Algorithmic Decision-Making in Health Care: Evidence from Post-Acute Care in Medicare Advantage

Jeffrey Marr*

September 6, 2024

Job Market Paper

Abstract

Health insurers use predictive algorithms to determine the necessary level of care and deny services they deem unnecessary. Using a difference-in-differences design, I study the partnership of a large Medicare Advantage insurer with a firm that uses a predictive algorithm to aid post-acute care coverage decisions. This partnership led to an immediate and sustained 13 percent decline in the length of skilled nursing facility stays. This effect was partially driven by large declines in longer skilled nursing facility stays (over 30 days). Despite reductions in health care use, I don't observe changes in health outcomes following the adoption of the predictive algorithm.

*Johns Hopkins University. Email: jmarr5@jhu.edu. I am grateful to Dan Polsky, Joe Levy, Mark Meiselbach, and Karen Shen for their guidance and support. I thank Yaa Akosa Antwi, Kali Thomas, and Ben Thornburg, as well as seminar participants at the Hopkins Business of Health Initiative and the AcademyHealth Annual Research Meeting, for their helpful comments. I received funding from the Lipitz Center and from the National Institute on Aging, National Institutes of Health (T32AG066576).

1 Introduction

Some health care is expensive but does little to improve health. How do we eliminate this low-value health care spending while preserving care that improves health outcomes? Cost-sharing is the most widely used and studied method for reducing health care use (Pauly, 1968; Manning et al., 1987). Yet, it can be a blunt instrument that reduces not only low-value care but cuts high-value care, potentially worsening health and increasing mortality (Chandra et al., 2024). Can a more targeted approach work better?

Prior authorization, where an insurer reviews and approves care before it happens, is intended to target reductions towards unnecessary care (Brot-Goldberg et al., 2023; Dillender, 2018; Eliason et al., 2021). In practice, doing so is difficult. Medical decision-making is complex and requires time and expertise to distinguish between necessary and unnecessary care. When conducted manually, prior authorization is costly to insurers, difficult to do in the timely manner needed for clinical decision-making, and susceptible to human error (McKinsey & Company, 2021).

Insurers increasingly use algorithms to predict patients' needed level of care and deny coverage beyond what they determine is necessary (Ross and Herman, 2023a). Algorithms may improve prior authorization by quickly processing large amounts of relevant information and making recommendations based on historical data. This may allow insurers to precisely target reductions in health care use. However, relying on an algorithm may miss important details that are not observable to the model, leading to denials of needed care. As a result, policymakers, the press, and the courts have scrutinized insurers' use of algorithms (Centers for Medicare and Medicaid Services, 2023; Chu and Nadler, 2023; Ross and Herman, 2023a).

Predictive algorithms increasingly guide decision-making in a wide array of settings—not only in health care but also in criminal justice and child protective services (Ludwig et al., 2024). For example, in the criminal justice setting, judges use predictive algorithms during sentencing or pre-trial detention hearings to assess the risk of re-offending (Albright, 2014; Sloan et al., 2023; Stevenson, 2018). In that setting, the goal of algorithmic decision-making is to target pre-trial detention or longer sentences to a subset of defendants with a high risk of reoffending. In some ways, it is analogous to the goal of algorithms used by health insurers: to reduce unnecessary health care use while continuing to cover a targeted subset of patients who need more intense care. However, in the criminal justice setting—and in many other settings where predictive algorithms are used—the algorithm is designed to support a single human decision-maker. The situation is more complicated when algorithms are used in prior authorization by health insurers. In the health care setting, algorithms are used to support the decision-making of one agent—the insurer—in monitoring another agent—the provider. Both agents influence the final decision of how much care to provide but have very different objectives. While the insurer may be trying to reduce unnecessary care, the provider may be trying to induce demand.

I study a large health insurer's adoption of a predictive algorithm in its prior authorization process. I examine how health care use changes after the adoption of algorithmic decision-making. I also examine whether health outcomes changed to determine whether the reductions in health care use are sufficiently well-targeted to minimize adverse consequences for patients. I study these questions in the post-acute care setting, where patients receive rehabilitation services after a hospitalization. Post-acute care can be delivered at a skilled nursing facility (SNF), at an inpatient rehabilitation facility (IRF), or through a home health agency. There is substantial 'waste' in the US post-acute care system, where expensive care is provided that does not improve patient outcomes (Chandra et al., 2013; Doyle et al., 2017; McGarry et al., 2021; Regenbogen et al., 2019). One reason for this may be the incentives that post-acute care providers face. For example, Medicare pays SNFs on a per diem basis at rates that are well above their marginal costs (Medicare Payment Advisory Commission, 2024). This creates an incentive to induce demand by unnecessarily extending stays. As a result, reducing unnecessary post-acute care has been a broad focus throughout the health care system (Barnett et al., 2019; Biniek et al., 2019; Huckfeldt et al., 2017; McWilliams et al., 2017). One strategy to reduce post-acute care use has been the widespread adoption of prior authorization by Medicare Advantage (MA) plans (Kaiser Family Foundation, 2024b). Insurers that use prior authorization often contract with an outside firm—including NaviHealth, CareCentrix, and myNexus—to directly manage patients' post-acute care and share in any savings. These firms all use predictive algorithms to aid their decision-making (Chu and Nadler, 2023).

I leverage the partnership of one of these companies, NaviHealth, with a large MA insurer to evaluate the causal effect of algorithmic decision-making on post-acute care use and health outcomes. This publicly announced partnership with Blue Cross Blue Shield of Michigan (BCBS MI) began on June 1, 2019. Using a difference-in-differences design and administrative data, I compare health care use and outcomes for BCBS MI enrollees before and after this change, using traditional Medicare (TM) beneficiaries in Michigan as my control group.

I find that the NaviHealth partnership led to an immediate and sustained decline in SNF length of stay. The average length of stay declined by 2.3 days, a 13% decline relative to the pre-period mean of 18.4 days. This overall effect was driven, in part, by large declines in longer SNF stays. For example, SNF stays over 30 days, which comprised 12.8% of stays in the pre-period, declined by 7.1 percentage points, a decline of 56%. Declines in length of stay were larger for patients admitted to for-profit SNFs compared to non-profit SNFs, suggesting that the NaviHealth algorithm may successfully identify and reduce induced demand by SNFs. Despite concerns about algorithmic bias or discrimination by decision-makers who exercise discretion in how to use the algorithm (Albright, 2014; Davenport, 2023; Obermeyer et al., 2019), I find that the effect of the NaviHealth partnership on SNF length of stay was similar across a number of patient subgroups, including for white and black patients. In contrast to these substantial effects on the intensive margin of post-acute care use, I find no change in the extensive margin of post-acute care (where patients are discharged). For example, I find no effect on the probability of discharge to a SNF. I also find little evidence that, conditional on going to a SNF, there were substantial changes in which SNF a patient chooses, as measured by the SNF's quality rating, distance, or for-profit status. This suggests that the NaviHealth algorithm is most effective at reducing health care use once patients are admitted to a SNF. The lack of observed effects on the extensive margin may be related to what BCBS MI was doing to limit post-acute care use before the adoption of algorithmic decision-making, including their previous prior authorization system. Despite reductions in SNF length of stay, I do not observe any change in patient outcomes, as measured by 90-day readmissions and mortality. This suggests that the care that was denied was not sufficiently high-value to affect these important—but difficult to influence—outcomes.

These results significantly contribute to the literature in two ways. First, my research adds to the growing literature on the use of algorithms to aid human decision-making by evaluating a high-stakes, real-world application (Abaluck and Gruber., 2016; Agarwal et al., 2023; Albright, 2014; Angelova et al., 2023; Bundorf et al., 2024; Grimon and Mills, 2022; Gruber et al., 2020; Kleinberg et al., 2018; Sloan et al., 2023; Stevenson and Doleac, 2022). Within this literature, my work is unique in investigating the role of algorithmic decision-making when one agent uses it to monitor another agent with a different objective. Second, my work contributes to the literature on the use of managed care techniques to reduce health care use (Afendulis et al., 2017; Agafiev Macambira et al., 2022; Baker et al., 2020; Currie and Fahr, 2005; Curto et al., 2019; Duggan and Hayford, 2013; Duggan et al., 2018; Jung et al., 2024; Kuziemko et al., 2018; Layton et al., 2022). A recent strand of this literature has focused on the effects of prior authorization as one mechanism to reduce health care use and to target these reductions at unnecessary care (Brot-Goldberg et al., 2023; Dillender, 2018; Eliason et al., 2021). My work builds on this line of research by investigating how insurers use technology in prior authorizations.

2 Background and Related Literature

2.1 Medicare Advantage, Post-Acute Care, and Prior Authorization

A majority of Medicare beneficiaries are enrolled in MA. Medicare pays MA plans a risk-adjusted capitated payment and, in turn, MA plans are expected to cover their patients' qualifying medical expenses (Kaiser Family Foundation, 2024a). MA plans therefore have an incentive to provide care efficiently. A wide literature showing that MA plans reduce health care use relative to TM (Afendulis et al., 2017; Curto et al., 2019; Duggan et al., 2018; Jung et al., 2024). That research echoes work in the context of Medicaid, where there have been similar efforts for private managed care organizations to cover enrollees (Agafiev Macambira et al., 2022).

MA plans have focused on reducing post-acute care use (Achola et al., 2024; Huckfeldt et al., 2017, 2024; Prusynski et al., 2024; Skopec et al., 2020). Huckfeldt et al. (2017) and Skopec et al. (2020) compare post-acute care use among beneficiaries discharged with heart failure, joint replacement, and stroke. They do not find any substantial differences in the probability of SNF use after a hospitalization in MA but do find that, conditional on SNF use, the number of SNF days was substantially lower. They also show substantially lower probabilities of inpatient rehabilitation facility (IRF) and home health use in MA.[1] Both of these studies examined health care outcomes, finding that readmissions were lower in MA with no differences in mortality, suggesting that MA's reduction of post-acute care use does not harm beneficiaries. However, more recent research has highlighted that, while there are no differences in outcomes in administrative data, patients in MA report worse functional status related outcomes (Achola et al., 2024). While these studies provide a crucial understanding of how post-acute care use differs in MA, they are limited by favorable selection into MA. Recent quasi-experimental work has used exogenous changes in Medicare enrollment for retired Ohio state workers to provide causal evidence that MA decreases total institutional post-acute care days and IRF stays without changing readmissions or mortality (Huckfeldt et al., 2024). Overall, there is reason to conclude that MA substantially reduces post-acute care.

[1] These results do not necessarily imply that MA does nothing to the extensive margin of SNF use. High acuity patients that would use an IRF in TM may be diverted to a SNF in MA while low acuity patients who would have used a SNF in TM may be diverted home. The lack of an effect of MA on the extensive margin may be the result of these two effects in the opposite direction.

Yet a key question remains: how does MA reduce post-acute care use? There are several possibilities. First, MA plans may use different cost-sharing arrangements than TM. For example, first-dollar cost sharing used by some MA plans has been shown to substantially reduce SNF use (Keohane et al., 2017). Second, MA plans use restrictive networks. As a result, MA patients choose systematically different post-acute care providers than TM patients who live in the same area (Meyers et al., 2018; Schwartz et al., 2019). On the extensive margin, restrictive networks can imposing hassle costs that reduce the probability that people will use services (Atwood and Lo Sasso, 2016; Gruber and McKnight, 2016; Wallace, 2023). On the intensive margin, networks may steer patients to efficient providers or providers where the plan can exert greater control over the care process (Rahman et al., 2018). There is limited research on the role of networks in affecting the intensity of post-acute care use. However, some have suggested anecdotally that limited networks, for example for IRFs, may reduce health care use in this setting (Huckfeldt et al., 2024). Finally, plans can use prior authorization to limit care. Prior authorization requires providers to seek approval from insurers before they perform a service—e.g., admit a patient to an IRF, keep a patient at a SNF for an additional day. In 2023, 99% of MA beneficiaries were subject to prior authorization for SNFs (Kaiser Family Foundation, 2024b). Of the potential mechanisms, prior authorization is the most direct way an insurer can affect spending by choosing which services they deem necessary or unnecessary. Yet, to my knowledge, there is no evidence on the effects of this practice in post-acute care.

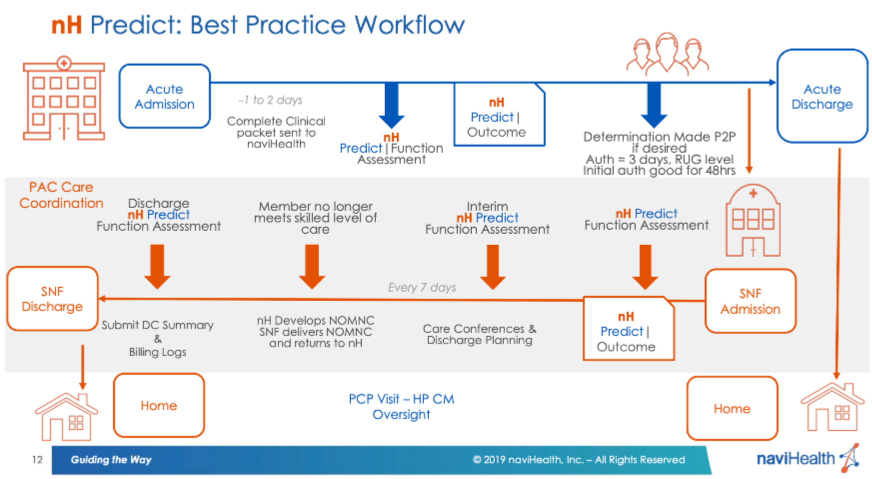
The effect of prior authorization has been the subject of substantial recent research in other health care settings. For example, in the context of Medicare Part D, Brot-Goldberg et al. (2023) show that prior authorization for prescription drugs substantially reduces use with no (observed) mortality increase. Similarly, Agafiev Macambira et al. (2022) show that prior authorization for prescription drugs and substitution to cheaper alternatives reduces overall spending in Medicaid managed care. In another context, Eliason et al. (2021) show that prior authorization for non-emergent ambulances in TM reduced use without harming patients. My research builds on this work by investigating how insurers use technology to conduct prior authorization.

2.2 NaviHealth and Algorithmic Decision-Making

The media, the courts, and policymakers have scrutinized NaviHealth's algorithmic decision-making (Chu and Nadler, 2023; Ross and Herman, 2023b,c; Jayapal, 2023). In part, this likely reflects their widespread reach. NaviHealth works with many of the largest MA carriers, including UnitedHealthcare, Humana, Cigna, and several Blue Cross Blue Shield carriers (Ross and Herman, 2023b,c; NaviHealth, 2019). In 2020, it was acquired by Optum, the owner of the largest MA carrier (UnitedHealthcare), in a deal valued at $2.5 billion (Ross and Herman, 2023a). NaviHealth is one of several similar firms, like CareCentrix and myNexus, that use an algorithm to make post-acute care decisions (Chu and Nadler, 2023). NaviHealth directly manages the post-acute care process for insures and shares in the savings (Ross and Herman, 2023a).

NaviHealth uses a proprietary prediction algorithm, called nH Predict, to guide its decision-making. While its model is proprietary, there is a considerable public information on how NaviHealth intervenes at each step of the post-acute care process. NaviHealth's promotional materials detail this process (Figure 1).

Figure 1. NaviHealth Post-Acute Care Process



Source: Obtained from the NaviHealth publicly available presentation to providers that work with Blue Cross Blue Shield of Michigan (NaviHealth, 2019).

The process begins when a patient is admitted to an acute care hospital. At this stage, the hospital is required to submit information to NaviHealth, including a history, clinical notes, cognitive and functional assessment, and information on the patient's living situation, among other factors. This information is preferably submitted through NaviHealth's portal but can also be faxed. The patient's information is fed into the nH Predict algorithm, which guides the initial authorization. This authorization is to approve the post-acute care setting after discharge—e.g., SNF, home health, or home without formal post-acute care. At this stage, there is an opportunity for a peer-to-peer discussion between the hospital clinician and a physician at NaviHealth (NaviHealth, 2019). Patients are then discharged to the chosen post-acute care setting. NaviHealth primarily discusses the SNF setting, a focus I adopt in my work. At hospital discharge, NaviHealth tries to influence which SNF a patient chooses, recognizing that there are substantial differences in practice patterns and quality across SNFs. Once admitted to a SNF, NaviHealth initially provides the SNF with a 3-day authorization. At this time, the SNF is required to submit further information to NaviHealth, including the assessment that SNFs complete for all patients at the time of admission, as well as information on medications and any evaluations by a physical or occupational therapist. NaviHealth then produces predictions of several factors: the expected length of stay at the SNF, readmission risk, functional improvement, care burden after discharge, and therapy intensity. This report is shared with the SNF and is the basis of decision-making. Some of this information is also included in a separate report given to patients and their caregivers. SNFs are required to continue to submit functional assessments at regular intervals, which nH Predict uses to adjust its initial recommendations. At some point, NaviHealth determines that the patient is ready to be discharged. A physician at NaviHealth must review and make the final determination (NaviHealth, 2019). In a pre-post designed case study, NaviHealth claims to have substantially reduced the probability of SNF use (13%) and the number of SNF days conditional on use (36%), while also reducing readmissions by 3% (NaviHealth, 2018).

A 2023 investigative reporting article detailed two anecdotes where patients had their payments cut-off by NaviHealth after a set number of days at a SNF, even though contemporaneous clinical notes indicated that they continued to need SNF level care and a federal judge later ruled that the denials were improper. There is concern that the algorithm's recommendations may not account for key patient details that are observed by nursing home clinicians (Ross and Herman, 2023a). Partially as a result, there has been considerable focus on NaviHealth and its competitors by policymakers. In recent rulemaking, the Centers for Medicare and Medicaid Services (CMS) re-emphasized that MA plans are subject to the TM coverage determinations when making authorization decisions (Centers for Medicare and Medicaid Services, 2023). In response, members of Congress emphasize the need to actively monitor the use of algorithms in authorization decisions, explicitly referencing NaviHealth and its competitors (Chu and Nadler, 2023; Jayapal, 2023). In addition, both UnitedHealthcare and Humana, the two largest MA insurers, are being sued in class-action matters for using NaviHealth's algorithm to reject post-acute care treatment. That lawmakers are focused on this policy issue emphasizes the need to understand, empirically, the effects of algorithm-aided prior authorization.

2.3 Algorithmic Decision-Making

Human decision-making is flawed in many high-stakes settings. Judges sentence low-risk defendants to pre-trial detention, physicians fail to screen high-risk patients for a heart attack, and consumers pick health insurance plans that are worse than others available to them (Abaluck and Gruber., 2016; Mullainathan and Obermeyer, 2022; Kleinberg et al., 2018). These errors in judgment involve prediction problems: for example, who is high-risk and who is low-risk? As a result, there have been efforts to augment human decision-making with algorithmically based predictions or recommendations. Early research on the topic has highlighted the possibility that algorithms can, in theory, improve on human performance (Kleinberg et al., 2018; Mullainathan and Obermeyer, 2022).

More recent work has directly evaluated the use of algorithms in decision-making, with mixed results (Albright, 2014; Grimon and Mills, 2022; Gruber et al., 2020; Sloan et al., 2023; Stevenson, 2018; Stevenson and Doleac, 2022). For example, Sloan et al. (2023) found that providing judges with risk assessment tools reduced pre-trial detention without increasing violent crime. On the other hand, Stevenson (2018) found that a similar intervention modestly increased pre-trial release but with a corresponding increase in pre-trial crime. My work is methodologically similar to this line of research, which examines how the exogenous introduction of an algorithm affects decision-making while evaluating potential unintended consequences.

This literature raises two important issues about the use of algorithms in decision-making. First, algorithms interact with the incentives that decision-makers face. For example, when setting bail, judges are concerned that releasing a low-risk defendant will cause them reputational harm if the defendant commits a crime. Albright (2014) argues that providing judges with a recommendation reduces the reputational costs of this type of error, so they become more willing to release low-risk defendants. It is critical to understand not only how a given algorithm works but also how it fits into the decision-making process and interacts with the agents that are involved. Second, there are concerns that algorithms worsen (or perpetuate) racial disparities in outcomes (Albright, 2014; Davenport, 2023; Obermeyer et al., 2019). For example, Obermeyer et al. (2019) show that a common risk prediction model in health care understates the health needs of black patients relative to white patients; black patients with the same risk score tend to be considerably sicker. If used for decision-making, this algorithm would steer resource use in a biased way. In another example, (Albright, 2014) shows that judges override algorithmic recommendations in ways that worsen racial disparities. However, in the child protective services setting, (Grimon and Mills, 2022) show that the introduction of algorithms can reduce racial disparities. Investigating differential effects of algorithms is therefore an important part of understanding algorithms and how they are used in practice.

3 Conceptual Framework

Health insurers have used managed care strategies like prior authorization for decades to monitor and reduce unnecessary care by their enrollees. What role do predictive algorithms play in changing health care use beyond what insurers already do?

The decisions to start and end post-acute care involve different providers, settings, and incentives. First, hospital clinicians must decide on the post-acute care setting. These clinicians are unlikely to have any financial interest in which setting the patient goes to.[2] Hospital clinicians have to make one post-acute care-related decision that involves a substantial documentation and planning, which should be easily observable to the insurer. Second, clinicians in the SNF determine how long the patient stays there. The SNF has a clear incentive to keep patients longer, particularly given the high rates Medicare pays (Medicare Payment Advisory Commission, 2024). SNF clinicians do not make one decision but a number of repeated decisions. The length of stay is likely not determined at the beginning of the stay but is the result of continued assessments to determine whether the patient still requires SNF level care. It would be more costly for the insurer to monitor each time a decision is made about continued care than it would be for the insurer to make a decision at one fixed point in time.

With full information, insurers have some level of care that maximizes their objective function, which includes total episode costs, reputational costs from patient dissatisfaction, and the constraint of following Medicare's legal requirements. However, observing what level of care it deems necessary is difficult and requires them to exert costly effort (e.g., employing their own clinicians to review patient records). In deciding the level of effort to exert, the insurer must consider the costs of obtaining information on the patient, and weigh these against the benefits of acquiring this information. For example, assessing a patient's need for an extra SNF day may require the insurer to hire an expert with access to a large amount of clinical and functional information about the patient. If the costs of obtaining information exceeds the insurers' potential benefit of reduced health care use, the insurer will choose not to monitor care.

The use of an algorithm reduces the cost to the insurer of obtaining information on

[2]Hospitals are paid on DRG basis, so they would prefer shorter lengths of stay and earlier discharge to post-acute care (Morrisey et al., 1988) However, it is not clear how much this affects choice of post-acute care setting.

the patient. Rather than having a person collect and evaluate a patients' clinical and functional status, this is done by the algorithm and checked by a NaviHealth employee. As a result, insurers may be most likely to use algorithms to change decision-making in settings where the cost of obtaining the information was previously high or the benefit to the insurer of reducing care is relatively low. Under this framework, I hypothesize that the use of a predictive algorithm is most likely to impact the intensive margin of SNF care. Monitoring this margin is costly without an algorithm because it requires evaluating the decision-making of the SNF at multiple points in time. However, the benefit to the insurer is relatively small: reducing a SNF length of stay by a few days is only likely to modestly reduce overall episode spending, meaning insurers are unlikely to perform this monitoring in the absence of a predictive algorithm. On the other hand, I expect that the extensive margin of SNF use is less likely to be affected by the introduction of the algorithm. The difference between even a relatively short SNF stay and discharge to home without formal care are likely to be large for the insurer. As a result, it is likely that the plan would be willing to exert costly effort to monitor this margin even in the absence of an algorithm.

4 Data and Empirical Strategy

The main methodological challenge in evaluating the effects of NaviHealth—and algorithm-aided prior authorization more broadly—is that the dates of its adoption by insurers are rarely publicly reported. For example, court records show that NaviHealth partners with UnitedHealthcare and Humana, the two largest MA carriers, but the dates that these partnerships began are not publicly known. I exploit one NaviHealth partnership that was publicly announced—between NaviHealth and BCBS MI that went into effect on June 1, 2019. BCBS MI is Michigan's largest MA insurer. Before June 1, 2019, it had a prior authorization system for post-acute care but was not using NaviHealth's algorithm. Using a difference-in-differences design and administrative data, I estimate the causal effect of this partnership by comparing health care use and outcomes for BCBS MI enrollees before and after the partnership to TM enrollees in Michigan over the same period.

4.1 Data and Sample

My primary data source is Medicare administrative data. I use the 2019 hospitalizations from MedPAR, which contains the near universe of hospitalizations in both TM and MA. I exclude stays that began within 90 days of the last hospital discharge and those that began after hospice was initiated. I limit to patients who live in Michigan, were enrolled in Part A and B during the month of hospital discharge and were enrolled in either TM or a BCBS MI plan. I then exclude hospitalizations from hospitals that had 10 or fewer Michigan-residing patients in either TM or BCBS MI throughout the year. Restricting by geography and hospital ensures that the comparison group accurately reflects underlying local trends in health care utilization and outcomes.

I link hospitalizations to SNF stays that began on the day of or day after hospital discharge using a combination of MedPAR and MA encounter data.[3] I use publicly available information on SNFs from Medicare and LTCFocus.org to obtain information on the SNF that patients go to. Further details on the data and sample construction can be found in Appendix A.

The sample includes 284,846 hospitalizations (211,155 in TM and 73,691 in BCBS MI) at 161 hospitals (see Table 1), of which 44,484 are followed by a SNF admission. The TM and BCBS MI samples differ along many observable characteristics. For example, BCBS MI patients are slightly older than those in TM, on average. These differences in TM and BCBS MI patient composition are not problematic for my research design, which does not rely directly on any assumptions about the observable or unobservable similarities between the treatment and comparison groups. In addition, two facts about BCBS MI are somewhat unique. First, most BCBS MI patients are enrolled in a Preferred Provider Organization (PPO). Nationally, a majority of MA beneficiaries are in Health Maintenance Organizations (HMO), which tend to have narrower networks and have more restrictions than PPO plans (Kaiser Family Foundation, 2024a). Second, the share of patients dually eligible for Medicare and Medicaid is low (3.9%). Nationally, 19.1 percent of Medicare beneficiaries were dually eligible in 2019 (Centers for Medicare and Medicaid Services, 2020).

[3]I use MedPAR, which contains the universe of TM paid SNF stays and a relatively small share of MA paid SNF stays, to do so where available. I use MA encounter data for MA patients when a MedPAR record is not available.

Table 1. Sample Characteristics

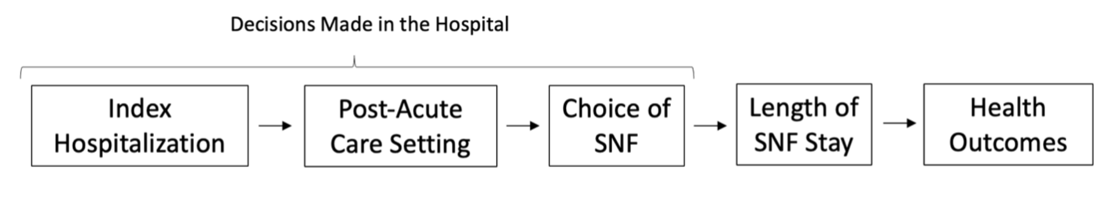
Characteristic	TM	BCBS MI
N	211,155	73,691
Number of Unique Patients	185,971	66,118
Number of Hospitals	161	161
Number of SNF Admissions	33,078	11,406
Age (Years), mean (SD)	73.16 (13.01)	76.68 (9.18)
Female	118,173 (56.0%)	38,205 (51.8%)
Dual Eligible	56,066 (26.6%)	2,899 (3.9%)
Enrolled in PPO	0 (0.0%)	59,792 (81.1%)
Race/Ethnicity		
American Indian/Alaska Native	933 (0.4%)	96 (0.1%)
Asian/Pacific Islander	2,145 (1.0%)	332 (0.5%)
Black	30,164 (14.3%)	9,207 (12.5%)
Hispanic	3,835 (1.8%)	1,115 (1.5%)
Non-Hispanic White	170,203 (80.6%)	61,785 (83.8%)
Other	1,210 (0.6%)	362 (0.5%)
Unknown	2,665 (1.3%)	794 (1.1%)
Type of Hospitalization		
Medical	144,570 (68.5%)	46,231 (62.7%)
Surgical	66,585 (31.5%)	27,460 (37.3%)
Condition		
Septicemia	15,609 (7.4%)	4,994 (6.8%)
Joint Replacement	12,703 (6.0%)	6,277 (8.5%)
Heart Failure	9,644 (4.6%)	3,434 (4.7%)
Stroke	5,888 (2.8%)	2,348 (3.2%)
Other	167,311 (79.2%)	56,638 (76.9%)

Notes: Observations are acute care hospitalizations for Medicare patients living in Michigan (see appendix A for more details on sample construction). Percentages are in parentheses for categorical variables and standard deviations are in parentheses for age. The number of SNF admissions should not be interpreted as a percent of total hospitalizations because SNF admissions after December 2019 hospitalizations were excluded from analysis. Type of hospitalization and condition were determined based on the diagnosis related group. All differences between TM and BCBS MI were statistically significant at 0.1% level using chi-squared for categorical variables and a t-test for age. The absolute value of the standardized mean difference was greater than 0.1 for age, dual eligibility, PPO, type of hospitalization but not for any of the individual race/ethnicity, condition, or female variables. *Abbreviations:* TM, traditional Medicare; BCBS MI, Blue Cross Blue Shield of Michigan. PPO, preferred provider organization; SD, standard deviation. *Source:* Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

4.2 Outcomes

I follow patients through the post-acute care process as detailed by NaviHealth (see Figure 1). Figure 2 shows how patients go through the post-acute care process in my data. I focus on those admitted to a SNF, given the NaviHealth's emphasis on that setting.

Figure 2. Post-Acute Care Process



First, I examine the choice of post-acute care setting (i.e., SNF, IRF, home health, hospice/death, home without formal post-acute care, other), as indicated on the hospitalization record. As shown in Figure 1, NaviHealth sees itself of working to intervene on this margin by conducting an authorization of this decision and providing the hospital clinician with an opportunity for a peer-to-peer to discuss this (NaviHealth, 2019). Examining this margin is important because it allows me to see whether NaviHealth encourages substitution across post-acute care settings, for example from SNFs to home health. Second, among those admitted to a SNF, I examine the characteristics of the SNF they were admitted to (profit status, distance from the patient's ZIP code tabulation area centroid to the SNF, and five-star rating). NaviHealth claims that it engages in active steering to preferred SNFs. If true, we might expect to see an increase in the quality and the distance to the chosen SNF (NaviHealth, 2019). However, because these are not necessarily the measures that NaviHealth uses to select SNFs, it would also be plausible to see different effects, though ultimately this is an empirical question. Third, among those admitted to a SNF, I examine the length of stay. One of the key goals of NaviHealth is to reduce the variability in SNF length of stay caused by different SNF practice patterns. As a result, much of their data collection and algorithm targets this margin after the SNF admission (NaviHealth, 2019). Investigative reporting has contained a number of anecdotes of patients' length of stay being affected by the NaviHealth cutting off payment after a pre-specified number of days. Finally, I examine health outcomes, as measured by 90-day mortality and readmissions. The goal of post-acute care is to improve functioning and prevent adverse outcomes. If high-value care is screened out by NaviHealth, it is likely that these outcomes will worsen as a result. I examine these outcomes for the full set of hospitalizations and for the subset that were admitted to a SNF, as this subset of patients may be particularly affected by NaviHealth in a way that is not true of patients who would have been discharged home regardless. Summary statistics for these outcomes can be found in Appendix Table B1.

4.3 Empirical Strategy

I use a difference-in-differences design to examine health care use and outcomes for BCBS MI and TM enrollees before and after the NaviHealth partnership. I estimate the following model using ordinary least squares regression:

$$y_{ict} = \alpha + \beta_1 \left(\text{PostJune}_t \times \text{BCBS}_c\right) + \mathbf{X_i} + \tau_t + \text{BCBS}_c + \epsilon_{ict} \tag{1}$$

Where y_{ict} is an outcome observed for hospitalization i for carrier c at time t; $\text{PostJune}_t \times \text{BCBS}_c$ is equal to 1 for BCBS MI patients discharged on or after June 1, 2019 and 0 otherwise; $\mathbf{X_i}$ is a vector of covariates (age, race/ethnicity, sex, DRG, ICU use, diagnosis count, and 30 indicators for chronic conditions calculated using the diagnoses on the hospitalization record); τ_t are month fixed effects; BCBS_c is an indicator for being in BCBS MI; and ϵ_{ict} is an error term. β_1 parameter of interest, the average treatment effect on the treated (ATT). I cluster standard errors at the insurer-county level.

My difference-in-differences design relies on the assumption that, in the absence of the NaviHealth partnership, health care use and outcomes for BCBS MI and TM patients would have had parallel trends after June 2019. While this assumption is ultimately untestable, I display event-studies to examine pre-partnership trends and visualize effects of the intervention. To do so, I estimate the following model:

$$y_{ict} = \alpha + \sum_{\substack{j=-5 \\ j \neq -1}}^{6} \beta_j \left(\text{BCBS}_c = 1 \ \& \ \text{Month}_t = j\right) + \text{BCBS}_c + \tau_t + \epsilon_{ict} \tag{2}$$

Where, in addition to the variables defined above, $(\text{BCBS}_c = 1 \ \& \ \text{Month}_t = j)$ are a series of monthly event-time indicators equal to 1 for BCBS MI patients discharged in the given month j (relative to June) and zero otherwise (Clarke and Tapia-Schythe, 2021). I use May 2019 as the reference month. In addition, I examine the heterogeneity of effects on health care use and outcomes and investigate changes to the distribution of SNF length of stay.

4.4 Robustness

I consider two main robustness checks. First, I alternatively use Humana in Michigan—rather than TM—as the comparison group. Humana is another large MA carrier in the state.[4] This analysis ensures that my results are not due to an MA-specific shock that might have affected both Humana and BCBS MI patients but not TM patients. TM is the comparison group in my preferred approach because of the quality of the data, the high number of patients enrolled, and the fact that TM does not use algorithm-aided prior authorization at any time during the study period—Humana's practices during this period are unknown. Nonetheless, using Humana is useful in ensuring the robustness of my main approach. Second, I examine the publicly announced partnerships of two additional considerably smaller MA carriers (MVP and Horizon) with NaviHealth to ensure that the effects I observe for BCBS MI are not idiosyncratic. These analyses are discussed in greater detail in Appendix C.

[4]Humana is the third largest MA insurer in the state behind BCBS MI and Priority Health. I use Humana rather than Priority Health because Priority Health was an early adopter of NaviHealth (NaviHealth, 2018).

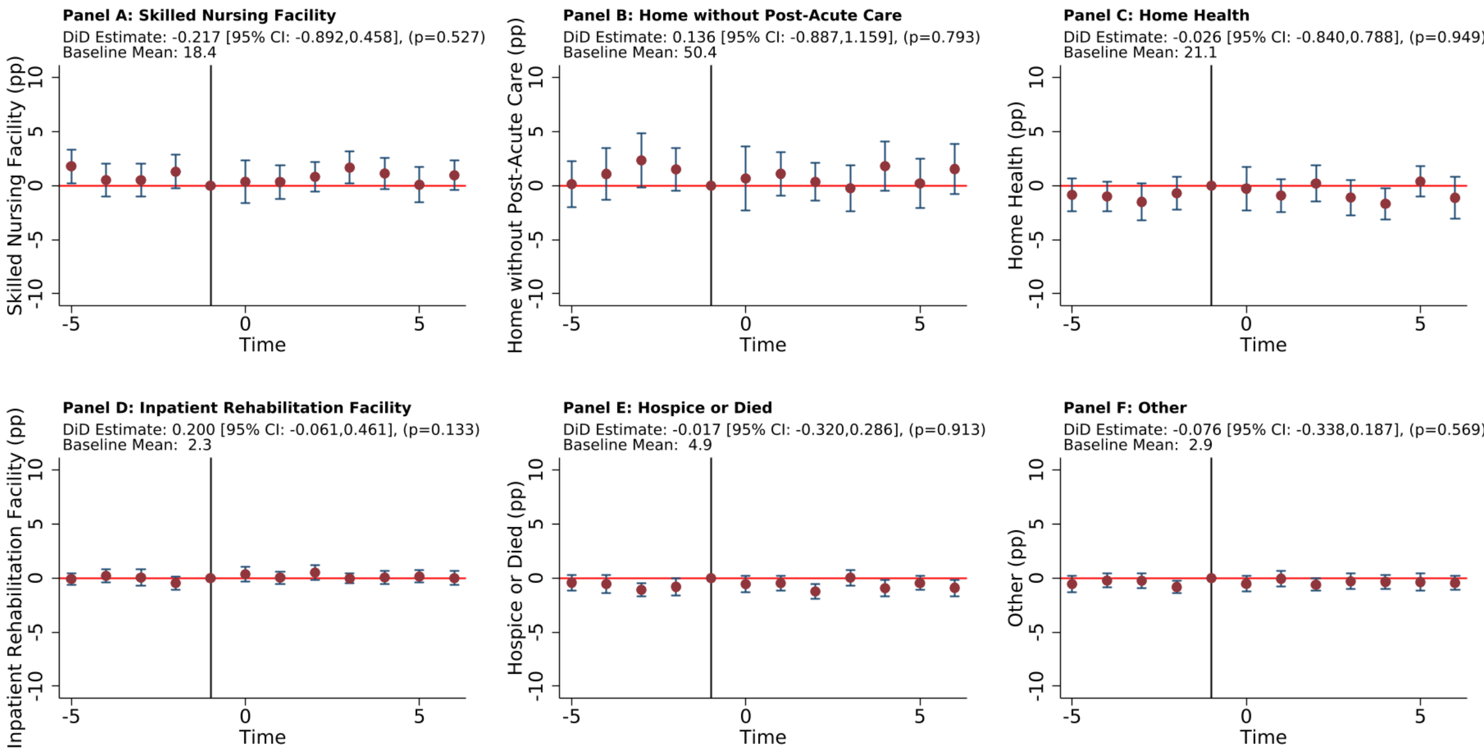
5 Results

5.1 Effect of NaviHealth on Utilization

I first evaluate the impact of BCBS-MI's partnership with NaviHealth on patients' hospital discharge settings. NaviHealth conducts an initial authorization of this decision, and the in-hospital clinician can have a peer-to-peer discussion with a NaviHealth clinician to appeal, if need be. However, I find no effects of the NaviHealth partnership on discharge setting (Figure 3). The estimates are relatively precise. For example, the null effect on the probability of discharge to SNF was -0.217 percentage points (95% CI: -0.892,0.458). Estimates for all main outcomes in this section can be found in Appendix Table B2. There are also concerns that prior authorization can affect the timeliness of hospital discharge, though the directions of the effect are theoretically ambiguous. Hospital length of stay could be shortened if the algorithm-aided process is faster or lengthened if it takes longer than the previous system. However, in Appendix Figure B2, I also find no effect of the NaviHealth algorithm on hospital length of stay.

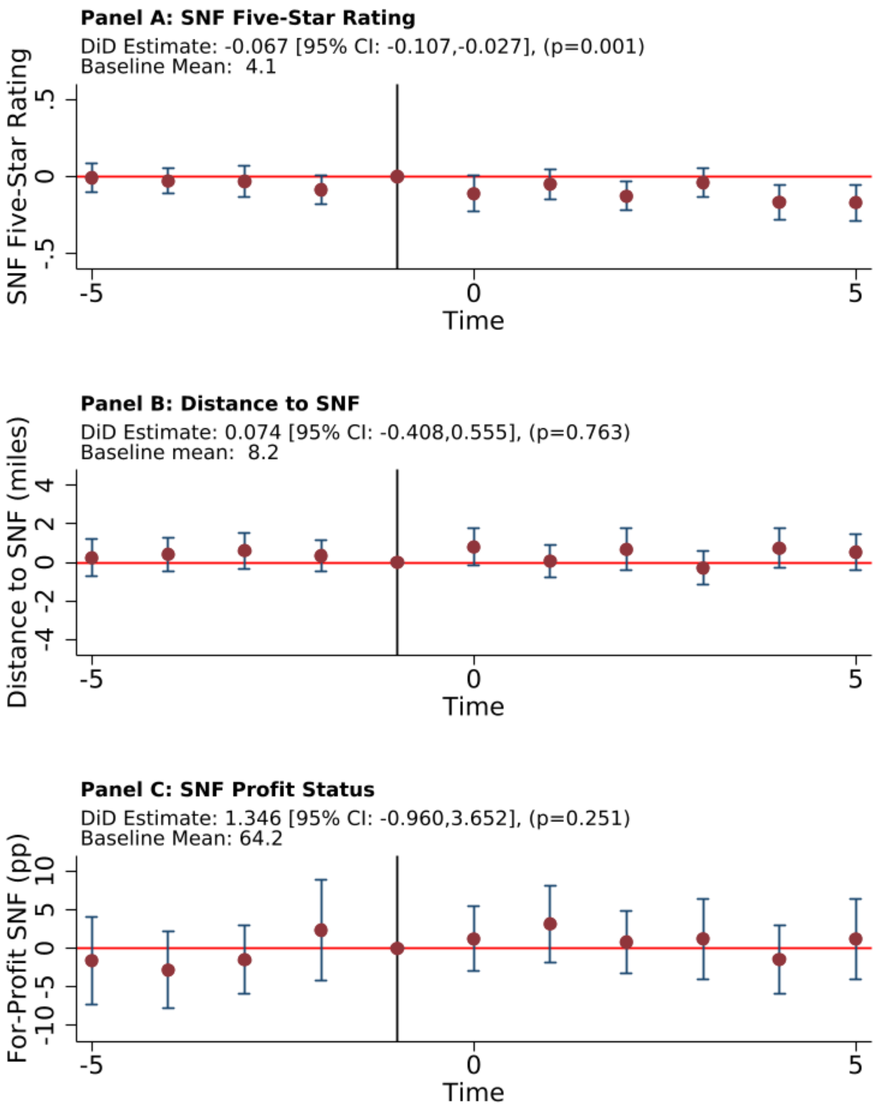
I next examine whether the algorithm changes characteristics of the SNF that a patient was discharged to. This is another margin NaviHealth claims to operate on by steering patients to their preferred SNFs. I find no evidence of systematic changes in patients' distances to their SNFs or in the probability of patients' choosing a for-profit SNF (Figure 4). I do find that the NaviHealth partnership modestly reduced the quality of the SNFs that patients were admitted to. It reduced the five-star rating of the chosen SNF by 0.067 stars (95% CI: -0.107,-0.027). However, this is only a 1.6 percent decline relative to the pre-period mean of 4.1 stars for BCBS MI patients. This change is very modest and was only statistically significant during 3 of the 6 post-partnership months. Overall, there is little evidence that decision-making in the hospital changed, with no effects on which post-acute care setting was chosen and, conditional on going to a SNF, which SNF was chosen. Yet, intervening on decisions made in the hospital is only one way that NaviHealth can reduce health care use.

Figure 3. Effect of NaviHealth Partnership on Discharge Setting



Notes: Each panel shows ordinary least squares event-study estimates of the effect of the Blue Cross Blue Shield of Michigan partnership with NaviHealth in June 2019. Time is relative to June 2019. There is a vertical May 2019, which is used as the reference month (time = -1). Pooled difference-in-differences estimated with covariates are shown above each figure. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. The six outcomes are mutually exclusive and exhaustive. They were obtained from the discharge destination code on the index hospitalization record. Standard errors were clustered at insurer-county level.
Abbreviations: pp, percentage point.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure 4. Effect of NaviHealth Partnership on Skilled Nursing Facility Choice

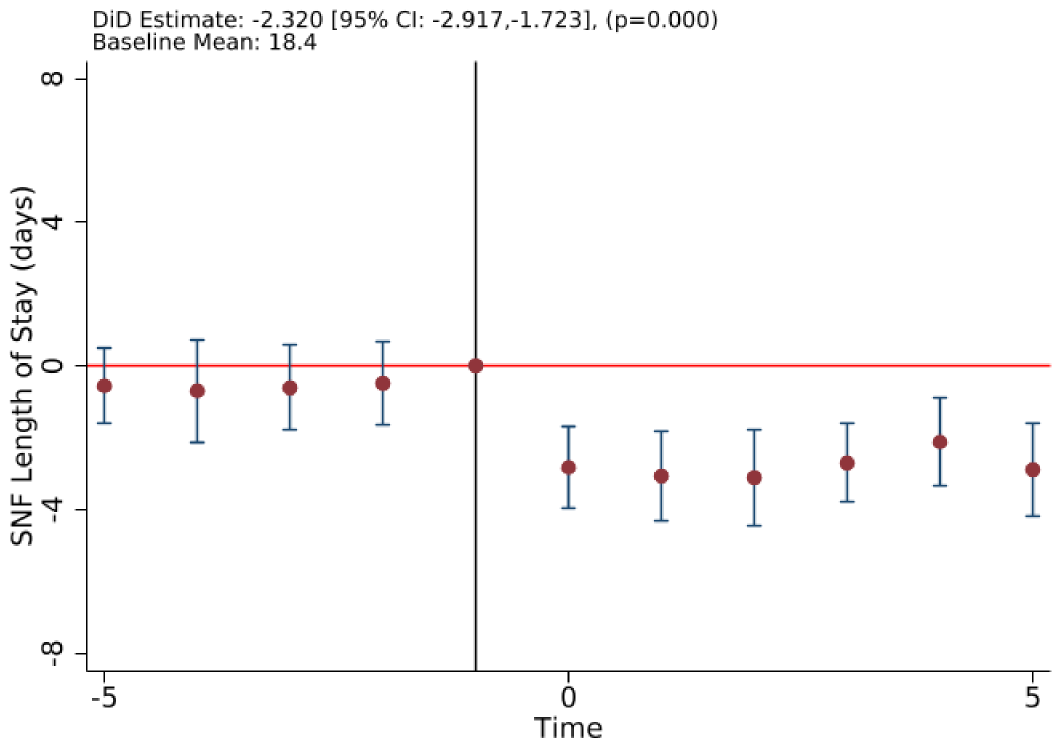


Notes: Each panel shows ordinary least squares event-study estimates of the effect of the Blue Cross Blue Shield of Michigan partnership with NaviHealth in June 2019. Time is relative to June 2019. There is a vertical May 2019, which is used as the reference month (time = -1). Pooled difference-in-differences estimated with covariates are shown above each figure. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Only patients admitted to a skilled nursing facility with non-missing information on the relevant outcome variable were included. Standard errors were clustered at insurer-county level.
Abbreviations: SNF, skilled nursing facility; pp, percentage point.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, the Medicare Master Beneficiary Summary File (MBSF), and publicly-available skilled nursing facility data.

In Figure 5, I show that the NaviHealth-BCBS MI partnership resulted in an immediate and sustained reduction in the average SNF length of stay. The 2.320 day decline in SNF length of stay (95% CI: -2.917,-1.722) was a 12.6 percent decline relative to the pre-period BCBS MI mean of 18.4 days. While this average change is substantial, I next examine how the partnership changed the distribution of SNF length of stay. Figure 6 shows the distributions of length of stay before and after the partnership for TM (Panel A) and BCBS MI patients (Panel B). The TM distribution is essentially unchanged before and after June 1. However, the length of stay distribution for BCBS MI compressed substantially after the partnership, as the right side of distribution meaningfully shifted inwards.

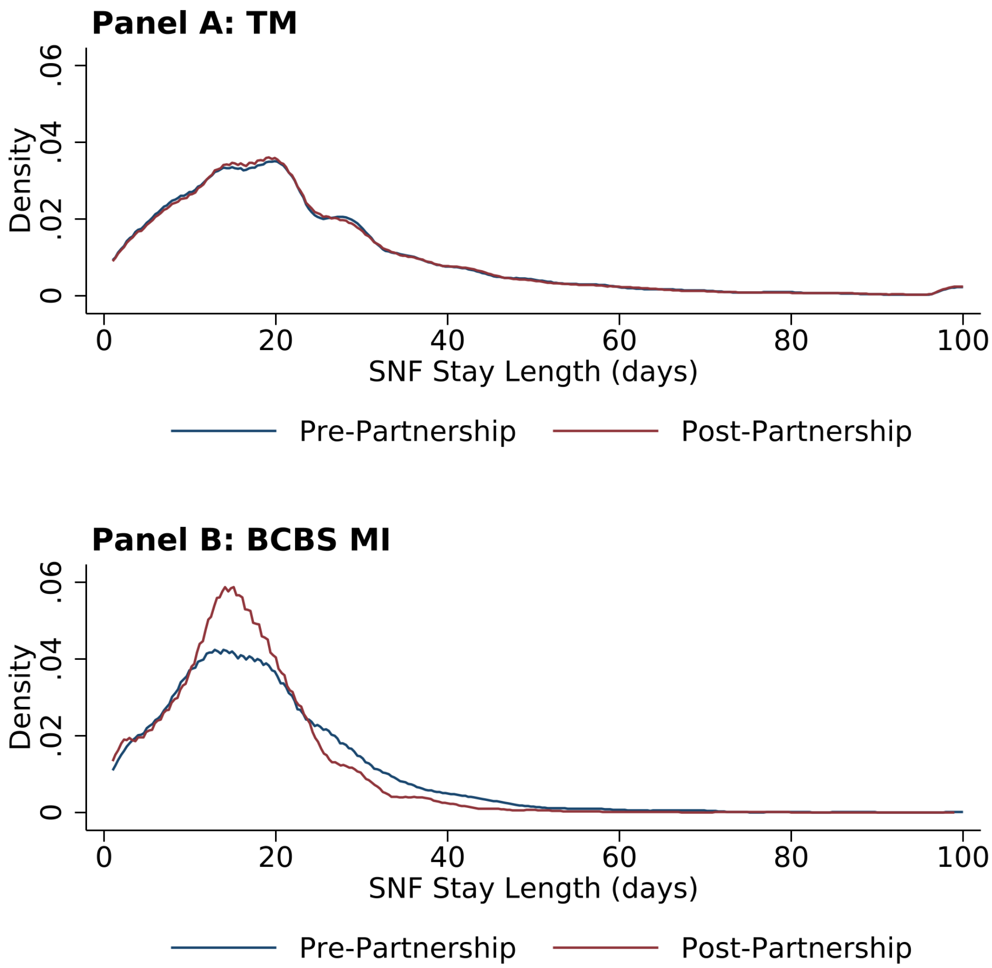
I examine the change in the SNF length of stay distribution more concretely in Appendix Table B3, where I show the distribution of SNF length of stay by insurer before and after the partnership. I also use a difference-in-differences quantile regression to examine the effect at each percentile. There was no statistically significant effect at either the 10th or 25th percentile of the length of stay distribution. There was a modest effect at the median, which declined by 2 days. The 75th and 90th percentiles experienced even greater declines, of 4 and 6 days, respectively. These results imply that NaviHealth compressed the length of stay distribution, reduced variability in length of stay, and substantially lowered the prevalence of the longest SNF stays. In Figure 7, I examine the effect of the partnership on long SNF stays (defined as those over 30 days) using an event-study framework to further assess how these types of stays were affected. I find that the partnership led to a 7.137 percentage point decline (95% CI: -8.611,-5.663) in the probability of a SNF stay over 30 days. This represented a 55.8 percent decline relative the pre-period BCBS MI mean of 12.8 percent of SNF stays.

Figure 5. Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay



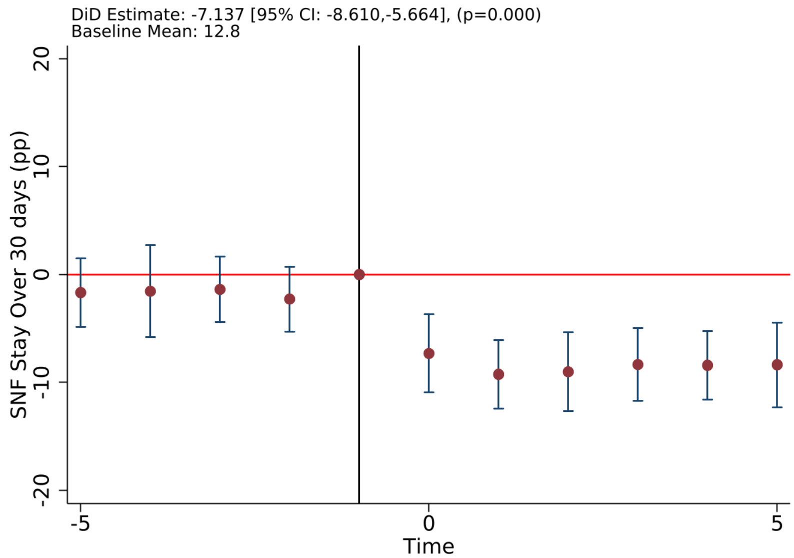
Notes: This figure shows ordinary least squares event-study estimates of the effect of the Blue Cross Blue Shield of Michigan partnership with NaviHealth in June 2019. Time is relative to June 2019. There is a vertical May 2019, which is used as the reference month (time = -1). Pooled difference-in-differences estimated with covariates are shown above the figure. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Only patients admitted to a skilled nursing facility were included. Standard errors were clustered at insurer-county level.
Abbreviations: SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure 6. Effect of NaviHealth Partnership on the Distribution of Skilled Nursing Facility Length of Stay



Notes: Each panel shows the distribution of skilled nursing facility length of stays before and after the June 1, 2019 partnership for traditional Medicare and Blue Cross Blue Shield of Michigan patients.
Abbreviations: SNF, skilled nursing facility; TM, traditional Medicare; BCBS MI, Blue Cross Blue Shield of Michigan.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure 7. Effect of NaviHealth Partnership on Longer Skilled Nursing Facility Stays



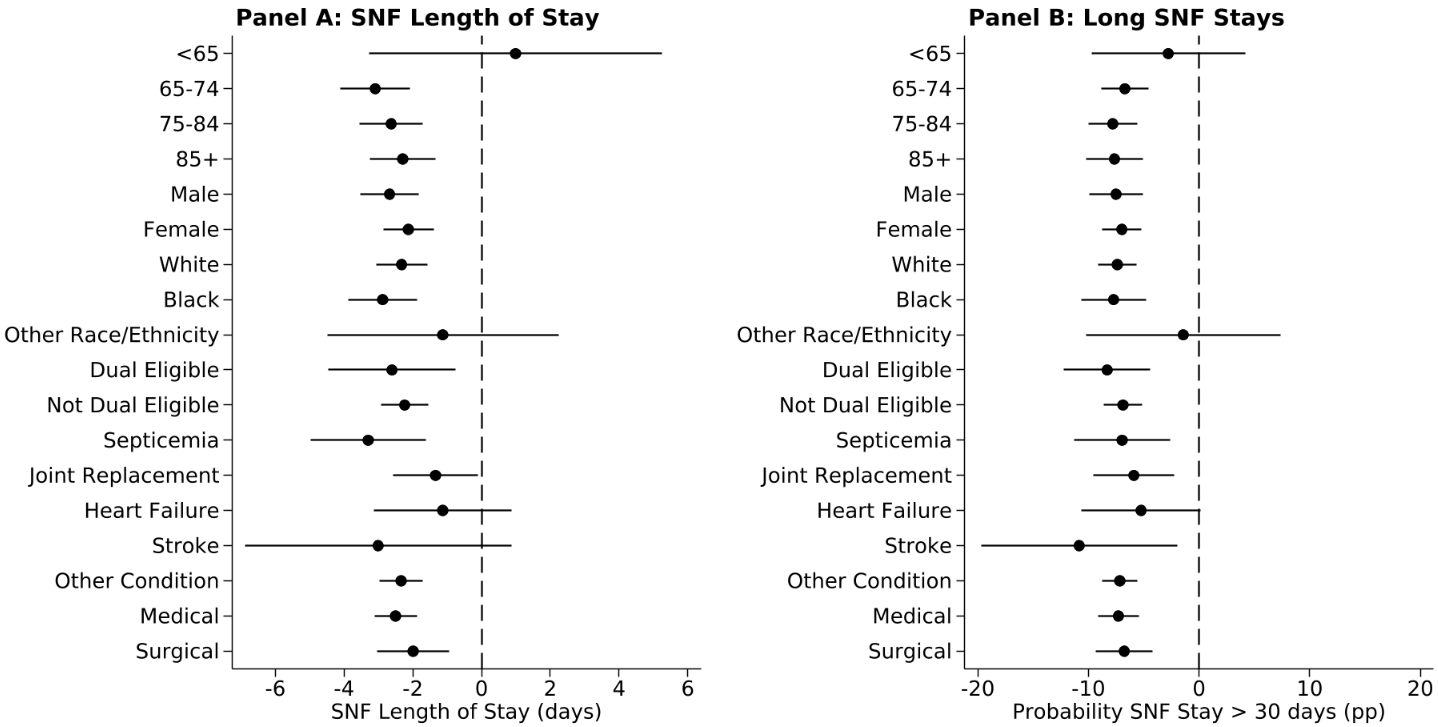
Notes: This figure shows ordinary least squares event-study estimates of the effect of the Blue Cross Blue Shield of Michigan partnership with NaviHealth in June 2019. Time is relative to June 2019. There is a vertical May 2019, which is used as the reference month (time = -1). Pooled difference-in-differences estimated with covariates are shown above the figure. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Only patients admitted to a skilled nursing facility were included. Standard errors were clustered at insurer-county level.
Abbreviations: SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

5.1.1 Heterogeneity in the Effect of NaviHealth on Length of Stay

I next examine whether the causal effect of the NaviHealth partnership on SNF length of stay varies by patient characteristics. A key concern for algorithmic decision-making is that algorithms may differentially target patient subgroups. For example, there are concerns that algorithms may worsen racial disparities. However, in this setting I find that the estimates are strikingly similar across all of the patient subgroups I examine (Figure 8). For example, I find no evidence that this algorithm differentially affects white and black patients. The effect was 2.301 days for white patients (95% CI: -3.051,-1.551), almost identical to the effect of 2.343 days for black patients (95% CI: -3.195,-1.491).

Next, I examine whether effect on SNF length of stay varies by the ownership of the SNF (Table 2). One factor motivating the use of an algorithm to reduce care is the recognition that there is wide variation in practice patterns, some of which may be driven by the incentive that SNFs have to induce demand by extending stays. In theory, we would expect that for-profit SNFs are more likely to engage in this behavior than non-profits (Bowblis et al., 2016). If so, the use of an algorithm which is intended to screen out unnecessary care may have larger effects at for-profit SNFs. Consistent with this hypothesis, I find that the effect is larger for patients treated at for-profit SNFs than it is for patients treated at non-profit SNFs. For example, the effect on length of stay is -2.728 days (95% CI: -3.557,-1.899) at for-profit SNFs but only -1.576 (95% CI: -2.535,-0.616) at non-profit SNFs (p-value of difference =.055). The effect on the probability of having a SNF stay over 30 days is -8.946 percentage points (95% CI: -10.89,-7.000) at for-profit SNFs but only -3.713 percentage points (95% CI: -6.499,-0.926) at non-profit SNFs (p-value of difference = .002). One likely driver of this is that there were the baseline differences in length of stay at for-profits and non-profits. The baseline mean length of stay was 19.0 days at for-profits (14.2 percent of stays over 30 days) while it was 17.4 days at non-profits (9.6 percent of stays over 30 days). The baseline rates may differ either because of differing practice patterns at non-profits and for-profits or because differences in case-mix. Regardless of the cause, the difference in the baseline values and the reduction in the variation in length of stay caused by NaviHealth resulted in larger effects at for-profit SNFs.

Figure 8. Heterogeneity in the Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay, by Patient Characteristics



Notes: Each panel shows a set of difference-in-differences estimates. Each value represents the effect estimate (and 95% confidence interval) for a separately estimated ordinary least squares regression including only the given patient subgroup. Long stays are those over 30-days. Only patients admitted to a skilled nursing facility were included. Condition and medical/surgical indicators were based on the diagnosis related group of the index hospitalization. Each regression is adjusted for covariates. Standard errors were clustered at insurer-county level.
Abbreviations: SNF, skilled nursing facility pp, percentage point.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Table 2. Heterogeneity in the Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay, by Skilled Nursing Facility Ownership

	For-Profit	Non-Profit	P-Value of Difference
Panel A: Skilled Nursing Facility Length of Stay			
Effect	-2.728***	-1.576**	.055
	[-3.557,-1.899]	[-2.535,-0.616]	
Baseline Mean	19.0	17.4	
Panel B: Skilled Nursing Facility Stay >30 Days			
Effect	-8.946***	-3.713**	.002
	[-10.89,-7.000]	[-6.499,-0.926]	
Baseline Mean	14.2	9.6	
N	29,093	11,391	

Notes: Entries are difference-in-differences estimates and 95% confidence intervals from ordinary least squares regressions with covariates. Regressions are estimated separately for each outcome on the subset of patients admitted to for-profit or non-profit skilled nursing facilities. Estimates for patients admitted to government owned skilled nursing facilities are not shown. The p-value is obtained by fully interacting each parameter in the difference-in-differences model, including all the covariates and fixed effects with an indicator for being treated at a for-profit facility. Only patients treated at for-profit and non-profit skilled nursing facilities were included. The relevant p-value was obtained from the coefficient on the interaction of for-profit with the difference-in-differences indicator (treatment times post). The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Standard errors were clustered at insurer-county level.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, the Medicare Master Beneficiary Summary File (MBSF), and publicly-available skilled nursing facility data.
$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

5.1.2 Robustness of the Effect of NaviHealth on Length of Stay

In Appendix Table B4, I demonstrate the robustness of the decline in the SNF length of stay and the probability of having a long SNF stay to alternative methods of obtaining CI and p-values. In addition to the point estimates and the CI obtained by clustering standard errors at the insurer-county, I show the CI and p-values for several other methods. I show CI obtained using robust SEs without clustering, analytic CI based on SEs clustered at the insurer level, and wild bootstrapped CI obtained using the method suggested by (MacKinnon and Webb, 2018). The un-clustered robust CI are very similar to my preferred method. The analytic CI clustered at the county are substantially narrower because cluster robust SEs are biased downwards when the number of clusters is small (Cameron et al., 2008). It is for this reason that I use the cluster at insurer-county level, following the guidance to always use the more conservative approach when there is uncertainty in the level to cluster at. In addition, clustering at the insurer-county considers the potential correlation of the error term within a geographic area over time. Because of the bias in analytic standard errors clustered by insurers, I also use the MacKinnon and Webb (2018) wild bootstrap method for very small number of clusters (there are only two when I cluster by insurer) (Roodman et al., 2019). This method under-rejects the null hypothesis in difference-in-differences designs, particularly when the treated cluster is smaller than the untreated cluster, as in my setting (MacKinnon and Webb, 2018). It should be used only as a conservative approach to obtaining CI but nonetheless is useful to compare to my preferred approach of clustering at the insurer-county. Despite differences in precision across methods, the results are robust across these alternate approaches.

In Appendix Figure B2, I show the estimated effect on length of stay across different specifications. I show models with and without covariates and with and without hospital fixed effects. I show models excluding October and November, where there was a modest decline in the completeness of the encounter data (see Appendix A for further discussion). I show models where I remove the exclusions that patients must live in Michigan and go to a hospital that sees at least 11 BCBS MI and TM patients. I also show a model where I use Poisson regression rather than OLS and show average marginal effects. The estimates are almost identical across all of these different specifications, emphasizing the robustness of this finding to alternate analytic choices.

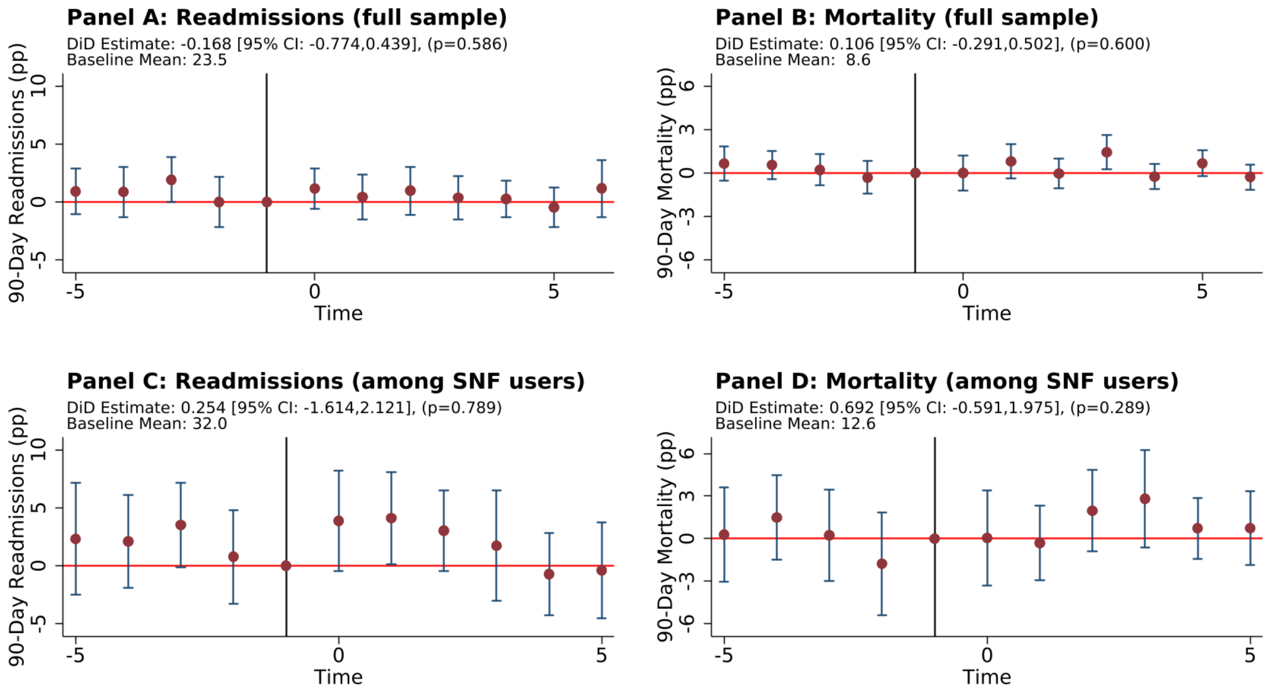
5.2 Effect of NaviHealth on Health Outcomes

I next examine whether the NaviHealth partnership affected health outcomes. This is important given that the major concern with algorithmic decision-making in this setting is that it screens out high-value care that would have improved health had it been provided. Because I find that the utilization reductions driven by NaviHealth occur at the SNF (rather than at the hospital discharge phase), I examine outcomes for both the full set of hospitalizations and the subset of patients who are subsequently admitted to SNFs. I find no effect of the NaviHealth partnership on 90-day mortality or readmissions, either for the full sample of hospitalizations or for patients admitted to a SNF (Figure 9). The full sample estimates are relatively precise. For example, the 95% CI for 90-day readmissions ranges from -0.774 to 0.439, relative to a mean of 23.5. However, because the set of SNF-admitted patients is smaller, the estimates for this subset are less precise, with the 95% CI for 90-day readmissions ranging from -1.615 to 2.122.

Even among patients admitted to a SNF, the algorithm-driven cuts are not likely to be binding for all patients. For example, as I show above, a patient who would have had a 11-day stay is unlikely to be affected. On the other hand, a patient who would have had a 33-day stay is likely to experience a major cut in their health care use (see Table B3).

Unfortunately, I do not observe the counterfactual length of stay for each patient. Instead, I use the TM sample of SNF admissions and baseline covariates to predict the probability of having a stay over 30 days (age, race/ethnicity, sex, DRG, ICU use, diagnosis count, and 30 indicators for chronic conditions calculated using the diagnoses on the hospitalization record, as well as the hospital length of stay). These variables explain 6.3 percent of the variation in the probability of a longer SNF stay in the TM sample. Using this predicted probability for all SNF users, I partition the sample into quartiles and estimate the effect of the partnership at each quartile. Critically, the point estimate on SNF length of stay is strictly increasing across quartiles. The point estimate of the effect of the partnership on NaviHealth is larger by over a day in the highest quartile of predicted longer stay probability -2.766 compared to -1.691 (Appendix Table B5. This is a useful "first stage" for this analysis: my method is able to divide patients into groups who are more or less effected by the utilization reductions. Having divided the sample in this way, I can further examine whether there are any adverse effects for any of these groups. I find no effects on readmissions for any of these groups (Figure 10). Even for those most likely to be exposed to the SNF length of stay reductions caused by the NaviHealth partnership, there was no effect on health outcomes.

Figure 9. Heterogeneity in the Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay, by Patient Characteristics

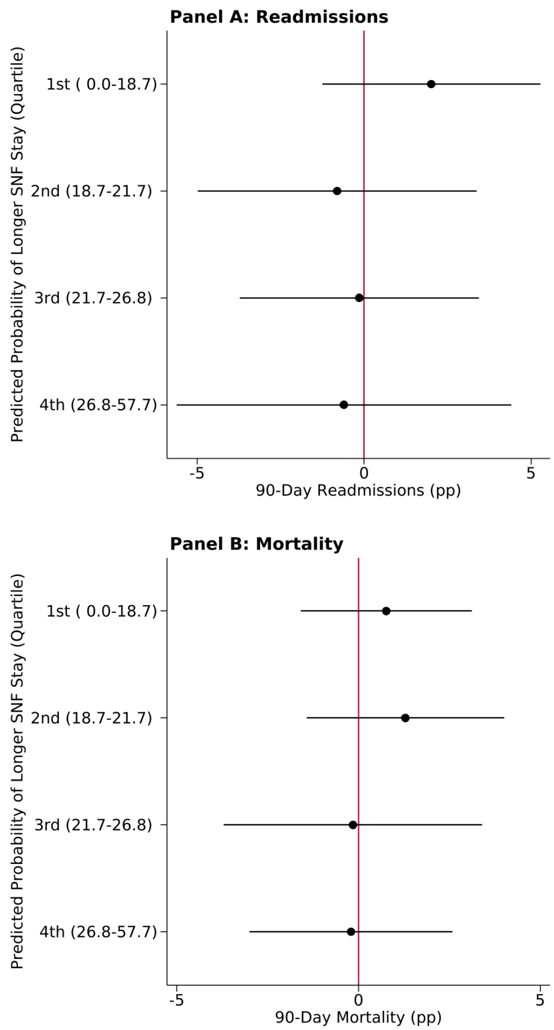


Notes: Each panel shows ordinary least squares event-study estimates of the effect of the Blue Cross Blue Shield of Michigan partnership with NaviHealth in June 2019. Time is relative to June 2019. There is a vertical May 2019, which is used as the reference month (time = -1). Pooled difference-in-differences estimated with covariates are shown above each figure. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Outcomes were measured at 90 days from hospital discharge. Standard errors were clustered at insurer-county level.

Abbreviations: SNF, skilled nursing facility pp, percentage point.

Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure 10. Heterogeneity in the Effect of NaviHealth Partnership on Health Outcomes, by Predicted Skilled Nursing Facility Length of Stay



Notes: Each panel shows four difference-in-differences estimates. Patients who were admitted to a skilled nursing facility were stratified by their predicted probability of having a skilled nursing facility stay over 30 days. This predicted probability was obtained by regressing an indicator for having a stay over 30 days on covariates (age, race/ethnicity, sex, diagnosis related group, intensive care unit use, diagnosis count, 30 indicators for chronic conditions calculated using the diagnoses on the hospitalization record, and hospital length of stay) for the sample of traditional Medicare skilled nursing facility users and applying the predicted probability from this model to the entire sample of patients admitted to a skilled nursing facility. Patient were then divided into quartiles based on this predicted probability. Values are difference-in-differences estimates of the effect on 90-day readmissions and mortality estimated separately (adjusting for covariates) for each quartile. Standard errors were clustered at insurer-county level.
Abbreviations: SNF, skilled nursing facility; pp, percentage point.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

5.3 Robustness to Alternate Comparison Group and Partnerships with Other Insurers

I consider two robustness checks. First, I use Humana patients in Michigan as the comparison group, where they are a large insurer with high quality SNF encounter data. Appendix Figure B4 shows an event-study of the effect on SNF length of stay and Appendix Table B5 shows difference-in-differences estimates for all outcomes. The results are robust to this alternative comparison group. Second, I consider two additional partnerships between NaviHealth and MA insurers. These results are discussed at length in Appendix C. They emphasize that NaviHealth's effects on SNF length of stay and lack of meaningful effects on other outcomes are not specific to the BCBS MI partnership but have broader generalizability.

6 Discussion

Firms increasingly use predictive algorithms to aid high-stakes decision-making. Using a difference-in-differences design and administrative data, I investigate the causal effect of a partnership between a large MA insurer and a firm that uses a predictive algorithm to aid its determination of what post-acute care is 'necessary'. I find that this partnership immediately reduced post-acute care use without observably worsening patient outcomes.

First, the partnership between NaviHealth and BCBS MI reduced SNF length of stay. This 13% decline was substantial and immediate. This effect was due not only to an economically meaningful average effect, but also to a substantial compression in the length of stay distribution. For example, the percentage of SNF stays over 30 days declined by 56%.Using a predictive algorithm reduced variation in SNF stay length by substantially reducing longer stays. These declines in utilization are similar in size across all of the patient subgroups that I examined, indicating that no patient groups were differentially screened out by the algorithm. For example, I find similar effects on SNF length of stay for white and black patients, despite broader concerns about racial bias in algorithms or racial discrimination in their use (Albright, 2014; Davenport, 2023; Obermeyer et al., 2019). I find that the effects on length of stay were larger at for-profit SNFs compared to non-profits. The decline in SNF stays over 30 days was more than twice as large for patients at for-profit SNFs as it was for patients at non-profit SNFs, reflecting the longer baseline stays at for-profit SNFs and the NaviHealth's compression length of stay distribution. In theory, we expect that for-profit firms would be more willing to induce demand by unnecessarily extending stays, though empirical evidence suggests that, on average, for-profit SNFs do not have (causally) longer stays (Bowblis et al., 2016). Nonetheless, my finding that the effects on length of stay are larger at for-profit SNFs is consistent with the idea that the use of a predictive algorithm may screen out induced demand by SNFs. While SNF length of stay declined significantly, this was the only margin where the partnership affected health care use. For example, despite NaviHealth's advertising claims that they affect the probability of SNF use, I do not observe any effect. In fact, my confidence intervals rule out declines in SNF use larger than 0.9 percentage points. The difference in the effects of SNF use on the extensive and intensive margins is most likely related to BCBS MI's behavior before partnering with NaviHealth. Like almost all MA plans (Kaiser Family Foundation, 2024b), BCBS MI used prior authorization in place before beginning to use NaviHealth's algorithm-aided approach. I expect that its prior authorization system monitored patients' discharge to SNFs closely. This one-time decision is likely not that costly to monitor while the benefits to the plan of avoiding an unnecessary SNF admission are substantial. On the other hand, monitoring decisions by SNFs about whether to continue care may be more costly. This is because these decisions are repeated continually and the benefit to plans of modestly reducing stay lengths is not as large as the benefit of avoiding a SNF admission altogether. The differences in the results for the probability of SNF use and the SNF length of stay are consistent with the idea that the use of predictive algorithms reduces the cost of monitoring providers and therefore, should have the greatest effect when pre-algorithm monitoring costs are high and the benefits to the plan of reducing health care use are modest.

Second, I find that the NaviHealth-BCBS MI partnership had no observable effect on health outcomes, as measured by 90-day readmissions and mortality. There were no effects in the entire sample of hospitalized patients, in the subset of patients who went to a SNF, and in the subset of SNF-admitted patients with the highest predicted SNF length of stay. Two factors may explain why I find no evidence of adverse effects on patients. On one hand, it may be that NaviHealth's utilization reductions are well-targeted. It is possible case that there were no adverse effects for the patients whose care was reduced by NaviHealth but had other patients not targeted by NaviHealth experienced similar sized utilization reductions there would have been effects on health outcomes. On the other hand, it is possible that the utilization reductions that I observe are insufficient to change patients' outcomes, regardless of the degree of targeting. Past work has shown very limited returns to additional days in a SNF (**?**Werner et al., 2023). As a result, even broad-based reductions in SNF stay lengths may not affect patient outcomes.

The lack ofobservable effects on health outcomes suggests that using an algorithm to make post-acute care coverage decisions in MA has little effect on overall welfare. In TM, a 13% reduction in SNF length of stay would save the federal government around 3.8 billion dollars a year (Medicare Payment Advisory Commission, 2024). However, in MA there is no direct connection between health care utilization and Medicare spending. MA plans are paid on a risk adjusted per-member per-month basis that is generally unaffected by the expenditures that the MA plan has for its patients. The reduction in SNF spending is therefore a pure transfer from the SNF to the health insurer. Welfare might be increased if patients value being discharged home earlier and if the algorithm reduces administrative burden for providers but decreased if patients are frustrated by the NaviHealth process and providers face added burdens. However, the lack of change in health outcomes and the lack of a direct mechanism for translating utilization reductions into savings for Medicare suggest that the broader impacts of algorithm-aided prior authorization in MA are relatively limited.

This study has three main limitations. First, I examine the effect of algorithmic decision-making on health outcomes using readmissions and mortality but lack information about more proximal health outcomes that are not contained in claims data. For example, I lack information on recovery of functional status or caregiving burden, which could both be plausibly affected by the intervention. Nonetheless, readmissions and mortality are clinically meaningful outcomes commonly used in the literature to determine whether changes in post-acute care use harm patients (McGarry et al., 2021; Werner et al., 2023). Second, I use length of stay from the MEDPAR and MA encounter data. While I provide evidence that BCBS MI has relatively complete encounter data, it is still possible that some SNF stays are missing from the data. Moreover, my measure of SNF length of stay is restricted to the Medicare-paid length of stay. If there is substitution to private-paid (or Medicaid-paid) nursing home use, I do not observe this. Finally, this study focuses on the partnership between NaviHealth and one large MA insurer. These results may not generalize to all NaviHealth partnerships, particularly after Optum's 2020 acquisition of NaviHealth, which may have changed NaviHealth's practices (Ross and Herman, 2023a). However, in Appendix C, I provide evidence from two other NaviHealth partnerships where I observe similar effects as what is observed for BCBS MI. It is unlikely that the effects in this study are specific to BCBS MI.

7 Conclusion

Insurers use predictive algorithms to reduce the level and variability of health care use. Despite concerns that the use of algorithms reduces necessary care, I find no evidence of worsening health outcomes.

References

Abaluck, J. and J. Gruber. (2016). Improving the Quality of Choices in Health Insurance Markets. https://www.nber.org/papers/w22917.

Achola, E. M., D. G. Stevenson, and L. M. Keohane (2024). Postacute care services use and outcomes among traditional medicare and medicare advantage beneficiaries. *JAMA Health Forum 4*(8), e232517.

Afendulis, C. C., M. E. Chernew, and D. P. Kessler (2017). The Effect of Medicare Advantage on Hospital Admissions and Mortality. *American Journal of Health Economics 3*(2), 254–279.

Agafiev Macambira, D., M. Geruso, A. Lollo, C. D. Ndumele, and J. Wallace (2022). The Private Provision of Public Services: Evidence from Random Assignment in Medicaid. https://www.nber.org/papers/w30390.

Agarwal, N., A. Moehring, P. Rajpurkar, and T. Salz (2023). Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology. https://www.nber.org/papers/w31422.

Albright, A. (2014). The Hidden Effects of Algorithmic Recommendations. https://www.goodrx.com/blog/health-insurance-aside-americans-still-struggle-to-pay-for-their-medications/.

Angelova, V., W. S. Dobbie, and C. Yang (2023). Algorithmic Recommendations and Human Discretion. https://www.nber.org/papers/w31747.

Atwood, A. and A. T. Lo Sasso (2016). The Effect of Narrow Provider Networks on Health Care Use. *Journal of Health Economics 50*, 86–98.

Baker, A. C. D. F. L. and C. C. Y. Wang. (2022). How Much Should We Trust Staggered Difference-in- Differences Estimates? *Journal of Financial Economics 144*(2), 370–95.

Baker, L. C., M. K. Bundorf, and D. P. Kessler (2020). The Effects of Medicare Advantage on Opioid Use. *Journal of Health Economics 70*, 102278.

Barnett, M. L., A. Mehrotra, and D. C. Grabowski (2019). Postacute Care — The Piggy Bank for Savings in Alternative Payment Models? *New England Journal of Medicine 381*(4), 302–303.

Biniek, J. F., A. Bloschichak, and S. Rodriguez (2019). Comparing Post-Acute Care Use and First Site of Care Among Medicare Advantage Enrollees and Medicare Fee-for-Service Beneficiaries. https://healthcostinstitute.org/hcci-originals-dropdown/all-hcci-reports/comparing-post-acute-care-use-and-first-site-of-care-among-medicare-advantage-enrollees-and-medicare-fee-for-service-beneficiaries.

Bowblis, J. R., J. Horowitz, and C. S. Brunt (2016). Ownership Status and Length of Stay in Skilled Nursing Facilities: Does Endogeneity Matter? *Journal of Applied Gerontology 35*(3), 303–320.

Brot-Goldberg, Z. C., S. Burn, T. Layton, and B. Vabson (2023). Rationing Medicine Through Bureaucracy: Authorization Restrictions in Medicare. https://www.nber.org/papers/w30878.

Bundorf, M. K., M. Polyakova, and M. Tai-Seale (2024). How Do Consumers Interact with Digital Expert Advice? Experimental Evidence from Health Insurance. *Management Science*.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-Based Improvements for Inference with Clustered Errors . *The Review of Economics and Statistics*.

Centers for Medicare and Medicaid Services (2020). Data Analysis Brief: Medicare-Medicaid Dual Enrollment 2006 through 2019. https://www.cms.gov/files/document/medicaremedicaiddualenrollmenteverenrolledtrendsdatabrief.pdf.

Centers for Medicare and Medicaid Services (2023). 2024 Medicare Advantage and Part D Final Rule (CMS-4201-F). https://www.cms.gov/newsroom/fact-sheets/2024-medicare-advantage-and-part-d-final-rule-cms-4201-f.

Chandra, A., M. A. Dalton, and J. Holmes (2013). Large Increases In Spending On Postacute Care In Medicare Point To The Potential For Cost Savings In These Settings. *Health Affairs 32*(5), 864—872.

Chandra, A., E. Flack, and Z. Obermeyer (2024). The Health Costs of Cost Sharing. *Quarterly Journal of Economics*.

Chu, J. and J. Nadler (2023). Letter to Chiquita Brooks-LaSure, Adminstrator of the Centers for Medicare and Medicaid Services. https://chu.house.gov/sites/evo-subsites/chu.house.gov/files/evo-media-document/chu-nadler-ma-ai-oversight-letter-11.3.2023.pdf.

Currie, J. and J. Fahr (2005). Medicaid Managed Care: Effects on Children's Medicaid Coverage and Utilization. *Journal of Public Economics 89*(1), 85–108.

Curto, V., L. Einav, A. Finkelstein, J. Levin, and J. Bhattacharya (2019). Health Care Spending and Utilization in Public and Private Medicare. *American Economic Journal: Applied Economics 11*(2), 302–332.

Davenport, D. (2023). Discriminatory Discretion: Theory and Evidence From Use of Pretrial Algorithms. https://papers.ssrn.com/abstract=4580390.

Dillender, M. (2018). What Happens When the Insurer Can Say No? Assessing Prior Authorization as a Tool to Prevent High-Risk Prescriptions and to Lower Costs. *Journal of Public Economics 165*, 170–200.

Doyle, J., J. Graves, and J. Gruber (2017). Uncovering Waste in US Healthcare: Evidence from Ambulance Referral Patterns. *Journal of Health Economics 54*, 25–39.

Duggan, M., J. Gruber, and B. Vabson (2018). The Consequences of Health Care Privatization: Evidence from Medicare Advantage Exits. *American Economic Journal: Economic Policy 10*(1), 153–86.

Duggan, M. and T. Hayford (2013). Has the Shift to Managed Care Reduced Medicaid Expenditures? Evidence from State and Local-Level Mandates. *Journal of Policy Analysis and Management 32*(3), 505–535.

Eliason, P. J., R. J. League, J. Leder-Luis, R. C. McDevitt, and J. W. Roberts (2021). Ambulance Taxis: The Impact of Regulation and Litigation on Health Care Fraud. https://www.nber.org/papers/w29491.

Grimon, M.-P. and C. Mills (2022). The Impact of Algorithmic Tools on Child Protection: Evidence from a Randomized Controlled Trial. https://www.dropbox.com/s/g9d5dqbztd41o9l/ChrisMills_JMP.pdf?dl=0&e=1.

Gruber, J., B. R. Handel, S. H. Kina, and J. T. Kolstad (2020). Managing intelligence: Skilled experts and ai in markets for complex products." working paper. working paper series. https://www.nber.org/papers/w27038.

Gruber, J. and R. McKnight (2016). Controlling Health Care Costs through Limited Network Insurance Plans: Evidence from Massachusetts State Employees. *American Economic Journal: Economic Policy 8*(2), 219–250.

Huckfeldt, P. J., J. J. Escarce, B. Rabideau, P. Karaca-Mandic, and N. Sood (2017). Less intense postacute care, better outcomes for enrollees in medicare advantage than those in fee-for-service. *Health Affairs 36*(1), 91–100.

Huckfeldt, P. J., V. Shier, J. J. Escarce, B. Rabideau, T. Boese, H. M. Parsons, and N. Sood (2024). Postacute Care for Medicare Advantage Enrollees Who Switched to Traditional Medicare Compared With Those Who Remained in Medicare Advantage. *JAMA Health Forum 5*(2), e235325.

Jayapal, P. (2023). Letter to Joseph Biden. https://jayapal.house.gov/wp-content/uploads/2024/03/2024-Medicare-Advantage-Letter-Final.pdf.

Jung, J., C. Carlin, and R. Feldman (2024). Medicare Advantage Has Lower Resource Use and Better Quality of Care than Traditional Medicare. *American Journal of Health Economics*.

Kaiser Family Foundation (2024a). Medicare Advantage in 2024: Enrollment Update and Key Trends. https://www.kff.org/medicare/issue-brief/medicare-advantage-in-2024-enrollment-update-and-key-trends/.

Kaiser Family Foundation (2024b). Medicare Advantage in 2024: Premiums, Out-of-Pocket Limits, Supplemental Benefits, and Prior Authorization. https://www.kff.org/medicare/issue-brief/medicare-advantage-in-2024-premiums-out-of-pocket-limits-supplemental-benefits-and-prior-authorization/.

Keohane, L. M., R. C. Grebla, M. Rahman, D. B. Mukamel, Y. Lee, V. Mor, and A. Trivedi (2017). First-dollar cost-sharing for skilled nursing facility care in medicare advantage plans. *BMC Health Services Research 17*(1), 611.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2018). Human Decisions and Machine Predictions. *The Quarterly Journal of Economics 133*(1), 237–293.

Kuziemko, I., K. Meckel, and M. Rossin-Slater (2018). Does Managed Care Widen Infant Health Disparities? Evidence from Texas Medicaid. *American Economic Journal: Economic Policy 10*(3), 255–283.

Layton, T. J., N. Maestas, D. Prinz, and B. Vabson (2022). Health Care Rationing in Public Insurance Programs: Evidence from Medicaid. *American Economic Journal: Economic Policy 14*(4), 397–431.

Ludwig, J., S. Mullainathan, and A. Rambachan (2024). The Unreasonable Effectiveness of Algorithms. https://www.nber.org/papers/w32125.

MacKinnon, J. G. and M. D. Webb (2018). The Wild Bootstrap for Few (Treated) Clusters. *The Econometrics Journal 21*(2), 114–135.

Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler, and A. Leibowitz (1987). Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *American Economic Review 77*(3), 251–277.

McGarry, B. E., D. C. Grabowski, L. Ding, and J. M. McWilliams (2021). Outcomes After Shortened Skilled Nursing Facility Stays Suggest Potential For Improving Postacute Care Efficiency. *Health Affairs 40*(5), 745–753.

McKinsey & Company (2021). Administrative simplification: How to Save a Quarter-Trillion Dollars in US Healthcare. https://www.mckinsey.com/industries/healthcare/our-insights/administrative-simplification-how-to-save-a-quarter-trillion-dollars-in-us-healthcare.

McWilliams, J. M., L. G. Gilstrap, D. G. Stevenson, M. E. Chernew, H. A. Huskamp, and D. C. Grabowski (2017). Changes in Post-acute Care in the Medicare Shared Savings Program. *JAMA Internal Medicine 177*(4), 518–526.

Medicare Payment Advisory Commission (2024). March Report to Congress: Chapter 6. https://www.medpac.gov/wp-content/uploads/2024/03/Mar24_Ch6_MedPAC_Report_To_Congress_SEC.pdf.

Meyers, D. J., V. Mor, and M. Rahman (2018). Medicare Advantage Enrollees More Likely To Enter Lower-Quality Nursing Homes Compared To Fee-For-Service Enrollees. *Health Affairs 37*(1), 78–85.

Morrisey, M. A., F. A. Sloan, and J. Valvona (1988). Medicare Prospective Payment and Posthospital Transfers to Subacute Care. *Medical Care 26*(7), 685.

Mullainathan, S. and Z. Obermeyer (2022). Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care. *The Quarterly Journal of Economics 137*(2), 679–727.

NaviHealth (2018). Case Study: A Nationally Recognized Non-Profit Health Plan Improved Its Already Below-Average Post-Acute Utilization. https://web.archive.org/web/20230408153827/https://navihealth.com/wp-content/uploads/naviHealth-PAC-Management-Case-Study.pdf.

NaviHealth (2019). Medicare Advantage Post-Acute Care Program. https://web.archive.org/web/20220401091849/https://partner-resources.navihealth.com/wp-content/uploads/BCBS-MI-SNF-Provider-Forum.pdf.

Obermeyer, Z., B. Powers, C. Vogeli, and S. Mullainathan (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science 366*(6464), 447–453.

Pauly, M. (1968). The Economics of Moral Hazard: Comment. *American Economic Review 58*(3), 531–537.

Prusynski, R. A., A. D'Alonzo, M. P. Johnson, T. M. Mroz, and N. E. Leland (2024). Differences in Home Health Services and Outcomes Between Traditional Medicare and Medicare Advantage. *JAMA Health Forum 5*(3), e235454.

Rahman, M., D. J. Meyers, and V. Mor (2018). The Effects of Medicare Advantage Contract Concentration on Patients' Nursing Home Outcomes. *Health Services Research 53*(6), 4087–4105.

Regenbogen, S. E., A. H. Cain-Nielsen, J. D. Syrjamaki, L. M. Chen, and E. C. Norton (2019). Spending On Postacute Care After Hospitalization In Commercial Insurance And Medicare Around Age Sixty-Five. *Health Affairs 38*(9), 1505–1513.

Roodman, D., M. O. Nielsen, J. G. MacKinnon, and M. D. Webb (2019). Fast and Wild: Bootstrap Inference in Stata using boottest. *The Stata Journal 19*(1), 4–60.

Ross, C. and B. Herman (2023a). Denied by AI: How Medicare Advantage Plans Use Algorithms to Cut Off Care for Seniors in Need. https://www.statnews.com/2023/03/13/medicare-advantage-plans-denial-artificial-intelligence/.

Ross, C. and B. Herman (2023b). Humana used Algorithm in 'Fraudulent Scheme' to Deny Care to Medicare Advantage Patients, Lawsuit Alleges. https://www.statnews.com/2023/12/12/humana-algorithm-medicare-advantage-patients-lawsuit/.

Ross, C. and B. Herman (2023c). UnitedHealth Faces Class Action Lawsuit over Algorithmic Care Denials in Medicare Advantage Plans. https://www.statnews.com/2023/11/14/unitedhealth-class-action-lawsuit-algorithm-medicare-advantage/.

Schwartz, M. L., C. M. Kosar, T. M. Mroz, A. Kumar, and M. Rahman (2019). Quality of Home Health Agencies Serving Traditional Medicare vs Medicare Advantage Beneficiaries. *JAMA Network Open 2*(9), e1910622.

Skopec, L., P. J. Huckfeldt, D. Wissoker, J. Aarons, J. Dey, I. Oliveira, and S. Zuckerman (2020). Home Health And Postacute Care Use In Medicare Advantage And Traditional Medicare. *Health Affairs 39*(5), 837–842. Publisher: Health Affairs.

Sloan, C., G. Naufal, and H. Caspers (2023). The Effect of Risk Assessment Scores on Judicial Behavior and Defendant Outcomes. *Journal of Human Resources*.

Stevenson, M. T. (2018). Assessing Risk Assessment in Action. https://papers.ssrn.com/abstract=3016088.

Stevenson, M. T. and J. L. Doleac (2022). Algorithmic Risk Assessment in the Hands of Humans. https://papers.ssrn.com/abstract=3489440.

Wallace, J. (2023). What Does a Provider Network Do? Evidence from Random Assignment in Medicaid Managed Care. *American Economic Journal: Economic Policy 15*(1), 473–509.

Werner, R. M., N. Coe, M. Qi, and R. T. Konetzka (2023). The Value of an Additional Day of Post-acute Care in a Skilled Nursing Facility. *American Journal of Health Economics*.

Wing, C., S. M. Freedman, and A. Hollingsworth (2024). Stacked Difference-in-Differences. https://www.nber.org/papers/w32054.

Appendix A: Details on Data Construction

1 Sample

1.1 Data on Hospitalizations

I use 100% MedPAR from 2015-2020. This long-time span of data is needed for the analysis of partnerships in 2017 (discussed in Appendix C.), the exclusion of hospitalizations that occurred soon after a past hospitalization, and the calculation of the 90-day readmissions outcome.

I include only records from short-term acute care hospitals and critical access hospitals, as indicated in the Medicare provider number.[5] I exclude stays that occurred in a specialty unit (e.g., psychiatric or rehabilitation unit).[6] I exclude admissions that occur within 90 days of the last discharge.[7] I exclude patients who were not enrolled in both Part A and Part B during the month of the hospital discharge, according to the MBSF. I exclude patients who had begun hospice care prior to their hospitalization, according to hospice claims from 2016-2019, which contain information on both TM and MA patients.

1.2 Exclusions

Using 2019 data, I make two BCBS MI specific exclusions. First, I include only those living in Michigan. Despite its name, BCBS MI does have some patients who live outside of Michigan, though the vast majority live in Michigan. Second, I include only patients treated at hospitals with at least 11 BCBS MI and 11 TM patients. Table A1 shows the number of hospitalizations excluded at each step. The result of this step is the data I use for analysis.

[5] https://resdac.org/cms-data/variables/MedPAR-provider-number

[6] https://resdac.org/cms-data/variables/provider-number-special-unit-code

[7] Before to this exclusion, however, I use this set of stays to determine whether a readmission occurred within 90 days of the index hospitalization.

Table A1. Sample Exclusions

Exclusion	TM	BCBS MI	Total
Outside Michigan	5,933,698	11,760	5,945,458
Hospital treats few TM or BCBS MI patients	8,983	1,254	10,237
Final Sample	211,155	73,691	284,846

Abbreviations: TM, traditional Medicare; BCBS MI, Blue Cross Blue Shield of Michigan. *Source:* Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

1.3 Outcomes

I follow patients through their index hospitalization and discharge, their admission to a SNF (where applicable), and their health outcomes at 90 days after discharge. First, I use the MedPAR discharge destination code to determine where the hospital discharged patient to.[8] I categorize patients as discharged to one of the following mutually exclusive and exhaustive categories: (1) home, no post-acute care (code 1); (2) SNF (codes 3 and 83); (3) home health (codes 6 and 86); (4) inpatient rehabilitation (codes 62 and 90) (5) died or hospice (codes 20, 50, 51, or a death date during the hospital stay according to the MBSF); (6) other. I also use the hospital length of stay from the MedPAR record. Second, I examine patients use of SNFs, which is discussed at greater length below. Third, I calculate 90-day readmission and mortality. I use MedPAR to obtain readmission data for acute care hospitals and use the MBSF to determine mortality. I exclude patients who die during the hospital stay from both measures. I also exclude patients from the readmissions measure denominator when their hospitalization ended in transfer to another hospital (codes 2, 82, 66, 94 or a MedPAR acute care admission on the day of the index discharge).

1.4 Covariates

I obtain enrollment as of the month of hospital discharge from the MBSF and link this to publicly available MA data to determine enrollment in BCBS MI.[9] In addition, I obtain the patients age, sex, race/ethnicity, dual eligibility at discharge, and date of death (where applicable) from the MBSF.

I obtain the diagnosis related group (DRG), a count of diagnoses, and a flag for ICU use from the MedPAR hospitalization record. In addition, I use the diagnosis codes on the

[8] https://resdac.org/cms-data/variables/destination-upon-discharge-facility-code

[9] https://www.cms.gov/data-research/statistics-trends-and-reports/medicare-advantagepart-d-contract-and-enrollment-data/monthly-enrollment-plan

hospitalization record to determine the presence of chronic conditions using the Chronic Conditions Warehouse 30 condition algorithm.[10] For descriptive purposes, I also group DRGs into medical or surgical hospitalizations based on the CMS definition and into four common specific conditions.[11,12]

2 Skilled Nursing Facility Stay Information

I used two data sources to obtain information on SNF stays: (1) MA encounter data and (2) MedPAR. MedPAR contains the universe of TM SNF stays as well as some, though not all, MA SNF stays. Therefore, the MA encounter data is needed for complete coverage of all SNF stays.

2.1 Cleaning Medicare Advantage Encounter Data

There are two major challenges to using the MA encounter data. First, the encounter data contains an organizational National Provider Identifier (NPI) but not a CMS certification number (CCN). Fortunately, MedPAR SNF records contain both. I therefore use these records to create a crosswalk of organizational NPI to CCN, allowing for multiple NPI to map to a single CCN. I do this separately for each year to reflect any annual changes. This method allows me to identify the CCNs on 95.9% of 2016-2019 MA encounter records.

Second, unlike MedPAR, there are often multiple MA encounters per stay. For example, a stay in June and July may contain two separate encounters for the June and July parts of the stay. I combine all encounters that overlap or are separated by zero or one day (i.e., an encounter that starts on the day or day after the last encounter ended) to begin on the earliest recorded date and end on the latest recorded date. I use the admission and discharge dates on the encounter where available to combine these stays, replacing the admission (discharge) date with the claim from (thru) date where it is not. I exclude chart review records and a small number of records where the admission date is after the discharge date. I combine stays across multiple providers because (1) I don't accurately observe the provider for all encounters (i.e., where I'm unable to determine CCN) and (2) this captures the full length of the episode, even in the small number of cases where a patient is observed at multiple facilities (2.2% of 2016-2019 stays). In cases where multiple

[10] 7/2023 version: https://www.cms.gov/medicare/coordination-benefits-recovery/overview/icd-code-lists

[11] Medical and surgical conditions are detailed here: https://www.cms.gov/icd10m/version372-fullcode-cms/fullcode_cms/P0371.html

[12] The conditions are septicemia (870, 871, 872), joint replacement (469, 470), stroke (61, 62, 63, 64, 65, 66), and heart failure (291, 292, 293). All other DRGs are classified as other.

facilities are part of the stay, the priority for which CCN is used is determined by CCN if observed for any of the encounters in the stay, then the earlier admission date, then the latest discharge date, then randomly.

I adopt the MedPAR convention for calculating length of stay: subtracting the admission date from the discharge date, replacing the number of days as 1 if the admission and discharge dates are the same.[13]

2.2 Cleaning MedPAR Data

The key goal of cleaning the MedPAR data is to ensure comparability between it and the MA encounter data. One issue is that a non-trivial share (3.0%) of records have lengths of stay over 100 days, which is implausible given the design of the Medicare SNF benefit. Indeed, many of these stays have significantly fewer Medicare paid days—a variable that in most cases is equal to the length of stay.[14]

I therefore primarily use the Medicare paid days to calculate the MedPAR SNF length of stay. In sensitivity analyses, I show that results are unchanged when using the main MedPAR length of stay variable to calculate the length of stay (see Figure B2).[15] To match MA encounter data, I combine stays that occur within 1 day into a single stay, regardless of where the stay occurs.[16] I use the record with the earliest admission date to obtain the CCN.

2.3 Obtaining Skilled Nursing Facility Information

I obtain 2019 facility information from nursing home compare (profit status, five-star rating) and LTCFocus.org (latitude/longitude). I calculate the distance, in miles, from the patient's ZIP code tabulation centroid to the SNF.[17,18] I exclude distances over 100 miles from analysis of this variable.

[13] https://resdac.org/cms-data/variables/days-beneficiarys-stay-hospitalsnf

[14] https://resdac.org/cms-data/variables/covered-days-care-chargeable-medicare-utilization-stay

[15] One additional issue is that some stays have zero Medicare paid days, which is also implausible given that MedPAR only includes Medicare paid stays. I exclude stays with zero days when using the Medicare paid variable, though these stays are included when using the alternative method in sensitivity analyses.

[16] When using the Medicare days, I calculate the end date as the admission date plus the Medicare days before this process. In sensitivity analyses when using the full length of stay, I calculate the end date as the admission date plus full length of stay (where it is missing).

[17] https://data.nber.org/distance/2019/centroid/

[18] https://udsmapper.org/zip-code-to-zcta-crosswalk/

2.4 Linking to Hospitalizations

I link hospitalizations to SNF stays that began the day of or the day after hospital discharge. If information is available from both the encounter data and MedPAR, I use the information from MedPAR. I exclude a small number of stays that are over 100 days. I also exclude hospitalizations that ended in December from analyses of SNF choice and length of stay.[19]

2.5 Examining Data Quality

There are concerns about the completeness of the MA encounter data. To examine the completeness of the encounter data, I calculate the share of hospitalizations with a SNF discharge setting that actually have a SNF stay. In Table A2, I show that the rates are comparable in BCBS MI and TM. I show this overall but also show it for DRG 470 to demonstrate that the similarity of this measure for these two groups is not driven by the composition of DRGs, which have varied levels of SNF follow-up. These results emphasize the completeness of the encounter data for BCBS MI.

Table A2. Percent of Hospitalizations with Skilled Nursing Facility Discharge Setting where a Skilled Nursing Facility is Observed

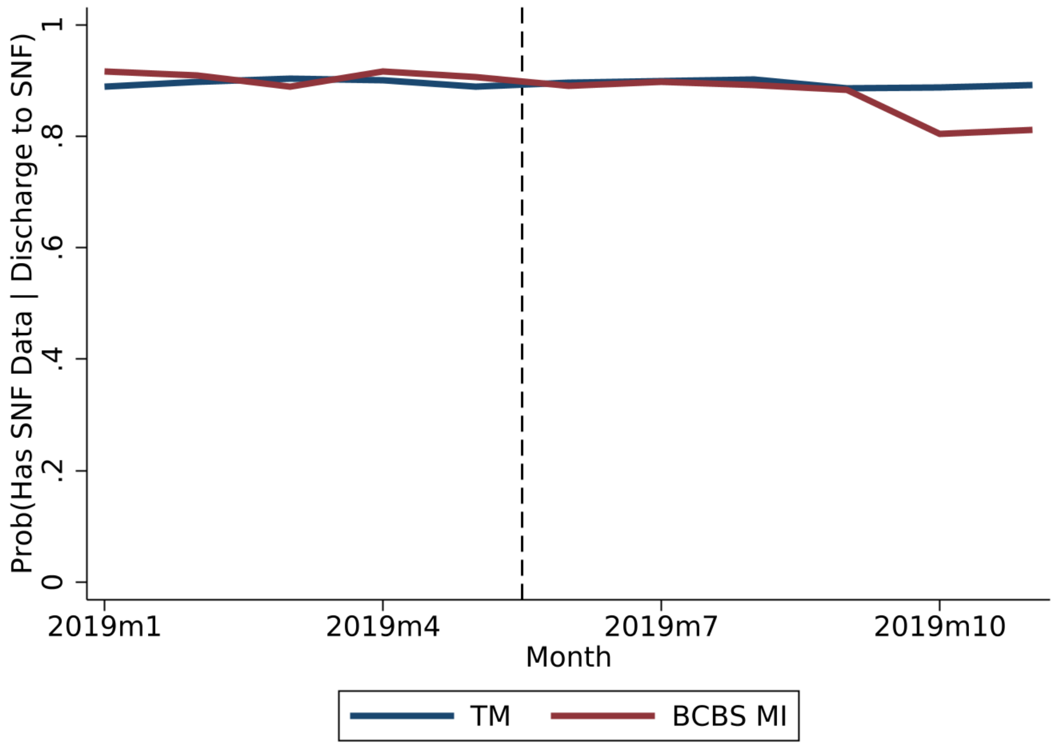
Insurer	All DRG	DRG 470
TM	89.5	95.5
BCBS MI	88.3	94.2

Abbreviations: SNF, skilled nursing facility; TM, traditional Medicare; BCBS MI, Blue Cross Blue Shield of Michigan; DRG, diagnosis related group.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

I also examined this measure of data completeness over time. There is a modest decline in this measure beginning in October (Figure A1). I interpret this as a modest decline in the completeness of the encounter data, potentially because some SNF stays are not recorded. However, the rate remains relatively high, suggesting that while data from this period is imperfect, it is still relatively complete. In a sensitivity analysis, I show that estimates are very similar when October and November are excluded (see Figure B2).

[19] One issue is that encounters are sorted into annual files by their end date. As a result, SNF stays ending late in the year may not be fully observed. I exclude December hospital discharges to account for this issue.

Figure A1. Proportion of Hospitalizations with Skilled Nursing Facility Discharge Setting where a Skilled Nursing Facility is Observed



Abbreviations: SNF, skilled nursing facility; TM, traditional Medicare; BCBS MI, Blue Cross Blue Shield of Michigan.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Appendix B: Supplemental Exhibits

Table B1. Outcome Means

Outcome	Pre		Post	
	TM	BCBS MI	TM	BCBS MI
Discharge Setting				
Skilled Nursing Facility	18.2	18.4	17.5	17.6
Home without Post-Acute Care	51.5	50.4	52.8	51.6
Home Health	18.3	21.1	17.9	20.8
Inpatient Rehabilitation Facility	3.1	2.3	3.0	2.4
Died or Hospice	4.3	4.9	4.1	4.6
Other	4.6	2.9	4.7	3.0
SNF Choice				
SNF Five-Star Rating	3.9	4.1	4.0	4.1
Distance to SNF	8.6	8.2	8.5	8.2
For-Profit SNF	68.1	64.2	67.2	65.1
SNF Stay				
SNF Length of Stay	23.5	18.4	23.6	16.2
SNF Stay > 30 Days	23.1	12.8	23.3	5.9
Outcomes				
90-Day Readmissions (full sample)	26.3	23.5	26.0	23.0
90-Day Mortality (full sample)	8.3	8.6	6.7	7.1
90-Day Readmissions (SNF users)	32.8	32.0	31.9	31.3
90-Day Mortality (SNF users)	12.9	12.6	11.1	11.8

Abbreviations: SNF, skilled nursing facility pp, percentage point.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Table B2. Effect of NaviHealth Partnership on Health Care Use and Outcomes

Outcome	Effect	95% Confidence Interval	Baseline Mean	N
Skilled Nursing Facility	-0.217	[-0.892, 0.458]	18.4	284,841
Home without Post-Acute Care	0.136	[-0.887, 1.159]	50.4	284,841
Home Health	-0.026	[-0.840, 0.788]	21.1	284,841
Inpatient Rehabilitation Facility	0.200	[-0.061, 0.461]	2.3	284,841
Died or Hospice	-0.017	[-0.320, 0.286]	4.9	284,841
Other	-0.076	[-0.338, 0.187]	2.9	284,841
SNF Five-Star Rating	-0.067**	[-0.107, -0.027]	4.1	42,892
Distance to SNF	0.074	[-0.408, 0.555]	8.2	43,026
For-Profit SNF	1.346	[-0.960, 3.652]	64.2	43,576
SNF Length of Stay	-2.320***	[-2.917, -1.723]	18.4	44,425
SNF Stay > 30 Days	-7.137***	[-8.610, -5.664]	12.7	44,425
90-Day Readmissions (full sample)	-0.168	[-0.774, 0.439]	23.5	273,604
90-Day Mortality (full sample)	0.106	[-0.291, 0.502]	8.6	278,498
90-Day Readmissions (SNF users)	0.254	[-1.614, 2.121]	32.0	44,353
90-Day Mortality (SNF users)	0.692	[-0.591, 1.975]	12.5	44,425

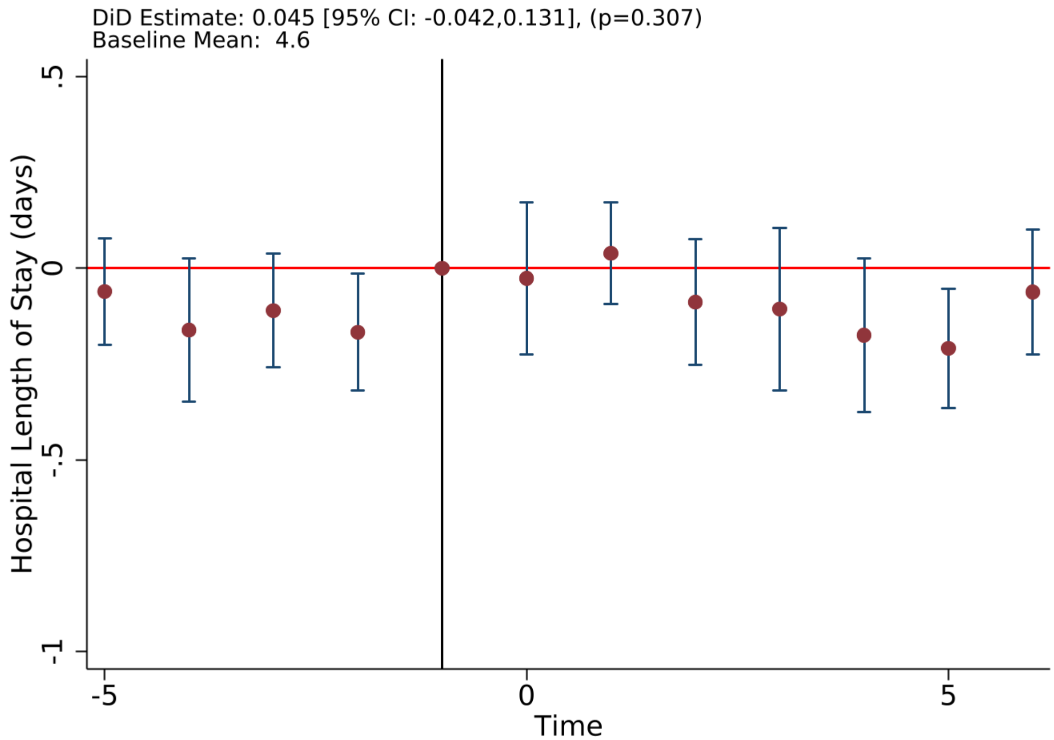
Notes: Entries are difference-in-differences estimates from ordinary least squares regressions. Only patients with non-missing values for the given outcome are included. Estimates are adjusted for covariates. A small number of singletons (due to diagnosis related groups with few patients) are excluded from each regression. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Only patients with non-missing values for the given outcome are included. Estimates are adjusted for covariates. A small number of singletons (due to diagnosis related groups with few patients) are excluded from each regression. Standard errors were clustered at insurer-county level.

Abbreviations: BCBS MI, Blue Cross Blue Shield of Michigan; SNF, skilled nursing facility.

Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, the Medicare Master Beneficiary Summary File (MBSF), and publicly-available skilled nursing facility data.

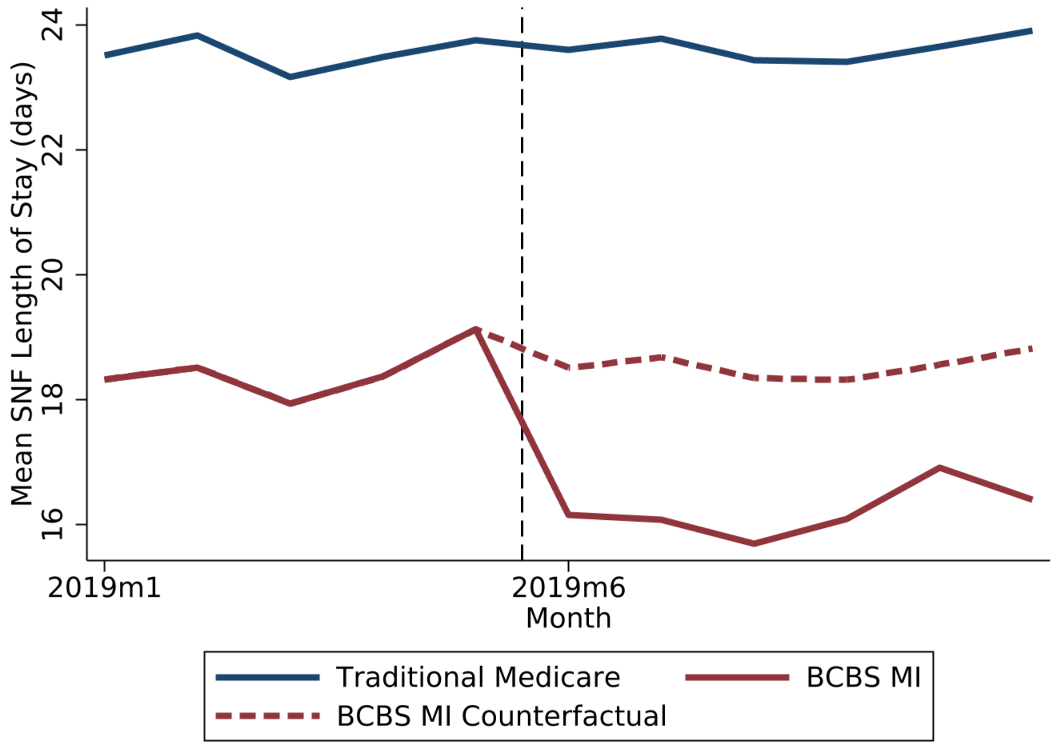
$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

Figure B1. Effect of NaviHealth Partnership on Hospital Length of Stay



Notes: This figure shows ordinary least squares event-study estimates of the effect of the Blue Cross Blue Shield of Michigan partnership with NaviHealth in June 2019. Time is relative to June 2019. There is a vertical May 2019, which is used as the reference month (time = -1). Pooled difference-in-differences estimated with covariates are shown above the figure. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Standard errors were clustered at insurer-county level.
Abbreviations: SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure B2. Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay



Notes: This figure shows the mean length of stay by month for traditional Medicare and Blue Cross Blue Shield of Michigan patients. Starting in June 2019, I calculate the Blue Cross Blue Shield of Michigan counterfactual as the traditional Medicare value in the given month minus the pre-period difference in means. The vertical line denotes the timing of the Blue Cross Blue Shield of Michigan partnership with NaviHealth.
Abbreviations: SNF, skilled nursing facility; BCBS MI, Blue Cross Blue Shield of Michigan.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Table B3. Effect of NaviHealth Partnership on the Distribution of Skilled Nursing Facility Length of Stay

Percentile	TM		BCBS MI		Effect	95% Confidence Interval
	Pre	Post	Pre	Post		
10	6	7	6	6	-1	[-2.488, 0.488]
25	12	12	11	11	0	[-1.545, 1.545]
50	20	20	17	15	-2***	[-2.804, -1.196]
75	29	29	24	20	-4***	[-5.670, -2.330]
90	45	45	33	27	-6***	[-8.293, -3.707]

Notes: This table shows the percentiles of the length of stay distributions for traditional Medicare and Blue Cross Blue Shield of Michigan patients before and after June 1, 2019. Differences difference-in-differences estimates were obtained using simultaneous a quantile regression, where length of stay was regressed on a BCBS MI indicator, a post June 2019 indicator, and the interaction of these variables, which is the effect estimate that is shown. Bootstrapped 95% confidence intervals were obtained using 20 bootstrap replications.

Abbreviations: TM, traditional Medicare; BCBS MI, Blue Cross Blue Shield of Michigan. *Source:* Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

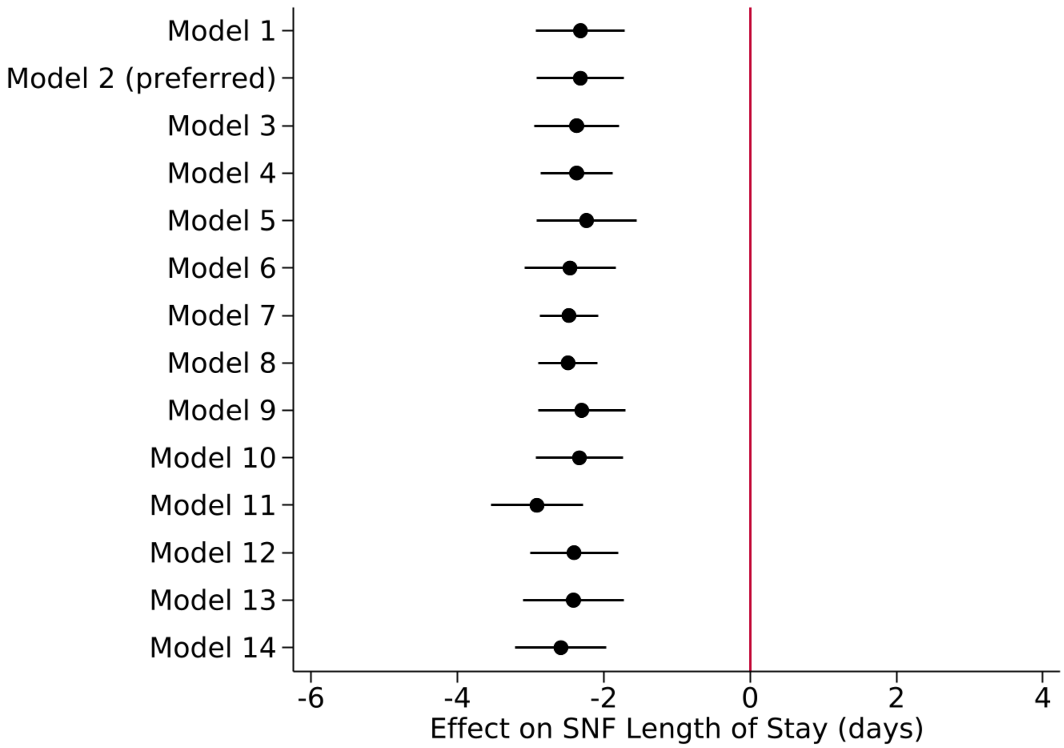
$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

Table B4. Robustness of the Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay using Alternative Methods of Obtaining Confidence Intervals/P-Values

	SNF Length of Stay		**SNF Stay >30 Days**	
Estimate	-2.317	-2.320	-7.137	-7.081
Clustered Insurer-County	[-2.926, -1.708]	[-2.917, -1.722]	[-8.528, -5.635]	[-8.611, -5.663]
	<.001	<.001	<.001	<.001
Robust	[-2.856, -1.778]	[-2.853, -1.787]	[-8.486, -5.677]	[-8.536, -5.739]
	<.001	<.001	<.001	<.001
Clustered Insurer (Analytic)	[-2.340, -2.294]	[-2.676, -1.963]	[-7.097, -7.065]	[-7.265, -7.009]
	<.001	0.008	<.001	0.001
Clustered Insurer (Wild Bootstrap)	[-4.168, -.2338]	[-6.727, 1.259]	[-8.178, -6.026]	[-8.501, -5.912]
	0.047	0.075	0.008	0.004
Covariates		X		X

Notes: Entries are difference-in-differences estimates from ordinary least squares regressions. Each row below the estimate shows confidence intervals and p-values obtained using different methods. Models are estimated with and without covariates: age, race/ethnicity, sex, diagnosis related group, intensive care unit use, diagnosis count, and 30 indicators for chronic conditions calculated using the diagnoses on the hospitalization record.
Abbreviations: SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure B3. Robustness of the Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay



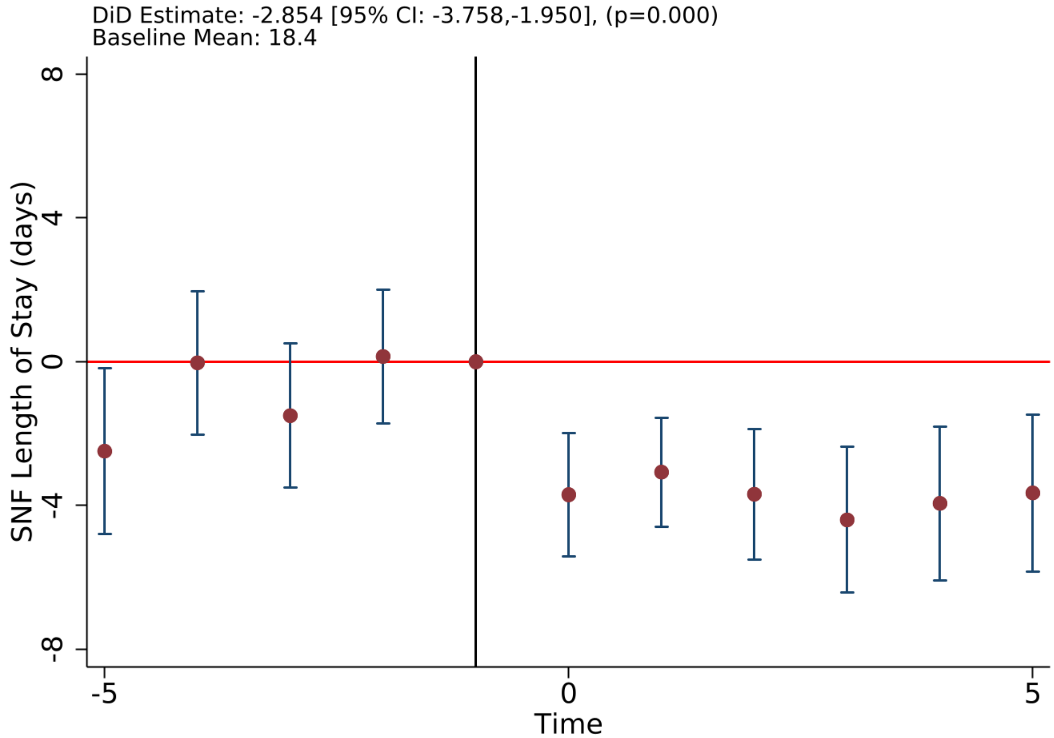
Notes: The models are as follows: Model 1: Equation (1) without covariates
Model 2: Equation (1)
Model 3: Equation (1) plus hospital fixed effect (FE)
Model 4: Equation (1) plus hospital FE, standard errors (SE) clustered at hospital
Model 5: Equation (1), non-duals only
Model 6: Equation (1), first observed hospitalization only
Model 7: Equation (1), not limiting to Michigan
Model 8: Equation (1) plus hospital FE, not limiting to Michigan
Model 9: Equation (1), limiting to Michigan but not restricting set of included hospitals
Model 10: Equation (1) plus hospital FE, limiting to Michigan but not restricting set of included hospitals
Model 11: Equation (1), estimated using Poisson regression (average marginal effect)
Model 12: Equation (1), excluding October and November discharges (see appendix A for discussion of modestly reduced data completeness during these months)
Model 12: Equation (1), using alternative definition of MedPAR SNF length of stay (see appendix A) top-coding at 100 days
Model 14: Equation (1), using alternative definition of MEDPAR SNF length of stay (see appendix A) excluding stays over 100 days
Abbreviations: SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Table B5. Effect of NaviHealth on Length of Stay by Predicted Length of Stay Quartile

Quartile	TM		BCBS MI		Estimate	95% CI
	Pre	Post	Pre	Post		
Panel A: SNF Length of Stay						
1	20.0	20.0	16.5	14.7	-1.691***	[-2.463, -0.919]
2	22.0	21.9	18.0	15.8	-1.988***	[-2.817, -1.160]
3	24.0	24.0	19.4	16.7	-2.642***	[-3.910, -1.374]
4	27.3	27.9	21.1	18.8	-2.766***	[-4.105, -1.428]
Panel B: SNF Stay > 30 Days						
1	15.7	14.9	8.5	3.4	-4.390***	[-6.246, -2.533]
2	19.6	20.0	11.2	5.1	-6.388***	[-8.694, -4.082]
3	24.0	23.8	15.4	7.7	-7.459***	[-10.406, -4.511]
4	31.5	33.0	18.7	8.9	-11.276***	[-15.281, -7.271]

Notes: Patients who were admitted to a skilled nursing facility were stratified by their predicted probability of having a skilled nursing facility stay over 30 days. This predicted probability was obtained by regressing an indicator for having a stay over 30 days on covariates (age, race/ethnicity, sex, diagnosis related group, intensive care unit use, diagnosis count, 30 indicators for chronic conditions calculated using the diagnoses on the hospitalization record, and hospital length of stay) for the sample of traditional Medicare skilled nursing facility users and applying the predicted probability from this model to the entire sample of patients admitted to a skilled nursing facility. Patient were then divided into quartiles based on this predicted probability. Values are the mean skilled nursing facility length of stay or percent with a stay over 30 days for each category. Effects are the difference-in-differences estimates for each subgroup. Estimates in each row are based on a separate ordinary least squares regressions with covariates. Standard errors were clustered at insurer-county level.
Abbreviations: BCBS MI, Blue Cross Blue Shield of Michigan; SNF, skilled nursing facility; CI, Confidence interval.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).
$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

Figure B4. Robustness of the Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay using Humana in Michigan as the Comparison Group



Notes: This figure shows ordinary least squares event-study estimates of the effect of the Blue Cross Blue Shield of Michigan partnership with NaviHealth in June 2019. Unlike in the main analysis, the comparison group consists of Humana patients living in Michigan. I also exclude patients that are treated at a hospital with fewer than 11 Humana or 11 Blue Cross Blue Shield of Michigan patients. Time is relative to June 2019. There is a vertical May 2019, which is used as the reference month (time = -1). Pooled difference-in-differences estimated with covariates are shown above the figure. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Only patients admitted to a skilled nursing facility were included. Standard errors were clustered at insurer-county level.
Abbreviations: SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Table B6. Robustness of the Effect of NaviHealth Partnership on Skilled Nursing Facility Length of Stay using Humana in Michigan as the Comparison Group

Outcome	Effect	95% Confidence Interval	Baseline Mean	N
SNF	-0.521	[-1.527, 0.486]	18.6	87,887
Home without Post-Acute Care	0.224	[-1.544, 1.992]	50.0	87,887
Home Health	0.520	[-1.049, 2.088]	21.1	87,887
Inpatient Rehabilitation Facility	-0.317	[-0.875, 0.240]	2.4	87,887
Died or Hospice	-0.319	[-0.864, 0.225]	5.0	87,887
Other	0.414	[-0.219, 1.046]	2.9	87,887
SNF Five-Star Rating	-0.083	[-0.213, 0.047]	4.1	13,035
Distance to SNF	0.315	[-0.528, 1.159]	8.2	13,098
SNF For-Profit	2.165	[-2.056, 6.386]	64.1	13,215
SNF Length of Stay	-2.854***	[-3.762, -1.946]	18.4	13,446
SNF Stay > 30 Days	-6.906***	[-9.274, -4.537]	12.8	13,347
90-Day Readmissions (full sample)	-0.094	[-1.429, 1.242]	23.6	84,324
90-Day Mortality (full sample)	-0.416	[-1.322, 0.490]	8.6	85,735
90-Day Readmissions (SNF users)	1.806	[-2.393, 6.005]	32.0	13,434
90-Day Mortality (SNF users)	0.090	[-2.557, 2.737]	12.5	13,446

Notes: Entries are difference-in-differences estimates from ordinary least squares regressions. Unlike in the main analysis, the comparison group consists of Humana patients living in Michigan. I also exclude patients that are treated at a hospital with fewer than 11 Humana or 11 Blue Cross Blue Shield of Michigan patients. Only patients with non-missing values for the given outcome are included. Estimates are adjusted for covariates. A small number of singletons (due to diagnosis related groups with few patients) are excluded from each regression. The baseline mean is the pre-June 2019 mean value for Blue Cross Blue Shield of Michigan patients. Standard errors were clustered at insurer-county level.
Abbreviations: BCBS MI, Blue Cross Blue Shield of Michigan; SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, the Medicare Master Beneficiary Summary File (MBSF), and publicly-available skilled nursing facility data.
$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

Appendix C: Analysis of Additional NaviHealth Partnerships

1 Overview

One limitation of this study is that it is based on a single insurer in one state. In this appendix, I analyze NaviHealth's partnership with two additional insurers, MVP and Horizon. While these insurers are smaller than BCBS MI, they both had publicly announced partnerships in 2017, during the period where I have data. I analyze each of these partnerships separately while also using the three staggered partnerships with NaviHealth and a stacked difference-in-differences design to generate pooled estimates. Doing so provides added confidence of the generalizability of my results beyond the case of the one insurer I examine in the main part of the paper.

2 Empirical Strategy

I observe three partnerships (see Table C1). I compare patients covered by each carrier in the listed state(s) to TM patients in the same state(s). I include discharges that occur the month of the partnership and the 5 months before and after the month of the partnership. I include only hospitals that treat more than 10 patients in both the treated carrier and in TM. The timing of each partnership, the number of hospitalizations, and the probability that patients discharged to a SNF are admitted to a SNF are shown in Table C1.[20] Note that these insurers are considerably smaller than BCBS MI. However, their SNF data appears comparable in completeness as TM patients in their state(s), emphasizing the feasibility of using their data for analysis.

[20] Information on these partnerships was obtained from the following websites that are no longer functioning. The MVP announcement can be found on the internet archive: https://web.archive.org/web/20221005105841/https://navihealth.com/wp-content/uploads/naviHealth_MVP_PR_FINAL.pdf. The Horizon announcement is no longer available but a PDF of the announcement can be provided upon request.

Table C1. Partnership Details

Carrier	Partnership Date	Partner		TM	
		N	Data Quality	N	Data Quality
BCBS MI	June 1, 2019	67,508	88.3	193,498	89.4
MVP	August 8, 2017	8,851	81.6	105,898	86.4
Horizon	February 1, 2017	8,150	90.8	173,674	91.5

Notes: Data quality is the percent of patients who were discharged to a skilled nursing facility who were admitted to skilled nursing facility. Note that the N and data quality measures for Blue Cross Blue Shield are slightly different in this appendix than they are in the main analysis because December 2019 is not included.
Abbreviations: SNF, skilled nursing facility; TM, traditional Medicare; BCBS MI, Blue Cross Blue Shield of Michigan.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).
$^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$

I first use a simple 2x2 difference-in-differences approach to estimate the effect of each partnership separately. One important detail to note is that announcements for both the BCBS MI and Horizon partnerships contain an "effective date" when the partnership began affecting patient care. The MVP announcement does not contain this level of detail; it contains only the date of the announcement. I use this as the basis for defining treatment timing.[21] I estimate the following model separately for each insurer and its TM comparison group:

$$y_{ict} = \alpha + \beta_1(\text{Post}_t \times \text{Treat}_c) + \mathbf{X_i} + \tau_t + \text{Treat}_c + \epsilon_{ict} \tag{C.3}$$

Where y_{isct} is an outcome observed for hospitalization i for carrier c at time t; Post_t is equal to 1 after the partnership; Treat_c is equal to 1 for the treated insurer; $\mathbf{X_i}$ is a vector of covariates (age, race/ethnicity, sex, DRG, ICU use, diagnosis count, and 30 indicators for chronic conditions calculated using the diagnoses on the hospitalization record); τ_t are month fixed effects. As in the main paper, β_1 is the parameter of interest, the average treatment effect on the treated (ATT). I cluster standard errors at the insurer-county level.

I then use a stacked difference-in-differences approach to obtain the pooled estimates (Baker and Wang., 2022; Wing et al., 2024). To do so, I create three "stacks" consisting of the 11 months of data for each carrier and their TM comparison group. For example, the Michigan stack includes both the BCBS MI and Michigan TM patients. Similarly,

[21] To be consistent with the monthly event study, I consider all MVP patients discharged in August or later to be treated.

the New Jersey and New York/Vermont stacks include the patients from those states. I then estimate the following event-study model:

$$y_{isct} = \alpha + \sum_{\substack{j=-5 \\ j\neq -1}}^{5} \beta_j I(t_{cj} - t_c^*) + \psi_{st} + \gamma_{sc} + \epsilon_{isct} \tag{C.4}$$

Where y_{isct} is an outcome observed for hospitalization i in stack s for carrier c at time t; $I(t_{cj} - t_c^*)$ are event-study indicators, which are always equal to zero for TM patients; ψ_{st} are stack by year-month fixed effects; γ_{sc} are stack by insurer fixed effects (i.e., TM patients have a separate fixed effect for each stack they are in); ϵ_{isct} is the error term. I cluster standard errors at the insurer-county level.

I also estimate the difference-in-differences equivalent to this equation:

$$y_{isct} = \alpha + \beta_1 \text{DiD}_{ict} + \mathbf{X_i} + \psi_{st} + \gamma_{sc} + \epsilon_{isct} \tag{C.5}$$

Where DiD_{ict} is equal to 1 after the partnership for the three MA carriers and zero otherwise.

3 Results

3.1 Effect on Skilled Nursing Facility Length of Stay

I first show the effect on SNF length of stay, as this is the main (non-null) result from the main analysis. The effect that I observed for BCBS MI is robust across the other two carriers (Figure C1). The MVP partnership with NaviHealth led to a 4.644 day decline in SNF length of stay (95% CI: -6.318,-2.970) while the Horizon partnership with NaviHealth led to decline of 3.852 days (95% CI: -4.978,-2.727). Note that both of these point estimates are modestly larger than the BCBS MI estimates, suggesting that if anything my use of BCBS MI for the main analysis may be conservative.

However, there are two minor issues with these results. First, there appears to be an anticipation effect of the MVP partnership. This could be interpreted as a violation of the parallel trends assumption. However, this trend is also consistent with the partnership going to effect during the month before the August announcement. Given the uncertainty in the effective date from the MVP announcement, the consistency in the magnitude of the change with the other NaviHealth partnerships, and the absence of effects along other margins at this time (see below), it is most likely that this is due to imprecision in my definition of the partnership timing. Second, Horizon experienced substantial enrollment growth in January, the month before the partnership with NaviHealth (see (Figure C2)). This shock could lead to a compositional change to the treatment group very close to when the partnership happened, which could be empirically problematic. To address this, I exclude the Horizon sample further to those that were enrolled in Horizon in January 2016. This reduces the sample size but does very little to the point estimates (Table C2), given that the levels of length of stay are similar between the existing and new Horizon enrollees (Figure C3). Overall, the evidence from these two partnerships emphasizes the generalizability of the substantial effect of NaviHealth on SNF length of stay beyond the BCBS MI partnership.

3.2 Effect on Other Measures of Health Care Use and Health Outcomes

I next examine other measures of health care use and health outcomes that were unchanged by the Blue Cross Blue Shield partnership. Similar to that partnership, I find no effects on the probability of discharge to a SNF—either for the two separate partnerships I examine or for the pooled estimate (Table C3). Also like the analysis in the main part of the paper, I find no effects on 90-day mortality or readmissions.

Table C3 also shows the estimates on the probability of having a SNF stay over 30 days. Each partnership had large effects on this outcome, as for BCBS MI in the main analysis. This emphasizes the role that NaviHealth consistently plays in reducing the variability of SNF stay length.

There are two additional statistically significant effects in Table C3 that are worth noting. The MVP partnership reduced the probability of being discharged to the "other" setting while the Horizon partnership decreased the probability of being discharged to the home health setting. Both of these estimates are relatively modest and imprecise (both are only statistically significant at the 5% level).

4 Appendix Conclusion

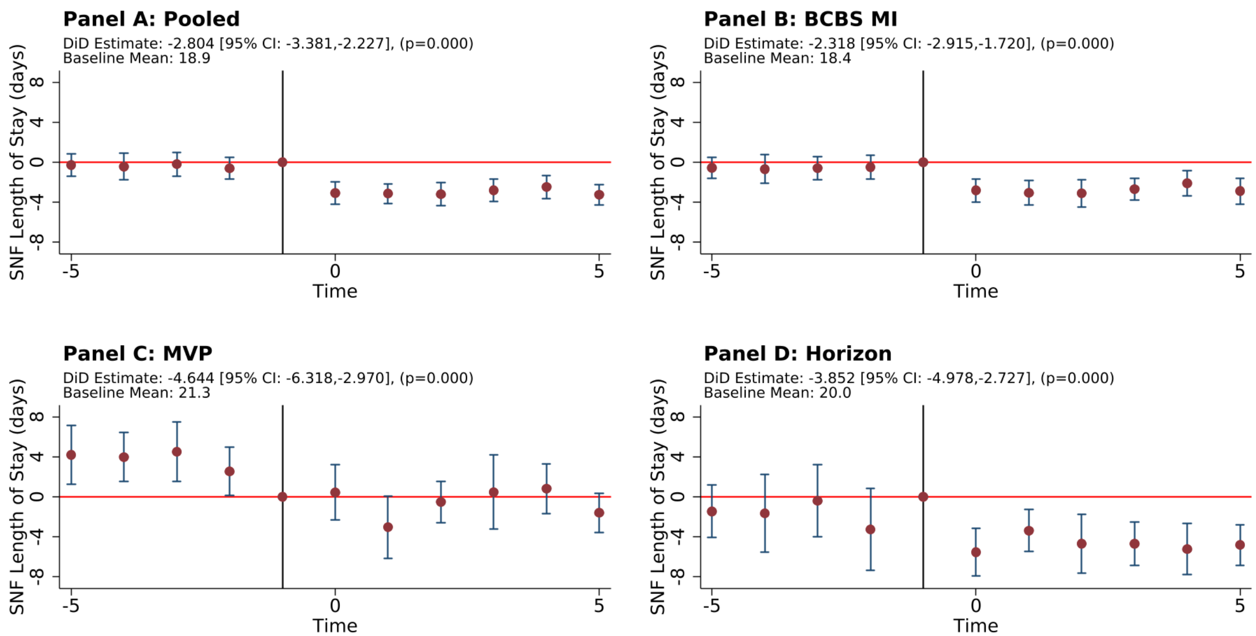
In this appendix, I present evidence from two other partnerships of Medicare Advantage insurers with NaviHealth. Each of these partnerships presents challenges. First, both insurers are small, so I observe fewer than 9,000 hospitalizations each compared to over 60,000 for Blue Cross Blue Shield of Michigan. Second, MVP does not provide the effective date of the partnership and my event-study results suggest that the partnership may have begun before it was publicly announced, though I have no concrete evidence of this. Third, Horizon experienced substantial enrollment growth the month before it partnered with

NaviHealth. Nonetheless, these results provide valuable insight into the generalizability of the results in the main paper exploiting the Blue Cross Blue Shield Partnership.

The NaviHealth partnerships with MVP and Horizon echo the three key findings of the main paper. First, there was no effect on the probability of discharge to a SNF. Second, there were large and immediate changes to SNF length of stay. This was driven in part by large declines in longer SNF stays over 30 days, suggesting that NaviHealth reduces not only the average length of stay but also the variability in length of stay. Finally, I do not observe any effects on readmissions or mortality. This suggests that the care that was screened out was not sufficiently high-value to impact health outcomes.

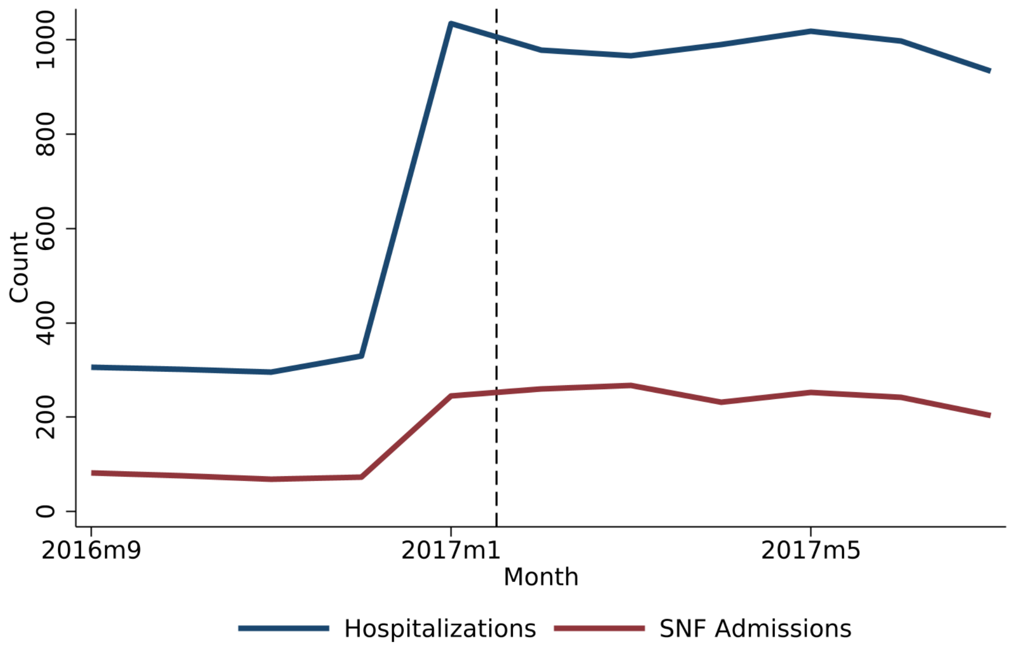
5 Additional Exhibits for Appendix C

Figure C1. Effect of NaviHealth Partnerships on Skilled Nursing Facility Length of Stay



Notes: Each panel shows ordinary least squares event-study estimates of the effect of three partnerships with NaviHealth. The pooled event-study (Panel A) is estimated using a stacked difference-in-differences approach. Panels B-D include only the given insurer and their geographically defined comparison group. Difference-in-differences estimates with covariates are shown above each figure. The baseline mean is the pre-partnership mean for patients covered by the treated insurer. Only patients admitted to a skilled nursing facility were included. Standard errors were clustered at insurer-county level.
Abbreviations: BCBS MI, Blue Cross Blue Shield of Michigan.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure C2. Horizon Enrollment Over Time



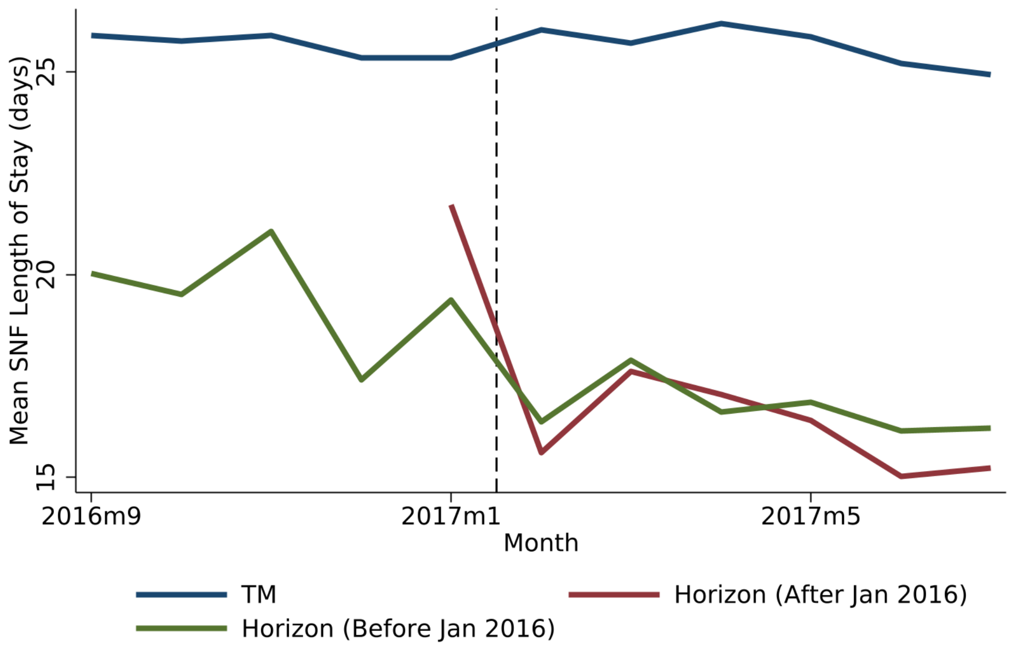
Notes: The vertical line shows the timing of the Horizon partnership with NaviHealth in February. Hospitalization and skilled nursing facility admission counts are based on the month of hospital discharge.
Abbreviations: SNF, skilled nursing facility.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Table C2. Effect of Horizon-NaviHealth Partnership on Skilled Nursing Facility Length of Stay

	Full Sample	**Excluding Recent Horizon Enrollees**
Estimate	-3.852***	-3.237***
	[-4.978, -2.727]	[-4.520, -1.953]
N	44,741	43,629

Notes: Entries are differences-in-differences estimates and 95% confidence intervals. In the third column, I exclude Horizon patients who were not enrolled in a Horizon plan in January 2016. Estimates were obtained using ordinary least squares adjusting for covariates. Standard errors were clustered at insurer-county level.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Figure C3. Length of Stay in New Jersey by Insurer and Timing of Enrollment



Notes: Values are mean skilled nursing facility length of stays by month. Horizon enrollees are split into those who were enrolled in Horizon in January and those that were not. Values for patients who enrolled in a Horizon plan after January 2016 but had a hospitalization ending before January 2017 are not shown.
Abbreviations: TM, traditional Medicare. Jan, January.
Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

Table C3. Effect of NaviHealth Partnerships on Health Care Use and Outcomes

Outcome	Pooled		MVP		Horizon	
	Effect	95% CI	Effect	95% CI	Effect	95% CI
Skilled Nursing Facility	-0.241	[-0.825, 0.343]	-0.868	[-2.015, 0.280]	1.547	[-0.201, 3.295]
Home without Post-Acute Care	0.369	[-0.558, 1.296]	1.627	[-0.541, 3.795]	1.014	[-1.268, 3.297]
Home Health	-0.262	[-1.007, 0.484]	-0.896	[-2.885, 1.093]	-2.271*	[-4.233, -0.309]
Inpatient Rehabilitation Facility	0.170	[-0.086, 0.426]	0.368	[-0.207, 0.942]	-0.676	[-1.775, 0.424]
Died or Hospice	0.107	[-0.151, 0.364]	0.554	[-0.342, 1.450]	0.182	[-0.710, 1.074]
Other	-0.143	[-0.391, 0.105]	-0.785*	[-1.417, -0.152]	0.204	[-0.853, 1.261]
SNF Length of Stay	-2.804***	[-3.381, -2.227]	-4.644***	[-6.318, -2.970]	-3.852***	[-4.978, -2.727]
SNF Stay >30 Days	-8.191***	[-9.531, -6.851]	-11.543***	[-15.985, -7.101]	-11.107***	[-14.554, -7.659]
90-Day Readmissions (full sample)	-0.134	[-0.665, 0.398]	-0.394	[-1.548, 0.759]	1.286	[-0.534, 3.105]
90-Day Mortality (full sample)	0.271	[-0.088, 0.631]	0.241	[-0.706, 1.188]	0.805	[-0.373, 1.982]
90-Day Readmissions (SNF users)	-0.359	[-1.808, 1.090]	-2.278	[-5.032, 0.475]	-0.927	[-4.312, 2.458]
90-Day Mortality (SNF users)	0.045	[-1.090, 1.179]	-2.203	[-4.732, 0.326]	-0.140	[-2.912, 2.631]

Notes: Entries are difference-in-differences estimates. The pooled estimates use a stacked difference-in-differences approach to aggregate the effects from the MVP, Horizon, and Blue Cross Blue Shield of Michigan NaviHealth partnerships. Estimates were obtained using ordinary least squares adjusting for covariates. Standard errors were clustered at insurer-county level.

Abbreviations: BCBS MI, Blue Cross Blue Shield of Michigan; SNF, skilled nursing facility.

Source: Author's analysis of the Medicare Provider Analysis Review file (MedPAR), Medicare Advantage encounter data, and the Medicare Master Beneficiary Summary File (MBSF).

$^{*}p < 0.05,^{**} p < 0.01,^{***} p < 0.001$