as they help consumers select better providers, incentivize providers to invest in quality, and guide regulatory agencies to make policies encouraging higher quality (such as pay-for-performance). However, the wide variation in the quality of care has made these measures subject to much debate. A primary concern with using these measures is that risk adjusting for observed patient characteristics does not sufficiently control for patient composition, resulting in quality measures reflecting a provider's patient mix rather than the provider's effect on the quality of patient care and patient health. Given the significant implications and wide use of these measures, it is therefore essential to validate whether existing observational quality measures can actually predict the causal effect of providers on patient outcomes.

The nursing home industry is a setting where the quality of care has remained infamously low despite significant efforts over the past three decades to improve care (National Academies of Sciences, Engineering, and Medicine 2022; Grabowski, Rantz, and Travers 2022). Patients struggle to differentiate between high and low quality nursing homes (Grabowski and Town 2011; Werner et al. 2012). Further, the extent to which measured "quality" actually matters for patient health or satisfaction is up for debate. While the Centers for Medicare and Medicaid Services (CMS) has tried to measure and publicize measures of nursing home quality via the public report card tool it introduced in 2002, it is clear that existing measures of nursing home quality are subject to rampant manipulation by facilities (Thomas 2014; Perraillon, Brauner, and Konetzka 2019; Silver-Greenberg and Gebeloff 2021). A lot of the data submitted were self-reported by facilities and often incorrect – making nursing homes seem safer than they actually are. COVID-19 further exposed flaws in nursing home quality measurement as there was little if any correlation between existing quality measures and a facility's ability to handle the pandemic (Abrams et al. 2020; Konetzka 2020; Chen, Chevalier, and Long 2021). When existing quality measures provide a badly distorted representation of care in nursing homes, the consequences for patients can be harmful and dangerous. They can not only lead to people choosing poor-quality nursing homes, but they can also result in an equilibrium where all nursing homes are relatively low-quality because there is little incentive to make costly investments in quality and patient health. More often than not, these consequences hurt disadvantaged populations the most – leading to greater inequities and disparities in nursing home care (Werner et al. 2012; Konetzka et al. 2015; Sharma et al. 2020).

Previous work on validating quality measures in healthcare has focused on hospitals and Medicare Advantage Plans (Doyle, Graves, and Gruber 2019; Hull 2020; Abaluck et al. 2021). A growing literature has begun to estimate nursing home quality. In particular, Einav, Finkelstein, and Mahoney 2022 develop a stylized model that estimates the value-added of nursing homes for Medicare patients and finds substantial variation in quality within a geographic market. Our research adds to the existing literature by using a quasi-experimental design based on nursing home closures to test whether differences in nursing home quality capture the causal impacts of nursing homes on quality and patient health or if they are biased by patient sorting. Moreover, a unique feature of the nursing home industry is the availability of rich patient health data from federally mandated health assessments. This provides us with the ideal setting to measure how heterogeneity in nursing home quality impacts not only mortality rates but also a wide range of patient health outcomes such as mental health, daily functioning, cognitive capacity, and much more. Finally, racial segregation and health disparities have been well documented across nursing home facilities (Smith et al. 2007; Konetzka et al. 2015; Rahman and Foster 2015). Our research will fill a critical gap in the literature by understanding how differences in the causal effects of nursing homes on patient health contribute to these disparities. Put together, our project involves the following three specific aims:

- *Aim 1:* Validate a set of standard outcomes-based quality measures based on how well they predict the causal nursing home effects on these measures. Specifically, we will leverage quasi-random patient reallocation to alternative nursing

homes following nursing home closures of different measured quality to test whether these measures reflect unobserved sorting or causal nursing home effects.

- *Aim 2:* Test whether nursing home quality measures predict the causal nursing home effects on patient health outcomes using rich patient health data on every nursing home resident in the United States. This unique feature of our setting allows us to holistically measure patient health beyond mortality and capture multiple dimensions of health such as mental health, cognitive capacity, and activities of daily living.
- *Aim 3:* Estimate a model of causal nursing home effects on patient health that allows us to assess the extent to which differences in these causal effects across nursing homes are used differentially by different racial and ethnic groups and contribute to health disparities among disadvantaged populations.

Methodology: This project will validate existing measures of nursing home quality and estimate causal nursing home effects on patient health outcomes. We start with a set of commonly used quality measures such as length of stay, mortality rates, and a subset of the measures used to construct CMS facility star ratings. With a flexible hierarchical empirical bayes model, we derive shrinkage estimates of these measures while also adjusting for observable enrollee characteristics such as age, sex, race, and dual-eligibility status. These estimates reflect the within-market differences in quality measures among observably similar patients. The variation in these estimates, however, reflects both nonrandom patient sorting (on unobservables) and the causal nursing home effects.

To validate whether these observational measures predict causal nursing home effects, we will use a quasi-experimental design based on the reallocation of patients to alternative nursing homes after a nursing home closure. This approach provides us with a source of plausibly exogenous variation in market share that we can leverage to measure the relationship between our observational measures and the true effects of nursing homes. If our observational measures of quality accurately reflect the true variation in causal nursing homes effects, we would expect to see outcomes improve in markets where a facility with low-measured quality closes and worsen in markets where a facility with high-measured quality closes. Letting i index patients, j index geographic markets, and t index calendar time, we will initially estimate event studies showing how patient outcomes, y_{ijt},change differentially in these markets using the following regression specification:

$$y_{ijt} = \gamma_j + \tau_t + \sum_{k \in K} \mu^{g,k} D_{it}^{g,k} + X_{it}'\Gamma + \epsilon_{ijt}$$

where γ_i are market fixed effects, τ_t are calendar time fixed effects, and X_{it} are time-varying patient characteristics. To allow for differential trends by the quality of the closed nursing home, we divide nursing homes into four groups indexed as $g \in \{1,2,3,4\}$ by their quartile of quality. With this classification, D_{ijt}^{g} is then an indicator for whether patient i is in a nursing home of quality group g that closes in relative event time k. Plots of our coefficient $\mu^{g,k}$ will show how patient outcomes change differentially according to closures of varying nursing home quality in our markets.

We then formalize this general quasi-experimental approach of leveraging nursing home closures with an instrumental variables "value-added" framework that builds on existing quality estimation work in educational and healthcare settings (Kane and Staiger 2008; Chetty, Friedman, and Rockoff 2014a; 2014b; Angrist et al. 2016; Abaluck et al. 2021). Our parameter of interest measures the relationship between observational quality and the true nursing home effects. This comes from the coefficient in a regression of patient outcomes on observational quality measures. If patients of different unobserved health sort nonrandomly into nursing homes, standard ordinary least squares estimation will produce biased estimates of the causal relationship between observational quality and outcomes. Thus, we instrument for changes in observational quality using the quasi-random variation in the nursing home choice set that arises from nursing home closures.

For the final part of the project, we will move away from "value-added" models that assess the relationship between quality measures and patient outcomes and towards a more general model of causal nursing home effects on patient outcomes. Specifically, we will estimate a two-way fixed effects model that will generate causal estimates of nursing homes on patient outcomes for all nursing homes in markets where there was at least one closure. This model identifies causal effects both for nursing homes that close *and* the nursing homes in the same market that patients substitute to. We will use these nursing home specific estimates to assess correlates with nursing home causal effects on patient health. We will also use these estimates to assess the extent to which racial disparities in health outcomes are due to patients from different racial and ethnic groups utilizing nursing homes with different causal effects on patient health.

Our research will make use of several CMS datasets spanning 2011 to 2016: the Provider of Services (POS) file, the Minimum Data Set 3.0 (MDS 3.0), the Medicare Denominator file, and the 100% MedPAR file. We focus on residents that can be matched to the Medicare Denominator file, which drops around 13% of mostly private-pay patients. This gives a sample of 11.5 million nursing home residents across almost 16,000 facilities. With the POS file, we identify over 500 nursing home closures during this timeframe. A unique feature of our project is the set of rich patient health data from MDS 3.0. These data come from federally mandated health assessments that capture over 300 health measures. Moreover, these assessments are taken periodically on all patients during their stay – providing us with detailed, longitudinal data to more holistically measure causal nursing home effects on patient health beyond mortality effects. Finally, these data will allow us to see how nursing home quality differentially impacts disadvantaged populations on the different dimensions of health captured by MDS 3.0 and assess the extent to which heterogeneity in quality exacerbates existing racial health disparities.

References:

Abaluck, Jason, Mauricio Caceres Bravo, Peter Hull, and Amanda Starc. 2021. "Mortality Effects and Choice Across Private Health Insurance Plans." *The Quarterly Journal of Economics* 136 (3): 1557–1610.

Abrams, Hannah R., Lacey Loomer, Ashvin Gandhi, and David C. Grabowski. 2020. "Characteristics of U.S. Nursing Homes with COVID-19 Cases." *Journal of the American Geriatrics Society* 68 (8): 1653–56. https://doi.org/10.1111/jgs.16661.

Angrist, Joshua, Peter Hull, Parag Pathak, and Christopher Walters. 2016. "Interpreting Tests of School VAM Validity." *American Economic Review* 106 (5): 388–92.

Chen, M. Keith, Judith A. Chevalier, and Elisa F. Long. 2021. "Nursing Home Staff Networks and COVID-19." *Proceedings of the National Academy of Sciences* 118 (1): e2015455118. https://doi.org/10.1073/pnas.2015455118.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593–2632.

———. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–79.

Doyle, Joseph, John Graves, and Jonathan Gruber. 2019. "Evaluating Measures of Hospital Quality: Evidence from Ambulance Referral Patterns." *Review of Economics and Statistics* 101 (5): 841–52.

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2022. "Producing Health: Measuring Value Added of Nursing Homes."

Grabowski, David, Marilyn Rantz, and Jasmine Travers. 2022. "Report: U.S. Nursing Home Care Is Ineffective, Inefficient, Inequitable, Fragmented, and Unsustainable." *STAT* (blog). April 6, 2022. https://www.statnews.com/2022/04/06/nursing-home-care-ineffective-inefficient-inequitable-fragmented-unsustainable/.

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

Hull, Peter. 2020. "Estimating Hospital Quality with Quasi-Experimental Data." *Working Paper*.

Kane, Thomas J., and Douglas O. Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." National Bureau of Economic Research.

Konetzka, R. Tamara. 2020. "Senate Testimony before the Senate Special Committee on Aging." https://www.aging.senate.gov/imo/media/doc/SCA_Konetzka_05_21_20.pdf.

Konetzka, R. Tamara, David C. Grabowski, Marcelo Coca Perraillon, and Rachel M. Werner. 2015. "Nursing Home 5-Star Rating System Exacerbates Disparities in Quality, by Payer Source." *Health Affairs (Project Hope)* 34 (5): 819–27. https://doi.org/10.1377/hlthaff.2014.1084.

National Academies of Sciences, Engineering, and Medicine. 2022. *The National Imperative to Improve Nursing Home Quality: Honoring Our Commitment to Residents, Families, and Staff*. Washington, D.C.: The National Academies Press. https://doi.org/10.17226/26526.

Perraillon, Marcelo Coca, Daniel J. Brauner, and R. Tamara Konetzka. 2019. "Nursing Home Response to Nursing Home Compare: The Provider Perspective." *Medical Care Research and Review : MCRR* 76 (4): 425–43. https://doi.org/10.1177/1077558717725165.

Rahman, Momotazur, and Andrew D. Foster. 2015. "Racial Segregation and Quality of Care Disparity in US Nursing Homes." *Journal of Health Economics* 39 (January): 1–16. https://doi.org/10.1016/j.jhealeco.2014.09.003.

Sharma, Hari, Marcelo Coca Perraillon, Rachel M. Werner, David C. Grabowski, and R. Tamara Konetzka. 2020. "Medicaid and Nursing Home Choice: Why Do Duals End Up in Low-Quality Facilities?" *Journal of Applied Gerontology* 39 (9): 981–90. https://doi.org/10.1177/0733464819838447.

Silver-Greenberg, Jessica, and Robert Gebeloff. 2021. "Maggots, Rape and Yet Five Stars: How U.S. Ratings of Nursing Homes Mislead the Public." *The New York Times*, March 13, 2021, sec. Business. https://www.nytimes.com/2021/03/13/business/nursing-homes-ratings-medicare-covid.html.

Smith, David Barton, Zhanlian Feng, Mary L. Fennell, Jacqueline S. Zinn, and Vincent Mor. 2007. "Separate And Unequal: Racial Segregation And Disparities In Quality Across U.S. Nursing Homes." *Health Affairs* 26 (5): 1448–58. https://doi.org/10.1377/hlthaff.26.5.1448.

Thomas, Katie. 2014. "Medicare Star Ratings Allow Nursing Homes to Game the System." *The New York Times*, August 25, 2014, sec. Business. https://www.nytimes.com/2014/08/25/business/medicare-star-ratings-allow-nursing-homes-to-game-the-system.html.

Werner, Rachel M., Edward C. Norton, R. Tamara Konetzka, and Daniel Polsky. 2012. “Do Consumers Respond to Publicly Reported Quality Information? Evidence from Nursing Homes.” *Journal of Health Economics* 31 (1): 50–61. https://doi.org/10.1016/j.jhealeco.2012.01.001.