Producing Health: Measuring Value Added of Nursing Homes∗

Liran Einav† Amy Finkelstein‡ Neale Mahoney§

February 2, 2022

Abstract

We develop a stylized model that allows us to estimate a value-added measure for nursing homes (“SNFs”) which accounts for patient selection both into and out of a SNF. We use the model, together with detailed data on the physical and mental health of almost 6 million Medicare SNF patients between 2011 and 2016, to estimate the value added for about 14,000 distinct SNFs. We document substantial heterogeneity in value added across SNFs. Nationwide, compared to a 10th percentile SNF, a 90th percentile SNF is able to discharge a patient at the same health level 7 days sooner; this is about one third of the median length of stay in our sample. Heterogeneity in value added within a market is almost as large as it is nationwide. Our results point to the potential for substantial gains through policies that encourage reallocation of patients to higher-quality SNFs within their market.

∗We thank seminar participants at the University of Pennsylvania, Rice University, and Washington University for comments and suggestions. We are grateful to Sam Wang, Alicia Weng, Connie Xu, and especially James Okun for superb research assistance, and to the Laura and John Arnold Foundation, and the National Institute on Aging (Finkelstein R01-AG032449) for financial support.

†Stanford University and NBER, leinav@stanford.edu
‡MIT and NBER, afink@mit.edu
§Stanford University and NBER, nmahoney@stanford.edu
1 Introduction

The production of health lies at the core of health economics (e.g., Grossman 1972), yet with the exception of mortality, the study of health has been hampered by a lack of consistently and comprehensively measured data on health outcomes. As a result, instead of focusing on health outputs, most research in health economics has focused on inputs into the production of health – such as health care utilization and health behavior – where extensive data is available.

Nursing homes represent an important exception to the general paucity of data on health. For decades, virtually all nursing homes in the United States have been required to administer detailed health assessments to all of their patients at the time of admission, and at regular intervals during their stay. These assessments capture multiple different measures of each patient’s physical health, mental health, daily functioning, and cognitive capacity. They present a rare opportunity to study the health production process across healthcare providers, and to estimate potential gains from reallocation of patients.

In this paper, we use this rich data on health outcomes at nursing homes to construct a holistic measure of patient health and to estimate nursing home value added in improving this health measure for about 14,000 nursing homes. Our estimation accounts for potential bias from endogenous selection of patients into nursing homes and endogenous timing of when patients are discharged. We use the estimates to examine heterogeneity in nursing homes quality across and within markets, and study its association with patient and nursing home characteristics and with alternative measures of nursing home quality.

The nursing home sector is an important setting for assessing the production of health, both because of the size of the sector and because of widespread concerns about the quality of nursing home care. In 2016, there were over 15,000 certified nursing homes in the US, providing care to over 1.3 million patients at a cost of $160 billion, or roughly 5% of national health expenditures (Harrington et al. 2018, CMS 2019b). For decades, policymakers and the public have expressed concerns about the quality of nursing home care (e.g., Rau 2018, Rau and Lucas 2018, Goldstein and Gebeloff 2019, Jacobs and Richtel 2019) and about the ability of patients (or their caregivers) to assess this quality (Wunderlich et al. 1996, GAO 1996, …).
To help patients make choices, the Centers for Medicare and Medicaid Services (CMS) construct “star ratings” of each nursing home that are widely used (CMS 2008a, GAO 2016) but also widely criticized (e.g., Thomas 2014; Rau 2018; Ryskina et al. 2018; Silver-Greenberg and Gebeloff 2021).

We focus on Medicare patients, for whom nursing homes are intended to provide short-term care designed to aid in the recovery from a hospitalization or medical condition. In 2016, about one fifth of hospitalized Medicare patients were discharged to nursing homes, at a cost to Medicare of nearly $30 billion, or about 8% of overall Medicare spending (MedPAC 2018). We combine the consistently and comprehensively measured health assessment data into a one-dimensional “health index” that measures how fit the patient is to return to the community; it is straightforward to use our approach to estimate value added for other health measures, as we briefly illustrate below. We define a nursing home’s value added as the average change in a patient’s health index between the assessment at admission and the one taken 30 days later.

Estimating a nursing home’s value added requires overcoming two econometric challenges. First, there is potentially non-random sorting of patients to nursing homes, even conditional on observable patient characteristics. This problem of “selection in” is a standard one in the education literature on value added, and we follow the standard approach of constructing a control function to address it (Dubin and McFadden 1984; Abdulkadiroğlu et al. 2020). We use distance between the patient’s residence and the nursing homes in her market as the excluded instrument in her nursing home demand model.

A second challenge in our setting is that patients may leave the nursing home before their 30-day assessment, either because they recover and return to the community (which we refer to as “discharged downstream”) or because their health worsens and they are sent back to the hospital, to a hospice, or die (“discharged upstream”). We model this “selection out” with an explicit model of the nursing home’s discharge decision that is similar in spirit to Heckman (1979).

We estimate the model by maximum likelihood using data from 2011 to 2016 on health
assessments for almost 6 million Medicare patients at about 14,000 different nursing homes. Our results indicate substantial heterogeneity in value added. The average (weighted by patient stays) nursing home increases the weekly probability of community discharge (health index) by 4.5 percentage points between admission and 30 days, with a 7.0 percentage point difference in value added between the 90th and 10th percentile nursing home. To put that in perspective, consider that in the cross section of admitted patients, a 1 percentage point higher health index is associated with approximately 1 day shorter length of stay in the nursing home. The 7.0 percentage point difference in value added between the 90th and 10th percentile nursing home is thus roughly equivalent to a nursing home discharging a patient at a given health level a week sooner, which represents about a one-third reduction compared to the median length of stay in our sample. Given that a nursing home day costs Medicare about $470 (MedPAC 2018), this means that the 90th percentile value-added nursing home saves Medicare about $3,300 per patient relative to the 10th percentile nursing home. Interestingly, our estimates of nursing homes’ value added are only weakly positively correlated with their CMS star rating (correlation of about 0.25).

We find considerable variation in nursing home value added both across geographic markets and within geographic markets. Value added is on average 3 percentage points higher in the 90th percentile market than in the 10th percentile market; markets in the South tend to have substantially lower value added. Within markets, the average difference between the 90th and 10th percentile nursing home is 6.6 percentage points, almost as large as the corresponding nationwide statistic of 7.0 percentage points. Even in markets in the top decile of average value-added, more than 25% of patients are in nursing homes that perform below the national median. This points to the potential for substantial gains from within-market reallocation of patients to higher value-added nursing homes. Our focus on geographic variation in health production complements an influential body of research associated with the Dartmouth Atlas (e.g., Congressional Budget Office 2008, Gawande 2009, Skinner 2011, Institute of Medicine 2013) which has documented substantial regional variation in Medicare spending without corresponding variation in proxies for health care quality. Variation in spending for post-acute care (mainly nursing homes) explains a substantial share of the
spending variation. Relative to this existing literature’s focus on inputs (i.e. spending and utilization), we use rich, consistently measured data on health outcomes to estimate and examine heterogeneity in health care quality. Our finding of substantial heterogeneity in quality across market suggests that there may be scope to improve the quality of care in low-value-added regions. Moreover, our findings of substantial variation in value added within markets is encouraging, since there are likely to be more policy levers for encouraging reallocating patients within markets than across markets.

In addition to the Dartmouth Atlas literature, our paper relates to several other distinct literatures. It complements the literature on value added of hospitals (Chandra et al. 2016; Doyle et al. 2019; Hull 2020) that, like most of the prior health literature, has focused on mortality as the singular outcome. Our paper also contributes to a growing literature in health economics on nursing home behavior (Grabowski et al. 2008; Hackmann 2019; Gandhi 2020; Gandhi et al. 2021; Gupta et al. 2021; Hackmann et al. 2021) and a literature in health services research that examines cross-sectional variation in nursing home process measures and outcomes (e.g., Castle and Ferguson 2010). Finally, our methodological approach connects to a well-developed literature on teacher and school value added (e.g., Todd and Wolpin 2003; Koedel et al. 2015; Abdulkadiroğlu et al. 2020). This education literature has mainly focused on bias from “selection in,” while in our setting selection both “in” and “out” are prominent concerns.

The rest of the paper proceeds as follows. Section 2 provides background on our setting and our data. In Section 3 we present our econometric model of value added, and Section 4 discusses implementation and estimation. Section 5 presents the results, and Section 6 briefly compares our estimates to other measures of nursing home quality and explores other health measures and specifications. Section 7 concludes.

---

2 Specifically, despite accounting for only 16% of Medicare spending, post-acute care accounts for 73% of the geographic variation in spending (Institute of Medicine 2013).
2 Setting and Data

2.1 Institutional Setting

Nursing homes provide both short-term care to patients recovering from a hospitalization or illness, and long-term care to patients in need of ongoing assistance with their daily living. Most short-term patients are covered by Medicare, which pays for short-term nursing and rehabilitation services for Medicare patients recovering from a surgical procedures (e.g., hip replacement) or a health event (e.g., stroke). Crucially, Medicare coverage is predicated on the expectation that the patient is on a path to recovery and return to the community (CMS 2019a), which is why we will focus on discharge to the community as a key marker of “success.”

By contrast, patients who need ongoing assistance typically receive nursing home coverage from Medicaid. Some patients are initially admitted with Medicaid coverage, but many others transition to Medicaid during their stay. Therefore, while over 70% of nursing home patients are covered by Medicare at the time of admission, only 14% of nursing home patients have Medicare as their primary payer at a given point in time. Because Medicaid patients have much longer stays than Medicare patients, at a given point in time 60% of nursing home patients have Medicaid as their primary payer. The remainder are covered by private insurance (including Medicare Advantage) or pay out of pocket (Harrington et al. 2018).

Almost all nursing homes (96% of beds) are certified to care for both Medicare and Medicaid patients (Harrington et al. 2018). Nursing homes that are certified to treat Medicare patients are referred to by Medicare as Skilled Nursing Facilities (SNFs, pronounced “sniffs”), a terminology we will adopt for the remainder of the paper.

During our 2011-2016 study period, Medicare reimbursed SNFs at a prospective daily rate, with the rate depending on both the SNF’s geographic location and a measure of the patient’s health on admission derived from health assessments (MedPAC 2021). Starting at the 21st day in the SNF, patients must pay daily co-pays (either directly or via supplemental coverage), and after 100 days Medicare coverage ends.

\[^3\]Authors’ calculation using data described in Section 2.2
To be eligible for Medicare or Medicaid reimbursement, a SNF must comply with federal requirements regarding residents’ civil rights and ethical treatment of patients, as well as the health care services they provide (42 U.S.C. §483 2016). States conduct an initial certification of a SNF, after which the SNF must be re-certified annually. During (re-)certification, a state inspector collects (facility-reported) data on staffing levels and nursing home characteristics (such as number of beds, occupancy rates, and ownership type). The inspector also records any quality deficiencies that are identified during inspection (CMS 2018; Harrington et al. 2018; LTCFocus, no date). As part of the re-certification process, certified SNFs are also required to assess their Medicare patients’ health at regularly defined intervals during their stay, producing the health assessment data we use below (CMS 2017).

The information collected during re-certification is aggregated by CMS into a “five-star” rating system. These ratings, which have been constructed since 2008, are then publicly posted on the Nursing Home Compare website. Patients and their families are encouraged to use the five-star ratings in choosing a facility, and many do (CMS 2008a; GAO 2016); the ratings are also used by insurers for determining their provider networks, and by federal regulators deciding on loans to facilities (12 U.S.C. §1715w 2014 CMS, no date). Extremely poorly-performing SNFs – less than 0.1% of facilities – are subject to twice as many inspections, as well as financial penalties and potential termination (CMS 2022).

2.2 Data Sources

Our primary data source is the Long-Term Care “Minimum Data Set” for Resident Assessment and Care Screening (hereafter, MDS). These data contain a series of federally mandated, standardized patient assessments that track the health status of all patients in Medicare and Medicaid certified nursing homes. The assessments were developed by a consortium of professionals in the late 1980s in response to a congressional mandate to create national resident assessment system for nursing homes (Morris et al. 1990). These data have since been widely used by researchers studying the economics of nursing homes (e.g., Grabowski et al. 2008; Cornell et al. 2019; Hackmann 2019; Gandhi 2020; Gupta et al. 2021; Hackmann et al. 2021). The assessment form has been revised several times since it was first created. We use version 3.0, which was in effect during the 2011-2016 time period of our study.
The MDS covers all nursing home patients, not just the Medicare patients who are our focus. Assessments are required for all patients at admission and at discharge. In addition, during our study period, assessments were also required for Medicare-covered patients at days 14, 30, 60, and 90. The rich, longitudinal information that the assessments provide on patient physical health, mental health, and cognition will be the backbone of our measurement of value added. The MDS also provides basic demographics (age, race, gender, and marital status), length of stay, and discharge destination.

We supplement the MDS with additional data on Medicare patients and on nursing homes. We merge the MDS with Medicare data to identify the patient’s 5-digit zip code of residence, whether the patient is dually eligible for Medicaid (a proxy for low income), whether the patient has each of 27 different chronic conditions at the start of the year of their SNF admission, whether the patient was discharged to the SNF from an acute care hospital, and the patient’s diagnoses for the prior hospital admission if she was discharged directly from a hospital. We use facility-level data from the OSCAR/CARPS system, created during the annual re-certification process, to measure the number of beds in the SNF, its occupancy rate, for-profit status, and whether the SNF is based within a hospital.

Finally, we use the data to define the market or “choice set” of SNFs for each patient. Standard market definitions for hospital choice are either too large (the approximately 300 Health Referral Regions, or HRRs) or too small (the approximately 3,000 Health Services Areas, or HSAs, which are subsets of HRRs) for our purposes. Our “Goldilocks” solution, therefore, is to create our own SNF markets, which are either HRRs (when they are not too large) or groups of HSAs within HRRs. The end result is 655 markets; the median market has 20 SNFs, the 10th and 90th percentile markets have 6 and 37 SNFs, respectively. More details on our market definition and the resultant markets are provided in Appendix A.

---

4Specifically, initial assessments must be conducted by day 8, day-14 assessments must be conducted between days 13 and 18, day-30 assessments during days 27-33, day-60 assessments during days 57-63, and day-90 assessments during days 87-93.

5The MDS does not include information on whether the patient is covered by Medicare. We use a crosswalk between the internal resident ID in the MDS and the Medicare beneficiary ID to identify patients covered by Medicare at admission to the SNF. We use the Master Beneficiary Summary File to provide additional demographics, the Chronic Conditions file for chronic conditions diagnosed in the year of SNF admission, and the MedPAR file to identify whether patient had a prior hospital admission and which of the 241 Clinical Classifications Software (CCS) diagnoses applied for that admission.
2.3 Sample Construction and Summary Statistics

We start with the full sample of all patients stay at SNFs during the five fiscal years that begin on October 1, 2011 and end on September 30, 2016. The start of the period corresponds to the start of a new schedule for required assessments. The data contain 24.1 million patient-stays, admitted to 16,355 distinct SNFs. We restrict attention to patients who, at the time of their SNF admission, are at least 65 years old and are covered by Medicare. This leaves a sample that covers 10.7 million patient-stays at 16,307 distinct SNFs.

We then make several additional sample restrictions, which bring our final analysis sample to 5.9 million patient-stays at 13,867 distinct SNFs. First, we restrict attention to a patient’s first stay within an episode of nursing home care; this accounts for the vast majority of the sample decline. Second, we require that the patient is not missing any of the demographic characteristics or measures from the initial health assessment that are used to construct the baseline values for our health index (discussed below). Third, we require that the patient enters the SNF from an acute hospital discharge with a diagnosis code from that hospital stay (which about 92% of the remaining sample does). Fourth, we drop a small number of patients whose health assessment on admission occurs after day 8. Finally, to ensure an adequate sample size for estimating value added, we require that the SNF have at least 50 stays that meet our sample requirements over the five-year period. Appendix Table A2 provides more detail about the number of observations that are dropped due to each of these restrictions.

Table 1 presents summary statistics on the patients in the sample, as well as the standard deviation of the SNF-level average across SNFs, with each SNF weighted by the number of patient-stays. As shown in Panel A, the average patient is 81 year ago, almost two-thirds are female, one third are married at admission, and 86% are white. Almost one fifth of patients

---

6 A stay is defined as a set of contiguous days within a facility between admission and discharge or death. To account for patients with brief interruptions in their nursing home stay, CMS defines an episode of nursing home care as a period of time that spans one or more stays and ends either when the patient (i) is discharged without a return anticipated, or (ii) is discharged and does not return within 30 days, or (iii) dies. About 85% of Medicare patients admitted to a SNF have an episode consisting of a single stay. The stay is the more natural unit of analysis when measuring health changes, since time between stays can be of varying length and Medicare’s requirements for patient assessment are based on the days within a stay (although Medicare cost-sharing rules are based on the cumulative number of days to date across stays within an episode) (CMS 2016).

---

8
are dually eligible for Medicaid (a marker of low income), and the average hospital stay prior to admission lasted 7.5 days. The variation across patients in race and Medicaid status has a considerable SNF-specific component (see column (2)).

Panel B presents statistics on length of stay. Median length of stay is 22 days (not shown), but some patients stay for much longer, so the average length of stay is 45 days. Almost 60% of the patients in our sample are discharged from the SNF prior to their 30-day assessment, while a non-trivial share (15%) remain in the SNF past for more than three months.

Panel C shows discharge destinations for the approximately three fifths of our sample who are discharged prior to their 30-day assessment. Nearly two thirds are discharged “downstream” (i.e. to the community), which is the good outcome, and about one third are discharged “upstream” (most of whom are sent back to the hospital).

Figure 1 shows the movement of patients through our environment. Every patient in our sample receives an initial health assessment. Our value added measure, described below, is based on the change in outcomes between the initial assessment and the assessment conducted at 30 days; specifically, it measures the change in the probability that a patient is discharged to the community within a week of their assessment. We chose the 30-day assessment because, relative to the 14-, 60- and 90-day assessments, it seemed to strike a balance between being long enough after admission for SNFs to have had the opportunity to have an effect and short enough so that there is still a significant share of patients at the SNF. Before the 30-day assessment, 37.3% of patients are discharged downstream, 18.4% are discharged upstream, and 2.9% are discharged elsewhere, leaving 41.4% of patients in the SNF for the 30-day assessment. Within a week of the 30-day assessment, 7.6% of patients are discharged downstream, 2.1% are discharged upstream, and 0.6% are discharged elsewhere, leaving 31.1% of all patients in the SNF one week after the 30-day assessment.

Table 2 reports summary statistics on our baseline sample of SNFs. The average SNF has just over 100 beds; the 5th percentile SNF has 40 beds and the 95th percentile one has 210 beds. For most SNFs, occupancy rates hover around 80-90%, although there is a small fraction with occupancy rates lower than 60%. More than two thirds of SNFs are for-profit, which is similar to the rate elsewhere in the post-acute care sector (for example, almost three
quarters of long-term care hospitals are for-profit) but much higher than the rate for acute
care hospitals (where less than one fifth are for-profit) (MedPAC 2020). About 4% of SNFs
are located on the premises of a hospital. On average, our baseline sample includes 421
patient admissions per SNF; the 5th percentile SNF has 62 patients (recall that we required
a minimum of 50 patients as part of our sample restrictions) and the 95th percentile has
more than a thousand.

2.4 Health Measures

The rich data from patient health assessments are a key advantage of our setting. We
draw on 35 health measures that are consistently recorded over our study period and have
minimal missing data at the 30-day health assessment. Appendix B explains our selection
of measures in more detail.

The 35 measures include a wide range of health outcomes. Many are physical health
measures – such as urinary incontinence, shortness of breath, or dehydration – or physical
limitations to activities of daily living, such as walking, dressing, or toileting. These limitations
are measured on a 5-point scale ranging from fully independent (requiring no help or staff
oversight at any time) to total dependence (requiring full staff help every time during a
7-day period). There are also a number of mental health measures – including a depression
screen, the use of anti-psychotics, and the use of anti-anxiety medication – and a measure
of cognition (delirium).

Table 3 lists all 35 measures and provides summary statistics for them at admission. It
shows the mean and standard deviation across patients, and the standard deviation across
SNFs (weighted by patients).

There is a great deal of variation across measures in the frequency of health problems.
For example, about two fifths of patients are taking anti-coagulants, about one fifth are
depressed, 9% are on anti-psychotics, and 2% are experiencing vomiting. The median patient
scores a 3 out of 4 in terms of activities of daily living for walking, dressing, and toileting.

\footnote{In Section 4 below, we explain how we handle missing outcomes.}
\footnote{The fourth column discusses how each measure is used in constructing our baseline summary health
index; we defer a discussion of this to Section 4.}
meaning that for each of these activities she requires extensive assistance but is not totally dependent on the staff, and 1 out of 4 on urinary incontinence, meaning that she experiences incontinence only occasionally. Naturally, there is considerable heterogeneity across patients in these measures. While, naturally, there is less heterogeneity in average patient health characteristics across SNFs (column (3)) than across patients (column (2)), column (3) shows that there is a non-trivial heterogeneity across SNFs. This heterogeneity underscores the importance of grappling with potential unobservable selection on initial health.

3 An Econometric Model of Nursing Home Value Added

3.1 Defining Value Added

To fix ideas, consider a population of patients, each denoted by $i$, who are randomly assigned to a set of SNFs, each denoted by $j$. Patients arrive at SNFs with baseline health described by the index $h_{i1}$ at the start of period 1 (initial assessment). We then observe health $h_{i2}$ at period 2 (30-day assessment) for all patients.

Given this setup, we define the value added of SNF $j$ as $\alpha_j$ within the equation

$$h_{i2} = \alpha_j + \theta h_{i1} + \epsilon_i,$$

where $\epsilon_i$ is a mean-zero iid error term and $\theta$ controls for the effect of baseline health. Thus, SNF value added, $\alpha_j$, can be interpreted as the average improvement in health at SNF $j$ between period 1 and period 2, conditional on baseline health. This conceptualization of value added is very similar to that used in the education literature. For instance, teacher value added in test scores is often based on a regression of test scores at the end of a year on teacher fixed effects and lagged test scores.

While the specification may be familiar, it imposes several restrictions that are worthy of discussion. First, it assumes that $\alpha_j$ is homogeneous across patients. This allows us to estimate a single average value added for each SNF; it would be straightforward to estimate

\[\text{footnote}{\text{In practice, we will estimate } \theta \text{ to be very close to 1 (0.98). As a result, our measure of SNF-value added effectively measures the average improvement in patient health between period 1 and period 2, and we will use the terms “health improvement” and “value added” interchangeably in what follows.}}\]
the model separately for different groups of patients and thus allow for heterogeneity in SNF value-added based on observables. Second, it assumes that the serial correlation in health (captured by $\theta$) does not vary across SNFs. This assumption is needed for the cross-SNF differences in $\alpha_j$ to be interpretable; if we allowed $\theta$ to vary by SNF, we would not be able to characterize the health improvement across SNFs with $\alpha_j$ alone.

A natural concern is that SNFs may strategically misreport the health measures that feed into our health outcomes. In the education setting where test scores are the outcome, teachers or principals may have incentives to “teach to the test” (Hoffman et al. 2001; Lazear 2006) and have been found to fraudulently change students’ answers (Jacob and Levitt 2003). In our setting, the fact that health measures are reported by nursing home staff raises similar concerns. However, incentives to shade assessments operate in offsetting directions. On the one hand, better patient health will improve a facility’s quality ratings. On the other hand, worse patient health will increase the daily reimbursement rate received for the patient, which has created concerns about SNF “upcoding” (Bowblis and Brunt 2014; Levinson 2015). Importantly for our purposes, any net incentives to up- or down-code patients operate in the cross section. Since our value added estimates are based on within-patient changes in health assessment, they should not be affected by time invariant differences in coding intensity across SNFs.

3.2 Key Challenges to Estimation

Estimating equation (1) for a given health measure would be straightforward if patients were randomly assigned to SNFs and we observed health for all patients in periods 1 and 2. In practice, neither of these conditions holds, and we explicitly address this in our estimation.

First, patients are not randomly assigned to SNFs. If there is a correlation between patient health improvements and SNF quality, then estimates of $\alpha_j$ may be biased, although the direction of this bias is unclear. For example, if more savvy patients are more likely to choose higher quality SNFs and are more likely to improve, this would bias upwards the “naive” estimates of $\alpha_j$ for these high-quality SNFs, stretching the distribution of $\alpha_j$. Alternatively, if SNF quality is particularly important for those who would not improve otherwise, estimates of $\alpha_j$ for high-quality SNFs would be biased downwards, compressing
the distribution. We call this issue “selection in,” and address it by using the patient’s distance to different SNFs as an instrumental variables that shifts SNF choice.

Second, some patients leave the SNF before receiving a follow-up health assessment. Specifically, Figure 1 showed that three fifths of patients are discharged prior to their period 2 (day-30) health assessment, with about two thirds of them discharged “downstream” to the community, and one third discharged “upstream” (to a hospital or hospice, or die at the SNF). If there is a correlation between patient health improvements and SNF discharge propensities, estimates of $\alpha_j$ using only those patients who remain in the SNF until period 2 may be biased. This bias will most likely generate a compression of estimates around the mean. SNFs with higher value added will have a sicker pool of patients at day 30 than they would without discharge, since they are likely to discharge to community the patients that improve the most, understating the health improvements. SNFs with lower value added will also have a sicker pool of patients at day 30, but because the sickest are more likely to die or be transferred to a hospital before the 30-day assessment, the observed pool of patients may be healthier, overstating health improvements. We call this issue “selection out,” and address it with an explicit model of the SNF’s discharge decision.

Controlling for Selection Out of the Nursing Home

Selection of patients out of the SNF can occur both “downstream” when they are discharged to the community, presumably once their health is sufficiently good, and “upstream” when their health deteriorates and they have to be discharged to a hospital or (less frequently) die.

To account for this selection in an econometrically-tractable fashion, we model it as a two-step sequential process. First, SNFs actively make downstream discharge decisions, $d_i^D \in \{0, 1\}$, based on a SNF-specific health discharge threshold. Then, upstream discharge decisions are realized for the remaining patients, $d_i^U \in \{0, 1\}$, as a result of stochastic health deterioration. We choose to model the two discharge decisions sequentially and not simultaneously (for example via an ordered probit) in order to allow for different amounts of noise in each direction; this makes the model more flexible and allows us to fit the data significantly better.
We assume that downstream discharge is a result of a stylized discharge model, in which SNFs decide whether to discharge each patient to the community (that is, downstream) using the following rule:

\[ d^D_i = 1 \iff h_{i2} \geq \lambda_j + \nu_i, \quad (2) \]

where \( \lambda_j \) is a SNF-specific discharge threshold and \( \nu_i \) is an iid error term, drawn from \( N(0, \sigma_\nu) \). That is, SNFs are more likely to discharge healthier (higher \( h_{i2} \)) patients, with a discharge threshold that varies across SNFs and is affected by unobservables, such as the availability of family members to assist in the discharge and the opportunity cost of SNF beds.\(^{10}\)

We model discharge upstream in a more statistical way, assuming that upstream discharges are the results of a probit stochastic process, which is a (declining) function of \( h_{i2} \). That is, we assume

\[ P(d^U_i = 1|d^D_i = 0) = \Phi(\gamma_0 + \gamma_1 h_{i2}), \quad (3) \]

where \( \Phi(\cdot) \) is the standard normal cumulative density function. We expect \( \gamma_1 \) to be negative, so that patients with better health are less likely to be discharged upstream. We allow \( \gamma_0 \) and \( \gamma_1 \) to be market specific in order to account for any potential differences in hospital access across markets.

We should note that although the way we describe the downstream model (as an active decision) and the upstream model (as a more passive decision) is different (and arguably more realistic), the two situations are very similar from a statistical and estimation perspective, giving rise to a “probit-like” model downstream and a probit model upstream.

Controlling for Selection Into the Nursing Home

The possibility that patients do not sort into SNFs randomly, even conditional on observables, is a standard concern in the value-added literature. Following the literature (Dubin and McFadden 1984; Abdulkadiroğlu et al. 2020), we use an instrumental variable – the distance between the patient’s residence and her potential SNF choices – to construct a control

\(^{10}\)This model is obviously a simplification, abstracting for example from the possibility that the discharge decision also directly depends on the of length of time spent in the SNF or the improvement in the health index. In Appendix C we show that this parsimonious model performs well compared to richer specifications.
function that accounts for this selection. Specifically, we first specify a discrete choice model of SNF choice which depends on distance to the SNF. We then show how this choice model yields closed-form selection correction terms in the value-added equation. We provide evidence in support of the exclusion restriction in Section 4.2.

Let the utility of patient $i$ from SNF $j$ be

$$u_{ij} = \delta_j - \tau m_{ij} + \eta_{ij},$$

(4)

where $\delta_j$ is the average utility from SNF $j$, $m_{ij}$ is (the log of) the distance between patient $i$ and SNF $j$, and $\eta_{ij}$ is an error term drawn iid from a Type I Extreme Value (logit) distribution. Note that we impose no restriction on $\delta_j$; it may be correlated with SNF value-added $\alpha_j$ and with any other characteristics of the SNF that affect patient demand.

Following Dubin and McFadden [1984], we account for selection on unobservables with a linear control function in the (de-meaned) logit errors. Conditional on the choice of SNF $c_{ij}$, period 1 health $h_{i1}$, and the errors $\eta_{ij}$, expected period 2 health can be written as

$$E[h_{i2}|c_{ij}, h_{i1}, \eta_{i1}, ..., \eta_{iJ_i}] = \alpha_j + \theta h_{i1} + \sum_{k \in J_i} \phi_k (\eta_{ik} - \mu_\eta) + \varphi (\eta_{ij} - \mu_\eta),$$

(5)

where $J_i$ is the number of SNFs in patient $i$’s choice set, and $\mu_\eta$ is the mean of the logit errors (Euler’s constant). Integrating out the $\eta$’s yields the estimating equation

$$E[h_{i2}|c_{ij}, h_{i1}, m_{i1}, ..., m_{iJ_i}] = \alpha_j + \theta h_{i1} + \sum_{k \in J_i} \phi_k \beta_{ik} + \varphi \beta_{ij},$$

(6)

where the $\beta$’s are functions of the logit choice probabilities $\hat{p}_{ik}$:

$$\beta_{ik} (j) = \begin{cases} -\log \hat{p}_{ik} & k = j \\ \frac{\hat{p}_{ik}}{1-\hat{p}_{ik}} \log(\hat{p}_{ik}) & otherwise \end{cases},$$

(7)

and the logit choice probabilities are predicted values from the choice model described in equation (4):

$$\hat{p}_{ij} = \frac{\exp(\delta_j - \tau m_{ij})}{\sum_{k=1}^{J_i} \exp(\delta_k - \tau m_{ik})}.$$

(8)
Appendix D provides the full derivation.

The parameters in equation (6) have an economic interpretation. The $\phi_k \beta_{ik}$ terms control for any unobservable correlation between the demand shocks for SNF $k$ and the value added at SNF $k$. This term would correct, for example, for the bias that would result from a SNF that is known to generate larger improvements in health receiving patients who are unobservably less likely to improve. The $\varphi \beta_{ij}$ term soaks up any correlation between the patient’s demand shock at the chosen SNF $j$ and patient’s health improvement at the chosen SNF, beyond that captured by $\phi_k \beta_{ik}$’s. This term would capture any Roy-type selection in which a patient has an idiosyncratically higher preference for a SNF at which they are idiosyncratically more likely to improve.

3.3 Intuition for Identification

To gain some intuition for the identification properties of the model, we consider a simpler version of the model in which there is a health production function (equation (1)) and downstream discharge (equation (2)) only. The object of interest is the SNF value added parameters, the $\alpha_j$’s in equation (1)).

This version of the model closely resembles the standard Heckman (1979) selection model for estimating the impact of wages on hours worked. In the Heckman setup, the selection problem is that wages are not observed for individuals who work zero hours. In our case, the problem is that period 2 health is not observed for patients who are discharged before period 2. Our key identification challenge arises because SNFs may vary not only in their value added $\alpha_j$’s, but also in their discharge threshold $\lambda_j$’s. Thus, if we were to run a simple probit of discharge on initial health and SNF fixed effects, the SNF fixed effects would capture the difference $\alpha_j - \lambda_j$.

To address this, we take advantage of the (observed) heterogeneity in patient health at admission. The key to identification in our setting arises from the fact that we observe many patients with different initial health $h_{i1}$ within a SNF, and each SNF is characterized by a single discharge threshold $\lambda_j$, thus allowing us to identify $\alpha_j$ and $\lambda_j$ separately.\footnote{The fact that we observe discharge decisions for many patients within each SNF is also a key difference from the original Heckman (1979) setup, in which each worker makes only one labor force participation decision.}
To see this, note that for any two SNFs \( j \) and \( j' \), there exists a pair of initial health levels \( h_{i1} \) at SNF \( j \) and \( h'_{i1} \) at SNF \( j' \), such that (downstream) discharge probability is equalized. This also implies that for these two cases the selection correction term in the value-added equation is identical. With the selection correction terms equalized, the differential improvements for patients \( h_{i1} \) and \( h'_{i1} \) at their respective SNFs allow us to recover the SNFs’ differential value added. We can then apply this argument for each pair of SNFs in order to identify the full set of \( \alpha_j \)’s up to a level and scale normalization. The normalization is not needed if there exists a SNF \( k \) and a health level \( h \), such that discharge probability for that health level at that SNF is zero.

Our control function for selection into the SNF is more standard, using distance as an instrument to identify the selection correction terms. Intuitively, people who live farther away from a given SNF \( j \) will only choose it if it provides them with a large \( \eta_{ij} \) relative to people who live nearby. By comparing health outcomes for people who live further vs. closer to the SNF, we can identify any correlation between unobserved determinants of choice and improvements in health.

4 Implementation

We discuss the key implementation decisions we make in order to bring the model to the data. Specifically, we describe the construction of our baseline health measure, provide support for our choice of distance as an instrument for SNF demand, and discuss our specification and estimation in more detail.

4.1 Health Index

An attraction of our setting is the rich, consistent, and comprehensively measured data from patient health assessments – including measures of physical health such as shortness of breath, physical independence such as with dressing, mental health such as depression, and cognitive ability such as delirium. Naturally, these measures speak to different dimensions of decision and an instrument that shifts that decision without affecting wages is typically needed.
health, with differing (and unknown) utility weights. Likewise, SNF value-added may also differ across these dimensions.

Our econometric framework can be applied to any measure of health in our data. Moreover, unlike test scores in education, which (debatably) provide a low-dimensional summary of student achievement, there is no established measure of health that we can take “off the shelf” for our analysis. Therefore, for our baseline health measure, we follow the spirit of Morris et al. (2018) and construct a univariate health index that combines the health assessment data in a way that is guided by SNFs’ purpose: shepherding Medicare patients to the point when they can safely be discharged back to the community.

Specifically, we use the health assessment data to construct a predicted probability that a patient is discharged to the community within 7 days of their 30-day assessment. We use these predicted values as our health index. From this perspective, a higher quality SNF is one that can move a given patient more quickly to the level of health that is conducive to community discharge.

Operationally, we construct the health index with a regression tree that uses the 30-day health assessment data to predict whether the patient is discharged to the community within 7 days of their 30-day assessment. This allows us to capture potentially important interactions among the health measures. We estimate the tree using five-fold cross validation to tune the complexity parameter. The resulting tree has 288 terminal nodes and an Area Under the Curve (AUC) of 0.71.

Loosely speaking, our health index can be thought of as a weighted average of the underlying 35 health measures, with the weights reflecting the importance of these measures in increasing the likelihood the patient can be discharged back to the community. To give some sense of what drives the health index, the last column of Table 3 shows the “importance” of each individual measure, defined as the sum of the additional $R^2$ for that variable across

---

12The choice to use the 30-day assessment rather than the initial assessment is not essential, but is motivated by two factors. First, this makes the construction of the health index orthogonal to our primary analysis, which relies on what happens to patients between admission and the 30-day assessment, but not after. Second, over the first two weeks in the SNF there is a non-trivial set of patients who get transferred to other SNFs for, we suspect, non-health reasons. The choice of a 7-day outcome horizon is admittedly arbitrary, attempting to trade off greater noise that can be generated from additional health events over longer periods against lower prediction quality of “rare” outcomes, which would be the case if the horizon is too short.
all the leaves of the tree, scaled so that the “most important” variable has an importance value of 100. The most important variable for predicting discharge to the community is needing help with dressing; other important variable include urinary incontinence, needing help with toileting, and needing help with transfers (e.g., from a bed to a chair).

Table 4 summarizes the resulting health index and how it correlates with subsequent outcomes. Panel A shows the index for the patients with complete health data who are at the SNF for their 30-day assessment, which is the sample on which we trained the regression tree. While the average patient in this sample has a health index of 0.133 (that is, a 13.3% chance to be discharged to the community within 7 days), those who actually get discharged to the community within the next seven days have a higher health index (of 0.187 on average). Similarly, those who get discharged elsewhere (predominantly to the hospital) within 7 days are associated with a lower health index (0.113 on average) relative to those who remain in the SNF (0.125).

Panel B shows a similar qualitative pattern when we apply the health index construction to the initial health assessment for all patients, a different sample from the one used to construct the health index. Relative to the average health index at admission (0.131), patients who are discharged to the community before the 30-day assessment are healthier (health index of 0.177), while those who remain in the SNF or discharged elsewhere are sicker (0.106 and 0.101, respectively).

The full distribution of the health index at admission is shown in Figure 2, Panel A. The mean is 0.13 and the standard deviation is 0.08. Only 5% of patients have a probability of discharge at admission of 30% or more, which is consistent with them being admitted for care.

To provide an alternative way of interpreting the health index, Panel B shows a scatter plot of the average SNF length of stay against health index at admission for each of the 228 unique values of the health index. The relationship between health at admission and length of stay is monotone, with a slope slightly steeper than -1. Thus, one can alternatively think

---

13 The 1,983,205 patients we trained the regression tree on are less than the 2,429,200 patients who are still in the SNF at 30 days because about 20% of them are missing a 30-day assessment. In Appendix E we explain the hot-decking procedure we use to fill in the health index for these patients, and in Appendix Table A5 we show the fit for all 2,429,200 patients.
about a SNF with a 1 percentage point higher value added as one that is able to get a patient to the same level of health one day faster.

4.2 Distance as an Instrument for Nursing Home Choice

We use the patient’s distance to each nursing home as the excluded instrument in our control function for SNF choice. This is a common instrument for choice of healthcare provider, including choice of nursing home (Grabowski et al. 2013; Gupta et al. 2021), hospital (McClellan et al. 1994; Cutler 2007; Card et al. 2019), and other medical providers (Einav et al. 2016). Our identifying assumption is that distance affects health improvements only through the choice of SNF. To investigate the plausibility of this assumption, we examine the stability of the relationship between distance and SNF choice as we saturate the specification with proxies for patient health.

Panel A of Figure 3 plots the probability of going to a SNF as a non-parametric function of distance for the sample of patients and SNFs in greater Chicago. The figure shows that distance is strongly predictive of SNF choice, with the expected probability dropping by almost 50% as distance increases from 1 to 5 miles. It also shows that the relationship between distance and SNF choice is well approximated with a log relationship, which motivates our functional form choice in the demand model (equation (4)). Finally, it shows that this relationship is virtually identical across specifications with different sets of controls. This type of “coefficient stability” indicates that distance is uncorrelated with observable indicators of improvements in patient health, raising confidence in the plausibility of the exclusion restriction.

Panel B extends the exercise in Panel A to all of the markets in our data. We estimate the model of SNF choice as a function of log distance (see equation (4)) in each market both without and with controls, and plot the market-specific coefficients on log distance from the specification without controls (x-axis) against the specification with the full set of covariates (y-axis). The coefficients are almost identical across specifications, suggesting that the stability of the SNF demand-distance relationship we saw for the Chicago market in Panel A generally holds for all markets.
4.3 Specification and Estimation

Estimation proceeds in two steps. We first estimate the SNF demand model market-by-market to construct the choice probabilities \( \hat{p}_{ik} \) that are inputs into the control function terms \( \beta_{ik} \) in equation (6). For computational tractability, we estimate the demand model only controlling for the health index at admission.\(^{14}\)

With the control function estimates in hand, we then jointly estimate the remaining components of the model by maximum likelihood. Specifically, we jointly estimate the health process (equation (6)), the downstream discharge model (equation (2)), and the upstream discharge model (equation (3)).\(^{15}\) One numerical challenge is that with thousands of SNFs in the data, we have thousands of parameters (recall that \( \alpha_j, \lambda_j, \) and \( \phi_j \) are all SNF-specific). To overcome this challenge, we partition the model parameters to two groups: four “national” parameters \( (\theta, \varphi, \sigma_\epsilon, \sigma_\nu) \), and all other parameters, which are all market or SNF specific. This allows us to have a nested estimation procedure. Conditional on the “national” parameters, we estimate the model market by market, which is relatively standard and fast. We then search numerically over the four “national” parameters that maximize the likelihood, where each time we evaluate the likelihood we estimate the (conditional) model market by market. In Appendix \( F \) we describe the likelihood function and our estimation approach in more detail.

5 Results

5.1 Estimates of Nursing Home Value Added

Panel A of Figure 4 shows the distribution of value added for the SNFs in our baseline sample, along with summary statistics for this distribution. Panel B reports estimates of the “national” parameters that we restrict to be the same across SNFs in all markets, as well as summary statistics for the \( \gamma_0 \)’s and \( \gamma_1 \)’s, which are allowed to vary across markets. In this

\(^{14}\)In Appendix Figure A3 we show that the addition of further controls does not affect the estimated coefficient on log-distance.

\(^{15}\)As shown in Table 1 there is a small share (approximately 3%) of patients who are discharged to another SNF or to “other” locations prior to the 30-day assessment. Because it is unclear how to classify such discharges, we do not use these patients for estimation.
figure, and all subsequent results, each SNF is weighted by the number of patient-stays so that the statistics are representative of the nursing home sector.

The average (and median) SNF value added is 0.045. Combined with our estimate of almost exactly 1 for the coefficient on health at admission ($\theta = 0.98$), this implies that the average SNF increases the probability of discharge to the community by about 5 percentage points between the initial and 30-day health assessments. The fact that value added is on average positive (health improves on average while at a nursing home) is consistent with the role of nursing homes in the Medicare system to aid patient rehabilitation and recovery (rather than to provide long-term custodial care).

The figure also indicates substantial heterogeneity across SNFs in value added. The difference in value-added between a 25th and 75th percentile SNF is 0.036; the difference in value added between the 10th and 90th percentile SNF is 0.07. We can benchmark the variation in value added across SNFs by comparing it to the variation in health at admission across patients. The interquartile range of value added across SNFs of 0.036 (Figure 4) is slightly more than one third of the inter-quartile range of health in admission across patients of 0.09 (Figure 2). Alternatively, going from a SNF at the 25th percentile to the 75th percentile of value added would generate the same health improvement (0.036) as moving a patient from the 5th to the 25th percentile in health at admission.

The heterogeneity in value added across SNFs can also be interpreted in terms of variation in length of stay and the associated Medicare spending at the SNF. Recall from Figure 2 that a one percentage point increase in the health index at admission is associated with about one less day in the SNF. Given this relationship, moving from a 10th percentile value-added SNF to a 90th percentile one can also be thought of as getting a patient to the same level of health seven days faster (relative to a median length of stay of 22 days). With a median

\[ \text{It is natural to be concerned that the estimated variation in value added across SNFs is biased upward due to estimation error. In Appendix Figure A4, we plot a kernel density of the baseline value added and a second kernel density of estimates where we use an empirical Bayes methodology to shrink our estimates towards the mean to account for this noise. The adjustment has a negligible effect on the dispersion in value added. This is not particularly surprising since we restricted our sample to SNFs with at least 50 patient stays. Given the similarity of the estimates, we use the non-adjusted estimates for the subsequent analysis. Additionally, to test for variation in value added over time within SNFs, we separately estimate our model on the first and second half of our sample after restricting to 5.5 million patients at 10,828 SNFs with sufficient sample size in both periods. We find that the correlation between value added estimated on these two samples is 0.72, suggesting value added is fairly stable over time.} \]
Medicare spending per SNF day of about $470 (MedPAC 2018), this corresponds to about $3,300 in Medicare savings.

Panels C and D of Figure 4 show the relationship between the SNF value added and the two other SNF-specific parameters estimated in the model: the discharge threshold $\lambda_j$ and the selection coefficients $\phi_j$. Panel C shows that the discharge threshold is only weakly negatively correlated with the value added (correlation of -0.02). While the magnitude is small, the negative correlation is consistent with higher value added SNFs discharging patients more quickly to the community, conditional on patient health. This relationship implies that moving patients to higher value-added SNFs would not come at the cost of increasing their length of stay (or Medicare SNF spending). Indeed, since higher value-added SNFs are able to improve health more quickly, the fact that they have similar discharge thresholds to lower value-added SNFs implies that this patient reallocation would primarily manifest itself by reducing length of stay.

Panel D shows that the correlation between the value added and $\phi_j$ is negative and large (correlation of -0.67). This negative relationship is consistent with a standard adverse selection mechanism, in which the most challenging patients in terms of expected health improvement have greater preferences for the highest value-added SNFs. Since distance is our excluded instrument, the estimates more specifically imply that patients with the lowest expected health improvements are disproportionately likely to travel longer distances to high-value SNFs. While we estimate important SNF-level selection based on $\phi_k$, the “Roy selection” coefficient $\varphi$ is small and insignificant (panel B). This implies that there is little correlation between patients’ unobserved health improvements and patients’ idiosyncratic preferences for SNFs.

5.2 Sources of Heterogeneity in Nursing Home Value Added

Figure 5 plots average SNF value added by market (Panel A) and difference in value added between the 90th and 10th percentile SNFs within each market (Panel B), as well as associated summary statistics (Panel C). Average nursing home value added is lower across a band of markets that stretches from Texas northeast through the Deep South and up through Appalachia. It is highest in the New England and the Upper Midwest. The
difference in average SNF value added between the 90th percentile market and the 10th percentile market is about 2.9 percentage points, or about two-fifths of the 90th to 10th percentile difference across SNFs.

We estimate considerable variation in nursing home value added within markets. In the average market, the difference in value added between the 90th and 10th percentile nursing homes is 6.6 percentage points (bottom row of Panel C), which is almost as large as the 7 percentage point 90-10 difference nationwide (top row of Panel C).\footnote{As shown in Panel B, the extent of dispersion in SNF value added within a market does not show any consistent pattern across regions. For example, high-dispersion markets – which can be thought of as less efficient perhaps – are not concentrated in low value-added markets such as the south, and are scattered across the country. Appendix Figure A5 confirms this with a scatter plot of the within market differences against the average value added at the market level.}

These results suggest that the potential gains from within-market reallocation are large. For example, if patients at the 10th percentile SNF within a market could be moved to the 90th percentile SNF within the same market, the gains would be on average equivalent to getting a patient home at the same health level 6.6 days sooner (relative to a median length of stay of 22 days), saving Medicare approximately $3,100 per patient. While the existing literature on geographic dispersion in healthcare utilization has focused primarily on cross-market variation (Skinner 2011; Institute of Medicine 2013), heterogeneity within markets may be more relevant for thinking about potential gains from reallocating patients across SNFs. It seems more feasible to imagine policies – such as payment incentives or information to consumers – that could reallocate patients across SNFs within their geographic market, rather than across the United States.

To further gauge the magnitude of within-market dispersion in value added relative to the overall distribution, Figure 6 plots the average within-market distribution of value added for markets at different deciles in the national distribution of value added. It shows substantial overlap in the distribution of SNF quality across markets of different average quality. For example, even in markets in the top decile of SNF value-added, about 25% of patients are in SNFs that perform below the national median. Likewise, even in markets in the bottom decile of value added, a 75th percentile SNF still performs about as well as a 25th percentile SNF in the top decile of markets.

Given the substantial variation in value added both across and especially within markets,
we examine how SNF value added correlates with characteristics of the SNF and of its patients. Table 5 summarizes these correlations both unconditionally (column (1)), and after conditioning out market fixed-effects (column (2)). Consistent with the substantial within-market heterogeneity, the results are fairly similar, and we focus on the within-market results in column (2). The relationship between SNF value added and SNF-level characteristics is fairly weak (Panel A). SNF value added is weakly decreasing in size (as measured by number of beds) and virtually uncorrelated with the SNF occupancy rate. For-profit SNFs, which account for 71% of facilities, have only slightly lower value added. SNF value added is much more strongly correlated with SNF patient characteristics (Panel B). In particular, there is very strong sorting to SNFs based on baseline patient health, with healthier patients (higher health index at admission) much more likely to go to higher value-added SNFs. Higher value-added SNFs also receive a lower share of Black patients, and a lower share of patients who are on Medicaid at the time of admission.

6 Relationship to Alternative Nursing Home Ratings

We explore how our value added compares to alternative measures of SNF quality, including the government “five-star” ratings and alternative measures we can construct with our data. Table 6 summarizes our findings, reporting the unconditional correlation between our value added and each of the alternatives (column (1)) and the partial correlation conditional on market fixed effects (column (2)). Conditioning on market has little impact and we focus our discussion on the conditional results.

6.1 Comparison to Government “Five-Star” Rating

Panel A compares our estimates of SNF value added to their “five-star” rating produced by the government and regularly used by patients, providers, and insurers. The overall rating is constructed from three sub-components that are constructed from the data collected during re-certification: a quality index which relies on the same health-assessment data we use in

---

18 Appendix Figure A6 shows scatter plots that correspond to each row of Table 5.
19 Value-added is on average 0.043 for for-profit SNFs, compared to 0.053 for non-profits.
20 Appendix Figure A7 produces scatter plots that correspond to each row of Table 6.
this paper, an inspection score based on deficiencies (e.g., abuse and neglect by staff and medication errors) identified during the three most recent annual state health inspections and the last three years of complaints, and a SNF-reported staffing measure (CMS 2008b).

The first row of Table 6 reports a fairly low positive correlation of 0.23 between our value added and the overall five-star ratings. The other rows in Panel A show the correlation between our value added and the star-rating’s sub-components.

The low correlation between our measure of value added and the government ratings is consistent with widespread concerns about the reliability of the government ratings, highlighted in a scathing New York Times critique (Thomas 2014). Among other issues, there is evidence that SNFs game the staffing measure by inflating self-reported estimates and up-staffing around the (known) inspection dates (Thomas 2014, Boccuti et al. 2015, Han et al. 2016, Rau 2018). The low correlation between our measure and the quality index component may reflect a number of underlying differences in the construction of these measures.

To get a sense of how different markets might fair under our value-added measure relative to the current star ratings, Figure 7 plots the average star ratings in each market against the average value-added estimate in each market. There is a lot of dispersion around the regression line, suggesting that a number of markets would benefit or suffer substantially under the alternative measures. For example, Manhattan (New York) and Chicago both score well under the star ratings (averaging 4.08 and 3.8 stars respectively, relative to the nationwide average of 3.5), but perform substantially below average by our value-added measure (with average scores of 0.006 and 0.018 respectively, relative to the average of 0.045). On the other hand, Renton, Washington (southeast of Seattle) averages a very low star rating of only 1.8, but an above-average value-added estimate of 0.060.

21 Unlike our measure of health improvement, the government’s quality rating is based on a cross-section of patient health. The government quality rating is also based only on a partially overlapping set of measures, and it includes all patients rather than just the short-stay Medicare patients that we focus on.
6.2 Comparison to 90-day Outcomes

We examine the correlation between our value-added measure and quality measures based on outcomes that we observe at 90 days after SNF admission, such as hospital readmission or death. The advantage of all these measures is that we can measure them for all patients, and therefore do not need to worry about “selection out (of the SNF)” concerns. One disadvantage, however, is that all these measures are relatively indirect measures of what the SNF is supposed to be doing – getting a patient to the point where they can safely be discharged to the community. Another limitation is that – because they are measured at 90 days – they are likely affected by events that happen to the patient outside of the SNF (recall that the median length of stay is 22 days), and therefore provide only a noisy measure of SNF quality. These caveats notwithstanding, their comparison to our value-added measure can still serve as an informative sniff-test (ahem) of our estimates, and also allow us to investigate potential concerns about multi-tasking.

To estimate SNF quality for these outcomes, we make two modifications to the baseline value-added specification described in Section 3. First, because these outcomes are observed both inside and outside of the SNF, we do not face the same issue of sample selection bias from patient discharge and therefore do not need to control for selection out. Second, unlike the health index that we observe at admission and at 30 days, we do not observe these outcomes at admission and therefore cannot estimate a value-added model that uses the value of the outcome at admission as a control. Instead, we control for the health index at admission as well as the control functions for SNF choice, as in Equation (6), which gives our estimates a slightly different interpretation.

Panel B of Table 6 shows these estimates for six different 90-day outcomes. Across outcomes, the correlation ranges from almost 0.5 to less than 0.2. Our measure of value added was designed to capture the extent to which SNFs are able to improve patient health in a way that allows them to be discharged to their homes. Consistent with this intent, our preferred measure has the largest correlation with the health measures that also reflect

---

Note that we “flip” the sign of “bad” outcomes, so that a positive correlation has the same qualitative interpretation across the six 90-day outcomes, regardless of whether they are “good” (e.g., home in 90 days) or “bad” (e.g., death or readmission).
whether the patient is able to transition to their homes (share of days in facilities, indicator for whether a patient is home at 90 days), and weaker for the other measures. The fact that SNFs that score better on our value-added measure also tend to score better on all of these other measures suggests that higher SNF value added does not come at the expense of addressing other unmeasured health issues that would show up in a higher death rate or higher readmission rate.

6.3 Comparison to Alternative Health Indices

As noted at the outset, our approach can be used to estimate value added for alternative health indices that can be constructed from the health assessment data. To illustrate this, we construct two standard health indices and compare the correlation between our preferred measure of value added with the measures of value added that use these alternative indices (Panel C of Table 6). The first alternative measure we consider, known in the clinical literature as the long-form ADL (Carpenter et al. 2006), is the sum of scores on the 7 ADL measures reported in Table 4. Our preferred value added measure has a fairly strong correlation with value added measured using ADLs (correlation of 0.79), which is consistent with the importance of ADLs in our health index (see Table 4). The second measure we consider is the quality measure proposed in Morris et al. (2018), which is constructed by normalizing ADL measures, diving them into three equal-sized groups, and summing them together. We find a moderately strong 0.41 correlation between our preferred value added measure and value added estimated using the Morris et al. (2018) quality indicator, which again reflects the importance of ADLs in the construction of our health index (see Table 4).

6.4 Comparison to More Basic Econometric Models

Finally, we considered the quantitative importance of different components of our econometric model for our value-added estimates. We start by considering a basic model in which value added is simply the average value of the 30-day health index at the SNF. We recover this measure, which we call *Average health*, from a regression of period 2 health on SNF

\[ \text{Average health} = \beta_0 + \beta_1 \text{SNF} + \epsilon \]

Since each measure takes value from 0-4, the aggregate measure ranges from 0 to 28, with higher values indicating that the patient needs more assistance.
fixed effects: \( h_{i2} = \alpha_j + \epsilon_{i2} \). A slightly more sophisticated measure, which we call \textit{Average health improvement}, is based on a regression of period 2 health on initial health and patient demographics: \( h_{i2} = \alpha_j + \theta h_{i1} + X' \alpha_x + \epsilon_{i2} \).

The first two rows in Panel D of Table 6 report the correlation between our preferred measure of value added and these two simpler measures. Our preferred value-added measure has 0.65 correlation with average health and 0.74 correlation with the health improvement. The correlations indicate that these more basic measures capture roughly two thirds of the variation reflected in our preferred measure.

The final two rows in Panel D of Table 6 explore the quantitative importance of the corrections we employ in our construction of SNF value added. In the third row, we report the relationship between our preferred value-added measure and a version of the average improvement model that accounts for the selection into the SNF (as we do in our baseline model), but does not account for the selection out (via the endogenous discharge). The fourth row reports on the reverse exercise, accounting for selection out but without correcting for selection into the SNF. Correcting for selection into the SNF actually reduces the correlation with our preferred measure to 0.56 (but controlling for market fixed effects brings this up to 0.84) relative to the model with no selection correction (correlation of 0.72), while correcting for selection out has virtually no impact (correlation of 0.71).

### 7 Conclusions

We developed a stylized model that allowed us to estimate and compare the health production process at almost 14,000 nursing homes. We find substantial heterogeneity in value added across nursing homes; moving from the 10th percentile of the nursing home value-added distribution to the 90th percentile is equivalent to being able to discharge a patient to the community almost 7 days faster (or almost one-third the median length of stay in our sample).

Strikingly, we find that the dispersion in nursing home value added \textit{within} markets is almost as large as the nationwide dispersion. In other words, there are at least some high-quality nursing homes in low-quality markets and at least some low-quality nursing homes
in high-quality markets. This points to the potential for substantial gains from policies that encourage reallocation of patients to higher quality nursing homes within their market.

Our findings suggest that an important area for further work is therefore to analyze the economic and policy forces that can achieve such market reallocation. Using our value-added measure as the dependent variable, it would be useful to study how it is affected by mergers or other changes in competition, or by the types of policies that are used to try to improve nursing home quality, such as changes in Medicare reimbursement, the imposition of minimum staffing ratios, or information campaigns (Wunderlich and Kohler 2001; Konetzka 2020).

There are also a number of natural ways to extend the analysis of nursing homes value added. Most obviously, while we focus on differences across nursing homes in improving health to the point that the patient can be safely discharged to the community, our approach can be applied to any of the many different underlying measures of physical and mental health. In addition, we have estimated a single value-added measure per nursing homes; future work could explore whether there is important heterogeneity in value added within a nursing home across different types of patients.

References


Long-Term Care: Facts on Care in the US. No date. “About Us”. [http://ltcfocus.org/about](http://ltcfocus.org/about)


Figure 1: Movement and Measurement of Patients

Notes: All statistics on discharges are calculated using all 5,871,691 patient-stays in our baseline sample. Statistics on discharge after the 30-day assessment are for discharges within 7 days of the 30-day assessment. Between the initial assessment and the 30-day assessment 2.9 percent of patients are discharged to a destination not accounted for by our upstream and downstream destinations, which includes discharges to other SNFs, MR/DD facilities, inpatient rehabilitation facilities, and other/unknown. Between the 30-day assessment and one week after the assessment, 0.61% of patients are discharged to a destination not accounted for by our upstream and downstream destinations.
Figure 2: Health Index

(A) Distribution of Health Index at Admission

(B) Health Index at Admission and Length of Stay

**Notes:** Panel (A) shows the distribution of our health index at admission ($h_1$). Panel (B) shows the relationship between length of stay and health index at admission ($h_1$) conditional on discharge to the community at any time after admission; it shows the relationship for each of the 228 unique values of $h_1$ that come from our regression tree. The line in Panel (B) is the linear fit between average length of stay and baseline health for patients with a baseline health measure below the 95th percentile (0.31) conditional on discharge to the community.
Figure 3: Distance Instrument for Selection In

(A) Chicago MSA

(B) All Markets

Notes: Panel (A) plots the non-parametric relationship between the probability of going to a SNF and distance to that SNF in the greater Chicago market; it shows this relationship with and without various controls (N = 212,937 patients at 294 SNFs). Panel (B) plots the market-by-market relationship between SNF choice and log distance from estimating the demand model in equation (4) on our baseline sample. We estimate this model without and with controls, and plot the coefficients from log distance on specification without controls against a specification with the full set of controls; this analysis is restricted to patients at SNFs in the 648 markets that have at least 2 SNFs (7 markets have a single SNF). When we control for demographics and health in Panel A, we control for the first ten principal components of demographics as well as the health index at admission. When we also control for chronic conditions in Panel A and B, we control for the first ten principal components of demographics, health, and indicators for the 20 most common Clinical Classification codes. The specific demographics are: above median age, female, married, Hispanic, White, dual eligible. The specific health measures are all taken from the 5-day assessment: walking, indwelling, falls, shortness of breath, depressed, delirium, vomiting, fever, dehydration, weight loss, and long-form ADL.
Figure 4: Value Added

(A) Distribution of Value Added

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ</td>
<td>0.98</td>
</tr>
<tr>
<td>ϕ</td>
<td>0.0</td>
</tr>
<tr>
<td>σ_ε</td>
<td>0.063</td>
</tr>
<tr>
<td>σ_ν</td>
<td>0.15</td>
</tr>
</tbody>
</table>

B. Upstream discharge parameters:
- Average γ_0: -0.35
- Std. Dev. of γ_0: 0.25
- Average γ_1: -1.6
- Std. Dev. of γ_1: 5.3

(C) Value Added vs. Discharge Threshold

Slope: -0.00
Correlation: -0.02
R-squared: 0.00
N = 13,888

(D) Value Added vs. Selection Coefficient

Slope: -2.69
Correlation: -0.67
R-squared: 0.45
N = 13,888

Notes: Figures show (A) the distribution of value added across SNF weighted by the number of stays in the SNF, (B) the national parameters from our model and summary statistics of the market level upstream discharge parameters (weighted by the number of patients in each market), (C) the relationship between value added (α_j) and the SNF specific downstream discharge thresholds (λ_j), and (D) the relationship between value added (α_j) and the SNF specific coefficients on the selection in control function (φ_j). These estimates are for 13,888 SNFs with value added estimate (i.e., the SNFs have some number of patients at 30 days). All figures are constructed weighting each SNF by its number of patients in our sample.
**Figure 5: Heterogeneity Across and Within Markets**

(A) Average Value Added by Market

(B) 90-10 Difference in Value Added By Market

(C) Distributions of Value-Added

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>10th</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>90th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Across SNFs</td>
<td>0.045</td>
<td>0.033</td>
<td>0.011</td>
<td>0.027</td>
<td>0.044</td>
<td>0.063</td>
<td>0.081</td>
</tr>
<tr>
<td>Across Markets</td>
<td>0.045</td>
<td>0.014</td>
<td>0.03</td>
<td>0.039</td>
<td>0.047</td>
<td>0.053</td>
<td>0.059</td>
</tr>
<tr>
<td>Within Markets</td>
<td>0.045</td>
<td>0.028</td>
<td>0.013</td>
<td>0.029</td>
<td>0.045</td>
<td>0.062</td>
<td>0.079</td>
</tr>
</tbody>
</table>

**Notes:** Panel (A) plots the average value added in each market. Panel (B) plots the difference between the 90th percentile and 10th percentile of the value added distribution within markets. Panel (C) shows summary statistics of value added across SNFs, across markets, and within markets. Both maps and summary statistics are constructed weighting each SNF by its patients. The color gradients follow quartiles of the corresponding market-level distribution. Data is shown for 655 markets covering 13,888 SNFs. Blank regions in the maps have no SNFs with patients in our sample.
Figure 6: Dispersion in Value Added Within Markets

Notes: Figure shows box-and-whiskers plots of SNF value added by deciles of the market-level distribution of value added. The lower (upper) “whisker” shows the 2.5th (97.5th) percentile. The bottom of each box is the 25th percentile, the middle is the median, and the top is the 75th percentile. The distribution depicted by each box-and-whisker plot is weighted by patients. The dashed horizontal line is the median across all SNFs in our sample, weighted by patients.
Figure 7: Correlating CMS Star Ratings and Value Added at the Market Level

Notes: Figure shows a scatter of the average CMS overall star rating against average value added at the market level. Market-level averages are weighted by the number of patients in each SNF. The points represent 655 markets covering the 13,550 SNFs (out of the 13,888 SNFs with estimated value added) for which we observe CMS star ratings. The black dot shows the patient-weighted average of the overall rating and value added. Certain outlier markets and other markets of interest are labeled. The line is a linear fit weighted by the number of patients in each market.
Table 1: Patient Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Mean (1)</th>
<th>Std. dev. across SNFs (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Patient characteristics at admission (N = 5,871,691)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>81.1</td>
<td>(8.32) 1.92</td>
</tr>
<tr>
<td>Female</td>
<td>0.64</td>
<td>0.06</td>
</tr>
<tr>
<td>Married</td>
<td>0.36</td>
<td>0.07</td>
</tr>
<tr>
<td>White</td>
<td>0.86</td>
<td>0.18</td>
</tr>
<tr>
<td>Black</td>
<td>0.08</td>
<td>0.14</td>
</tr>
<tr>
<td>Asian</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.03</td>
<td>0.09</td>
</tr>
<tr>
<td>Other/Unknown</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>Medicaid (aka &quot;dual&quot;)</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>Length of hospital stay (prior to SNF admission)</td>
<td>7.51</td>
<td>(7.55) 1.89</td>
</tr>
<tr>
<td>B. SNF length of stay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of stay</td>
<td>45.1</td>
<td>(101.49) 20.73</td>
</tr>
<tr>
<td>Share discharged before 30-day assessment</td>
<td>0.59</td>
<td>0.13</td>
</tr>
<tr>
<td>Share discharged before 60-day assessment</td>
<td>0.80</td>
<td>0.12</td>
</tr>
<tr>
<td>Share discharged before 90-day assessment</td>
<td>0.85</td>
<td>0.10</td>
</tr>
<tr>
<td>C. Discharge destination for patients discharged prior to day-30 assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community (downstream)</td>
<td>0.64</td>
<td>0.14</td>
</tr>
<tr>
<td>Acute care hospital (upstream)</td>
<td>0.27</td>
<td>0.11</td>
</tr>
<tr>
<td>Deceased or hospice (upstream)</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Another nursing home</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Other</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Notes: Table shows summary statistics of patient characteristics in our baseline sample (N = 5,871,691 patients at 13,941 SNFs). Column (1) reports the mean across patients, with patient-level standard deviations in parentheses. Column (2) reports the standard deviation of the SNF-level average across SNFs, with each SNF weighted by the number of patients.
### Table 2: SNF Characteristics

<table>
<thead>
<tr>
<th></th>
<th>No. of obs.</th>
<th>Mean (weighted mean)</th>
<th>Std. dev.</th>
<th>5th pct'ile</th>
<th>Median</th>
<th>95th pct'ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of beds</td>
<td>13,595</td>
<td>111.9 (128.3)</td>
<td>60.5</td>
<td>40</td>
<td>104</td>
<td>210</td>
</tr>
<tr>
<td>Occupancy %</td>
<td>13,595</td>
<td>82.2 (83.8)</td>
<td>12.6</td>
<td>56.9</td>
<td>85.9</td>
<td>96.0</td>
</tr>
<tr>
<td>For profit status</td>
<td>13,595</td>
<td>0.71 (0.70)</td>
<td>0.04</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital-based</td>
<td>13,595</td>
<td>0.04 (0.06)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients</td>
<td>13,941</td>
<td>421.2 (911.2)</td>
<td>454.3</td>
<td>62</td>
<td>272</td>
<td>1,273</td>
</tr>
</tbody>
</table>

**Notes:** Table shows summary statistics of SNFs characteristics in our baseline sample. All data except for the number of patients in the bottom row come from Online Survey Certification And Reporting (OSCAR), which are collected during annual inspections of SNFs and are missing for 346 SNFs. Means weighted by the number of patients in the SNF are shown in parentheses below the corresponding unweighted mean.
Table 3: Health Index Measures

<table>
<thead>
<tr>
<th>Health Measure</th>
<th>Mean (1)</th>
<th>Std. dev. (2)</th>
<th>Std. dev. across SNFs (3)</th>
<th>Importance (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has fallen (0-1)</td>
<td>0.07</td>
<td>0.25</td>
<td>0.03</td>
<td>50.0</td>
</tr>
<tr>
<td>Vomiting (0-1)</td>
<td>0.02</td>
<td>0.15</td>
<td>0.01</td>
<td>2.2</td>
</tr>
<tr>
<td>Fever (0-1)</td>
<td>0.02</td>
<td>0.15</td>
<td>0.02</td>
<td>0.8</td>
</tr>
<tr>
<td>Dehydration (0-1)</td>
<td>0.00</td>
<td>0.06</td>
<td>0.01</td>
<td>0.0</td>
</tr>
<tr>
<td>Life prognosis &lt; 6 months (0-1)</td>
<td>0.01</td>
<td>0.08</td>
<td>0.01</td>
<td>0.0</td>
</tr>
<tr>
<td>Anticoagulants (0-1)</td>
<td>0.37</td>
<td>0.48</td>
<td>0.11</td>
<td>33.2</td>
</tr>
<tr>
<td>Antibiotics (0-1)</td>
<td>0.38</td>
<td>0.48</td>
<td>0.07</td>
<td>19.0</td>
</tr>
<tr>
<td>Diuretics (0-1)</td>
<td>0.37</td>
<td>0.48</td>
<td>0.07</td>
<td>4.8</td>
</tr>
<tr>
<td>Indwelling catheter (0-1)</td>
<td>0.11</td>
<td>0.31</td>
<td>0.04</td>
<td>3.1</td>
</tr>
<tr>
<td>Weight loss (0-2)</td>
<td>0.10</td>
<td>0.43</td>
<td>0.10</td>
<td>16.1</td>
</tr>
<tr>
<td>Short-breathness (0-3)</td>
<td>0.32</td>
<td>0.73</td>
<td>0.21</td>
<td>32.0</td>
</tr>
<tr>
<td>Urinary continence (0-4)</td>
<td>1.23</td>
<td>1.30</td>
<td>0.35</td>
<td>68.9</td>
</tr>
<tr>
<td>Bowel continence (0-4)</td>
<td>0.89</td>
<td>1.19</td>
<td>0.38</td>
<td>59.6</td>
</tr>
<tr>
<td><strong>Mental</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed (0-1)</td>
<td>0.21</td>
<td>0.40</td>
<td>0.16</td>
<td>58.5</td>
</tr>
<tr>
<td>Antipsychotics (0-1)</td>
<td>0.09</td>
<td>0.29</td>
<td>0.05</td>
<td>39.7</td>
</tr>
<tr>
<td>Anti-anxiety meds (0-1)</td>
<td>0.17</td>
<td>0.38</td>
<td>0.05</td>
<td>0.7</td>
</tr>
<tr>
<td>Antidepressants (0-1)</td>
<td>0.35</td>
<td>0.48</td>
<td>0.07</td>
<td>23.6</td>
</tr>
<tr>
<td>Hypnotics (0-1)</td>
<td>0.07</td>
<td>0.26</td>
<td>0.04</td>
<td>1.2</td>
</tr>
<tr>
<td>Delirium (0-6)</td>
<td>0.31</td>
<td>0.96</td>
<td>0.31</td>
<td>54.7</td>
</tr>
<tr>
<td><strong>Restraints</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bed rail restraints (0-2)</td>
<td>0.03</td>
<td>0.24</td>
<td>0.18</td>
<td>0.4</td>
</tr>
<tr>
<td>Trunk rail restraints (0-2)</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td>Limb rail restraints (0-2)</td>
<td>0.00</td>
<td>0.04</td>
<td>0.01</td>
<td>0.0</td>
</tr>
<tr>
<td>Other bed restraints (0-2)</td>
<td>0.00</td>
<td>0.04</td>
<td>0.01</td>
<td>0.0</td>
</tr>
<tr>
<td>Trunk chair restraints (0-2)</td>
<td>0.00</td>
<td>0.06</td>
<td>0.01</td>
<td>0.0</td>
</tr>
<tr>
<td>Limb chair restraints (0-2)</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td>Chair prevent rise restraint (0-2)</td>
<td>0.00</td>
<td>0.05</td>
<td>0.02</td>
<td>0.0</td>
</tr>
<tr>
<td>Other chair restraints (0-2)</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Activities of Daily Living (ADL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking (0-4)</td>
<td>2.94</td>
<td>1.06</td>
<td>0.37</td>
<td>54.9</td>
</tr>
<tr>
<td>Hygiene (0-4)</td>
<td>2.50</td>
<td>0.89</td>
<td>0.40</td>
<td>74.5</td>
</tr>
<tr>
<td>Dressing (0-4)</td>
<td>2.74</td>
<td>0.69</td>
<td>0.22</td>
<td>100.0</td>
</tr>
<tr>
<td>Locomotion (0-4)</td>
<td>2.73</td>
<td>0.97</td>
<td>0.35</td>
<td>42.0</td>
</tr>
<tr>
<td>Transfer (0-4)</td>
<td>2.76</td>
<td>0.70</td>
<td>0.23</td>
<td>85.3</td>
</tr>
<tr>
<td>Toilet (0-4)</td>
<td>2.78</td>
<td>0.69</td>
<td>0.23</td>
<td>93.6</td>
</tr>
<tr>
<td>Bed-mobility (0-4)</td>
<td>2.69</td>
<td>0.74</td>
<td>0.26</td>
<td>44.2</td>
</tr>
<tr>
<td>Eating (0-4)</td>
<td>1.14</td>
<td>1.10</td>
<td>0.48</td>
<td>56.7</td>
</tr>
</tbody>
</table>

Notes: Table shows summary statistics at admission for the 35 health measures that are used in our health index. Higher values of each health measure indicate worse health. Appendix B provides more detail on the measurement of each variable. Standard deviation across SNFs in column (3) is weighted by the number of patients in the SNF. The importance measure in column (4) is defined as the incremental $R^2$ of that variable, scaled such that the “most important” variable has an importance of 100 and the importance of other variables is reported relative to this value.
### Table 4: Health Index

<table>
<thead>
<tr>
<th>A. Internal fit (h as of day-30 assessment, N = 1,983,205):</th>
<th>Share of baseline sample</th>
<th>Mean h</th>
<th>Std. dev. of h</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients assessed at day 30</td>
<td>0.333</td>
<td>0.133</td>
<td>0.085</td>
</tr>
<tr>
<td>Discharged to community w/in 7 days</td>
<td>0.044</td>
<td>0.187</td>
<td>0.095</td>
</tr>
<tr>
<td>Still at the SNF 7 days after</td>
<td>0.272</td>
<td>0.125</td>
<td>0.080</td>
</tr>
<tr>
<td>Discharged elsewhere w/in 7 days</td>
<td>0.017</td>
<td>0.113</td>
<td>0.078</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. External fit (h as of initial assessment, N = 5,871,691):</th>
<th>Share of baseline sample</th>
<th>Mean h</th>
<th>Std. dev. of h</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients assessed at day 30</td>
<td>1.000</td>
<td>0.131</td>
<td>0.082</td>
</tr>
<tr>
<td>Discharged to community before day-30 assessment</td>
<td>0.373</td>
<td>0.177</td>
<td>0.086</td>
</tr>
<tr>
<td>Still at the SNF for day-30 assessment</td>
<td>0.414</td>
<td>0.106</td>
<td>0.065</td>
</tr>
<tr>
<td>Discharged elsewhere before day-30 assessment</td>
<td>0.213</td>
<td>0.101</td>
<td>0.070</td>
</tr>
</tbody>
</table>

**Notes:** Table shows summary statistics for our health index. Panel A shows the health index at 30 days for the sample of 1,983,205 patients with non-missing 30-day assessment data (before we perform the hotdeck imputation); Panel B shows the health index at admission for our full sample of 5,871,691 patients.
Table 5: Correlations of Value Added Estimates with SNF and Patient Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Correlation (1)</th>
<th>Partial correlation controlling for market FEs (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. SNF characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNF occupancy</td>
<td>-0.051</td>
<td>-0.028</td>
</tr>
<tr>
<td>Number of beds</td>
<td>-0.159</td>
<td>-0.070</td>
</tr>
<tr>
<td>For profit</td>
<td>-0.145</td>
<td>-0.081</td>
</tr>
<tr>
<td><strong>B. Patient characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black patient share</td>
<td>-0.187</td>
<td>-0.093</td>
</tr>
<tr>
<td>Medicaid share</td>
<td>-0.373</td>
<td>-0.210</td>
</tr>
<tr>
<td>Health index at admission</td>
<td>0.529</td>
<td>0.297</td>
</tr>
</tbody>
</table>

Notes: Table shows SNF-level correlations between our value added estimates and characteristics of both (A) SNFs and (B) their patients. Correlations are weighted by the number of patients in the SNF and calculated based on value-added estimates for 13,888 SNFs covering 655 markets.
## Table 6: Correlations of Value Added Estimates with Alternative Measures

<table>
<thead>
<tr>
<th></th>
<th>Correlation (1)</th>
<th>Partial correlation controlling for market FEs (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. CMS star ratings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>Quality</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Inspection</td>
<td>0.17</td>
<td>0.21</td>
</tr>
<tr>
<td>Staffing</td>
<td>0.31</td>
<td>0.26</td>
</tr>
<tr>
<td><strong>B. 90-day outcomes (signed so that higher is better)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of days in facilities 90 days</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>Home at 90 days</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>Spending per day 90 days</td>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>SNF readmission 90 days</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td>Hospital readmission 90 days</td>
<td>0.19</td>
<td>0.15</td>
</tr>
<tr>
<td>Alive at 90 days</td>
<td>0.38</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>C. Alternative health indices</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADL measure</td>
<td>0.79</td>
<td>0.80</td>
</tr>
<tr>
<td>Morris et al. (2018) measure</td>
<td>0.41</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>D. Other specifications</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average health</td>
<td>0.65</td>
<td>0.64</td>
</tr>
<tr>
<td>Average health improvement</td>
<td>0.74</td>
<td>0.72</td>
</tr>
<tr>
<td>Average health improvement + select in</td>
<td>0.56</td>
<td>0.84</td>
</tr>
<tr>
<td>Average health improvement + select out</td>
<td>0.72</td>
<td>0.71</td>
</tr>
</tbody>
</table>

**Notes:** Table shows SNF-level correlations between our baseline value added estimates and alternative measures of SNF quality, with and without controlling for market fixed effects. These alternative measures include: (A) CMS star ratings, (B) various cumulative outcomes within 90 days of SNF entry defined from Medicare claims, (C) alternative health indices defined from MDS data, and (D) alternative econometric specifications of value added using the health index. Correlations are weighted by the number of patients in the SNF and calculated based on value-added estimates for 13,888 SNFs covering 655 markets, except for Panel A which is based on the 13,550 SNFs (covering 655 markets) for which we observe star ratings.
A Market Construction

In order to estimate the discrete choice model that gives rise to the control function we derive in Appendix D, we need to define each patient’s choice set. Within this choice set (or “market”), each patient faces a discrete choice problem: she must choose one of the SNFs in the market, or the outside option (a SNF that is not in the market). We would like to allow this choice set to be as large as possible – in order to retain as many patients as possible in the first stage – but are constrained by computational challenges when the choice set gets too big.

Table A1 shows summary statistics for two standard health care market definitions – Hospital Referral Regions (HRRs) and Hospital Service Areas (HSAs) – as well as for the market definition we end up constructing. We match patient stays to these geographic markets using the 9-digit zip code of their residence. HRRs are designed to approximate health care markets for tertiary medical care, each containing at least one hospital that performs operations on the heart and brain (Evalitative Clinical Sciences 1998). Since our sample consists of Medicare patients who have stayed at a hospital prior to choosing a SNF for their post-acute care, HRRs seem like a natural market for SNF choice. However, for computational purposes they are too large. There are only 306 HRRs, and almost a third of the HRRs have more than 50 SNFs.

An alternative to HRRs therefore is to use HSAs, which are subsets of HRRs and are designed to approximate local markets for health care. They are groups of ZIP codes where residents receive most of their hospitalizations at hospitals from the area (Evalitative Clinical Sciences 1998). There are an order of magnitude more HSAs than HRRs, but the result is that HSAs are too small for our purpose. The average HSA contains fewer than 5 SNFs, and about 30% of the patients go to a SNF outside of their HSA, whereas only 12% of the patients go to a SNF outside their HRR.

To address the issue of too-large HRRs, without reverting to an HSA market definition, we therefore split large HRRs into what we call sub-HRRs, which are collections of adjacent HSAs within an HRR. Ideally we want these sub-HRRs to be pretty concentrated so we make the restriction that each sub-HRR must have at least 80% of its stays in the largest 20 SNFs. 182 HRRs already satisfy this restriction, while the other 124 do not. Therefore, we partition these 124 HRRs into sub-HRRs, as follows:

1. If the geographic unit has at least 80% of its stays in its top 20 SNFs, terminate the procedure and start step 1 for the next geographic unit. If not, move to 2.

2. Split the geographic unit into 2 groups of HSAs by k-means, clustering on the geographic coordinates of HSAs. If there is only a single HSA remaining, terminate the procedure for this unit and start step 1 for the next geographic unit. Otherwise, move to step 3.
3. If any of the groups of HSAs has less than 80% of its stays in its largest 20 SNFs then go back to step 2. Otherwise, terminate the procedure for this geographic unit and start step 1 for the next geographic unit. If there are no geographic units remaining, the market construction procedure is complete.

Steps 1 through 3 will construct sub-HRRs out of large HRRs. By design, the stays in each sub-HRR are highly concentrated in the largest 20 SNFs. Table A1 shows the results of this procedure. It results in 655 sub-HRRs, with an average of 21 SNFs per market and almost no markets with more than 50 SNFs, which substantially reduces the computational burden relative to using HRRs (with an average of 45 SNFs per HRR), and only a modest increase (18% compared to 12%) in the share of patients who choose a SNF outside their market.

Figure A1 shows the distribution of the number of SNFs per sub-HRR. The average sub-HRR has 21 SNFs and 50% of sub-HRRs have 20 SNFs or fewer.

B Health Measures Used in Health Index

Table A3 shows the 35 different health measures that we use as inputs. To be included, the measure had to be available every year from 2011 to 2016, be included on the discharge assessment, and to not be missing for more than 10% of patient-stays at the admission and 30-day assessments. In addition, we followed the clinical literature (Morris et al. 2018) and dropped 6 measures that relate to activities of daily living (ADL) and urinary tract health that are considered redundant. Finally, we collapsed measures of delirium and depression into standard indices. Specifically, for delirium, we follow Morris et al. (2018) and construct the total delirium score as the sum of three measures: inattention, disorganized thinking, and altered level of consciousness. Each of these components is scored on a scale of 0 to 2 (with 0 meaning the behavior is not present, 1 meaning the behavior is continuously present, and 2 meaning the behavior is present but fluctuates); as a result, our total delirium score ranges from 0 to 6.

For the depression measure, we follow the literature and create a binary depression score if the depression score on the PHQ-9 is 5 or more (Kroenke et al. 2001). The PHQ-9 is a standard score which ranges from 0 to 27. To construct it, we sum the following nine items together (each measure ranges from 0 to 3 where 0 indicates that the patient does not experience the symptom and 3 indicates the patient experiences the symptom nearly every day): little interest or pleasure in doing things; feeling down, depressed, or hopeless; trouble falling or staying asleep, sleeping too much; feeling tired or having little energy; poor appetite or overeating; feeling bad about yourself; trouble concentrating on things; going or speaking so slowly, or being extremely fidgety; thoughts that you would be better off dead or hurting yourself in some way.

Binary variables indicate whether or not a patient receives some treatment or experiences some condition. Categorical variables that range from 0 to 2 include weight loss and various physical restraints. Weight loss is defined as a decrease in weight that is 5% or more in the last month or 10% or more in the past six months. This variable takes the value 0 if there is no weight loss, 1 if the weight loss coincides with a physician-prescribed weight loss regimen, and 2 if there is weight loss but the patient is not on a physician-prescribed weight
loss regimen. For restraints 0 indicates restraints are not used, 1 indicates they are used less than daily, and 2 indicates they are used daily.

The only variable that takes integer values from 0 to 3 is shortness of breath. It takes the value 0 if the patient experiences no shortness of breath, 1 if the patient experiences shortness of breath in one state (exertion, sitting, lying down), 2 if the patient experiences shortness of breath in two states (exertion, sitting, lying down), and 3 if the patient experiences shortness of breath in all three states (exertion, sitting, lying down).

Variables that take values from 0 to 4 include continence variables and ADLs. The continence variables take the value 0 meaning the patient is always continent, 1 meaning less than 7 episodes of incontinence, 2 meaning 7 or more episodes or urinary incontinence but at least one episode of continent voiding, 3 meaning always incontinent, and 4 meaning resident had an ostomy or no bowel movement for the past 7 days.

All ADLs take values from 0 to 4, with 0 meaning independent with no help or staff oversight at any time, 1 meaning supervision (oversight, encouragement or cueing), 2 meaning limited assistance (resident highly involved in activity and staff provide guided maneuvering of limbs or other non-weight-bearing assistance), 3 meaning extensive assistance (resident involved in activity, staff provide weight-bearing support), and 4 meaning total dependence (full staff performance every time during entire 7-day period.)

C Characterizing Discharge Decisions

We model the SNF’s decision of whether to discharge someone to the community before day 30 as a function of their current health (see equation (2)). This abstracts from the possibility that this discharge decision could depend not only on their current health, but also on their health improvement, or the length of time they have been at the SNF.

Here we provide empirical support for our modeling decision. Specifically, we examine whether or not a patient is discharged to the community in the week following the 30-day assessment, and estimate both our baseline model in which this decision is only a function of current health, and a variety of richer models. The results suggest our baseline model is a good approximation of the discharge decision. For purposes of exploring the appropriate functional form for the discharge model, we pool data across SNFs, and ignore any SNF-level heterogeneity.

Table A4 shows the results. The first row shows results for our baseline discharge model. The second row shows results of estimating an augmented version of our baseline discharge model, in which we allow the discharge decision to depend not only on the patient’s current health level, but also on their health improvement since admission. Controlling for health improvement has almost no effect on the impact of the health level on the discharge probability, and the coefficient on health improvement is an order of magnitude smaller than that on the health level.

Additionally, we estimated a probit model of discharge to the community that allows for discharge to depend (linearly) on the length of time in the SNF. To do this, recall that we measure health at admission (typically around day 5), and (for people still in the SNF) at 14 days, and at 30 days. We can think of our health measures therefore as occurring roughly at week 1, week 2, and week 4 in the SNF.
The results are reported in the third column of Table A4. They show similar effects of health each week on the probability of discharge, indicating that discharge probabilities as a function of health do not vary with length of time in the SNF. To see this more clearly, Figure A2 shows the predicted probability of being discharged to the community in weeks 1 through 4 as a function of a patient’s health in the beginning of the week. We see these functions all lie basically on top of each other, supporting the idea that discharge to the community in a given week is predominantly a function of health at the beginning of the week, not time in weeks.

D Controlling for Selection In

In Section 3, we briefly described the discrete choice framework used to construct the control function for selection into SNFs. In this section, we provide more detail on the derivation of this control function.

Recall that utility is given by

\[ u_{ij} = \delta_j - \tau m_{ij} + \eta_{ij} \]

where \( \delta_j \) is a SNF fixed effect, \( m_{ij} \) is the (log) distance patient \( i \) must travel to SNF \( j \), and \( \eta_{ij} \) is the patient \( i \)'s idiosyncratic preferences for SNF \( j \). SNF value added, \( \alpha_j \), is given by

\[ E[h_{i2}|h_{i1}, \eta_{i1}, ..., \eta_{ij}, c_i] = \alpha_j + \theta h_{i1} + g_j(\eta_{i1}, ..., \eta_{ij}) \]

where \( g_j(\eta_{i1}, ..., \eta_{ij}) \) allows period 2 health to vary flexibly with unobserved preferences (i.e., to allow for arbitrary selection patterns) and \( c_i \) is a categorical variable that indicates the SNF chosen by patient \( i \). We assume that distance does not affect health conditional on these unobserved preferences so \( m_{ij} \) is excluded from the outcome equation above.

Following Dubin and McFadden (1984), we parameterize \( g_j(\cdot) \) as a linear function of the unobserved logit shocks \( \eta_{i1}, ..., \eta_{ij} \) according to

\[ E[h_{i2}|h_{i1}, \eta_{i1}, ..., \eta_{ij}, c_i] = \alpha_j + \theta h_{i1} + \sum_{k=1}^{J} \psi_k(\eta_{ik} - \mu_\eta) + \varphi(\eta_{ij} - \mu_\eta) \]

where \( \mu_\eta \) is the mean of the logit shocks (Euler’s constant) and acts as a normalization. The term \( \sum_{k=1}^{J} \psi_k(\eta_{ik} - \mu_\eta) \) captures preferences that have effects on the outcome that are not specific to SNF \( j \) while \( \varphi(\eta_{ij} - \mu_\eta) \) captures selection that is specific to SNF \( j \). Roy-type selection is indicated by \( \varphi > 0 \).

Integrating over the unobserved logit shocks \( \eta_{i1}, ..., \eta_{ij} \) yields

\[ E[h_{i2}|h_{i1}, \eta_{i1}, ..., \eta_{ij}, c_i] = \alpha_j + \theta h_{i1} + \sum_{k=1}^{J} \psi_k \beta_{ik}(j) + \varphi \beta_{ij}(j) \]

where \( \beta_{ik}(j) = E[\eta_{ik} - \mu_\eta|c_i = j] \) are the control functions. Letting \( j \) indicate the chosen
alternative, the control functions are

$$\beta_{ik}(j) = \begin{cases} 
-\log \hat{p}_{ik} & k = j \\
\frac{\hat{p}_{ik}}{1-\hat{p}_{ik}} \log \hat{p}_{ik} & \text{otherwise}
\end{cases}$$

(12)

where \(\hat{p}_{ik}\) is the predicted probability that patient \(i\) chooses SNF \(k\) based on her distance to SNFs. Since \(\hat{p}_{ik} < 1\), \(\log \hat{p}_{ik} < 0\), meaning that the control function takes a positive value when \(k = j\) and a negative value otherwise.

We will now derive the resulting expressions for these control functions. Again, letting \(j\) denote the chosen alternative, we know from rearranging the utility function that

$$E[\eta_{ij} - \mu_\eta | c_i = j] = E[u_{ij} | c_i = j] - \delta_j + \tau m_{ij} - \mu_\eta$$

(13)

From Small and Rosen (1981), we know

$$E[u_{ij} | c_i = j] = \log \left[ \sum_{k=1}^{J} \exp(\delta_k - \tau m_{ik}) \right] + \mu_\eta$$

(14)

Substitution yields,

$$E[\eta_{ij} - \mu_\eta | c_i = j] = \log \left[ \sum_{k=1}^{J} \exp(\delta_k - \tau m_{ik}) \right] + \mu_\eta
- \delta_j + \tau m_{ij} - \mu_\eta
$$

(15)

For a non-chosen alternative \(k \neq j\) we have

$$E[\eta_{ik} - \mu_\eta | c_i = j] = E[u_{ik} | c_i = j] - \delta_k + \tau m_{ik} - \mu_\eta$$

(16)

By proprieties of conditional expectation, we know

$$E[u_{ik}] = \Pr(c_i = k)E[u_{ik} | c_i = k] + \Pr(c_i \neq k)E[u_{ik} | c_i \neq k]$$

(17)

Using the result from Small and Rosen (1981), as before, and substituting the above expression yields

$$\delta_k - \tau m_{ik} + \mu_\eta = \Pr(c_i = k) \left[ \sum_{k=1}^{J} \exp(\delta_k - \tau m_{ik}) + \mu_\eta \right] + \Pr(c_i \neq k)E[u_{ik} | c_i \neq k]$$

(18)
which implies
\[
E[u_{ik}|c_i \neq k] = \frac{1}{1 - \Pr(c_i = k)} \left[ \delta_k - \tau m_{ik} + \mu \eta - \Pr(c_i = k) \left[ \sum_{k=1}^{J} \exp(\delta_k - \tau m_{ik}) + \mu \eta \right] \right]
\]
Substitution then yields
\[
E[\eta_{ik} - \mu \eta | c_i = j] = \frac{1}{1 - \Pr(c_i = k)} \left[ \delta_k - \tau m_{ik} + \mu \eta - \Pr(c_i = j) \left[ \sum_{k=1}^{J} \exp(\delta_k - \tau m_{ik}) + \mu \eta \right] \right]

- \delta_k + \tau m_{ik} - \mu \eta
\]
\[
= \frac{1}{1 - \Pr(c_i = k)} \left[ \delta_k - \tau m_{ik} + \mu \eta \right] - \Pr(c_i = k) \left[ \sum_{k=1}^{J} \exp(\delta_k - \tau m_{ik}) + \mu \eta \right] - \delta_k + \tau m_{ik} - \mu \eta
\]
\[
= \frac{\Pr(c_i = k)}{1 - \Pr(c_i = k)} \left[ \delta_k - \tau m_{ik} - \log \left( \sum_{j=1}^{J} \exp(\delta_k - \tau m_{ik}) \right) \right]
\]
\[
= \frac{\Pr(c_i = k)}{1 - \Pr(c_i = k)} \left[ \log \left( \exp(\delta_k - \tau m_{ik}) \right) - \log \left( \sum_{j=1}^{J} \exp(\delta_k - \tau m_{ik}) \right) \right]
\]
\[
= \frac{\Pr(c_i = k)}{1 - \Pr(c_i = k)} \left[ \log \left( \frac{\exp(\delta_k - \tau m_{ik})}{\sum_{j=1}^{J} \exp(\delta_k - \tau m_{ik})} \right) \right]
\]
\[
= \hat{p}_{ik} \log \hat{p}_{ik}
\]
(19)

E Hot Deck Imputation Procedure

Approximately 20% of patients, who are still in the SNF 30 days after admission, do not receive a 30-day health assessment. Moreover, assessments are not missing at random, with patients with missing 30-day assessments more likely to be discharged within 7 days of when their assessment would have been (Appendix Table [A6]). We use a “hot-deck” procedure to impute the 30-day health index for patients with missing information.

The hot-deck procedure takes patients with missing 30-day assessments, and finds a set of “donor” patients who share observable information with the missing patients, but are not missing their 30-day assessments. The missing health index is imputed by random selection from these donors. We require donor patients to be from the same SNFs. We additionally select donors using health index calculated from the health assessments at admission and 14 days (if available), along with data on discharge in the week following what should have
We observe initial health assessment data for all patients \((h_1)\); 14-day assessment for 68% percent of patients \((h_{1,5})\), and discharge destination with 7 days of the 30-day assessment (if any) for all patients \((d_2)\). Appendix Table A7 summarizes the imputation procedure in more detail.

Appendix Table A8 shows summary statistics of patients who are still in a SNF at the time of the 30-day assessment after we impute missing health at 30 days as needed. By the end of this procedure, we successfully impute health at 30 days for 99.7% of patients who are missing health at 30 days and are in the SNF at the time of the 30-day assessment. We also see that patients missing health at 30 days are on average healthier than those not missing health at 30 days. Specifically, patients missing health at 30 days have an average discretized health index of 6.72 compared to an average of 5.75 for those not missing health at 30 days. The hot decking procedure imputes health consistent with discharge destinations in the sense that among patient-stays only missing health at 30 days the average imputed discretized health of those discharged to the community is 8.05, while the average of those discharged elsewhere (to ACH, deceased, etc.) is 4.85. Patients still in the SNF have average imputed discretized health of 5.87 which is between the average imputed health of those discharged elsewhere and those discharged to the community. This makes conceptual sense, because it seems likely that unhealthy patients are discharged elsewhere and the healthiest patients are discharged to the community, while patients in the middle remain in the SNF.

### F Estimation Details

#### F.1 Deriving the Likelihood

The health process and upstream and downstream discharge equations are specified in Section 3. In our setting, patients can be partitioned into three groups:

1. Discharged downstream before period 2 \((d_i^D = 1)\)
2. Discharged upstream before period 2 \((d_i^U = 1)\)
3. Still in the SNF by period 2 \((d_i^D = 0 \cap d_i^U = 0)\).

Given this partition, the likelihood for a market is given by

\[
L = \Pi_i \left[ \Pr(d_i^D = 1) \right] \left[ \Pr(d_i^U = 1) \right] \left[ \Pr(d_i^D = 0 \cap d_i^U = 0 | h_{i,2}) f(h_{i,2}) \right]^{(1-d_i^D)(1-d_i^U)}. \tag{20}
\]

where \(f(h_{i,2})\) be the probability density function of \(h_{i,2}\).

We will derive an explicit formula for \(L\) by considering each partition separately. First,

---

To facilitate matching, we discretize the health index into 16 bins of width 0.025, ranging from 0 to 0.4. In Table A7, this discretized health index is denoted by \(\tilde{h}_t\) where \(t = 1\) for the 5-day assessment and \(t = 1.5\) for the 14-day assessment.
consider patients with \( d_i^U = 1 \). We know

\[
\Pr(d_i^D = 1) = \Pr(h_{i2} \geq \lambda_j + \nu_i) = \Pr(\alpha_j + \theta h_{i1} + \sum_{k \in I_i} \phi_k \beta_{ik} + \varphi \beta_{ij} + \epsilon_i \geq \lambda_j + \nu_i)
\]

\[
= \Pr(\epsilon_i - \nu_i \geq \lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in I_i} \phi_k \beta_{ik} - \varphi \beta_{ij})
\]

\[
= 1 - \Pr(\epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in I_i} \phi_k \beta_{ik} - \varphi \beta_{ij})
\]

\[
= 1 - \Phi\left( \frac{\lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in I_i} \phi_k \beta_{ik} - \varphi \beta_{ij}}{\sqrt{\sigma^2 + \sigma^2_{\nu}}} \right)
\]

\[
= \Phi\left( \frac{\alpha_j + \theta h_{i1} + \sum_{k \in I_i} \phi_k \beta_{ik} + \varphi \beta_{ij} - \lambda_j}{\sqrt{\sigma^2 + \sigma^2_{\nu}}} \right).
\]

where \( \Phi(\cdot) \) is the standard normal cumulative density function. This is the case because \( \epsilon_i \) and \( \nu_i \) are independent normal random variables, meaning that \( \epsilon_i - \nu_i \sim N(0, \sqrt{\sigma^2 + \sigma^2_{\nu}}) \).

Next, consider patients with \( d_i^U = 1 \). Notice that we can rewrite the portion of the likelihood for these patients as follows

\[
\Pr(d_i^U = 1) = \Pr(d_i^U = 1|d_i^D = 0) \Pr(d_i^D = 0)
\]

\[
= \Pr(d_i^U = 1|d_i^D = 0) \Phi\left( \frac{\lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in I_i} \phi_k \beta_{ik} - \varphi \beta_{ij}}{\sqrt{\sigma^2 + \sigma^2_{\nu}}} \right).
\]

Now consider \( \Pr(d_i^U = 1|d_i^D = 0) \). Since we model \( \Pr(d_i^U = 1|d_i^D = 0, h_{i2}) = \Phi(\gamma_0 + \gamma_1 h_{i2}) \) we have that

\[
\Pr(d_i^U = 1|d_i^D = 0) = \int \tilde{\phi}(\epsilon_i) \Phi(\gamma_0 + \gamma_1 (\alpha_j + \theta h_{i1} + \sum_{k \in I_i} \phi_k \beta_{ik} + \varphi \beta_{ij} + \epsilon_i)) d\epsilon_i
\]

\[
(23)
\]

where \( \tilde{\phi}(\cdot) \) is the probability density function of \( \epsilon_i \) conditional on \( d_i^D = 0 \) (see Section F.3 for a derivation of \( \tilde{\phi}(\cdot) \)). Together, this yields

\[
\Pr(d_i^U = 1) = \int \tilde{\phi}(\epsilon_i) \Phi(\gamma_0 + \gamma_1 (\alpha_j + \theta h_{i1} + \sum_{k \in I_i} \phi_k \beta_{ik} + \varphi \beta_{ij} + \epsilon_i)) d\epsilon_i
\]

\[
. \Phi\left( \frac{\lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in I_i} \phi_k \beta_{ik} - \varphi \beta_{ij}}{\sqrt{\sigma^2 + \sigma^2_{\nu}}} \right).
\]

\[
(24)
\]

Lastly, consider patients who remain in the SNF in period 2, meaning that \( d_i^D = 0 \cap d_i^U = 0 \). We know

\[
\Pr(d_i^D = 0 \cap d_i^U = 0| h_{i2}) f(h_{i2}) = \Pr(d_i^D = 0| h_{i2}) \Pr(d_i^U = 0| d_i^D = 0, h_{i2}) f(h_{i2}).
\]

\[
(25)
\]
It is trivial that \( f(h_{i2}) = \phi \left( \frac{h_{i2} - \alpha_j - \theta h_{i1} - \sum_{k \in J_i} \phi_k \beta_{ik} - \varphi_{ij}}{\sigma_{\epsilon}} \right) \) where \( \phi(\cdot) \) is the standard normal probability density function. Similarly, by assumption we know that \( \Pr(d_{i}^U = 0|d_{i}^D = 0, h_{i2}) = 1 - \Phi(\gamma_0 + \gamma_1 h_{i2}) \). Now consider \( \Pr(d_{i}^D = 0|h_{i2}) \). We know that

\[
\Pr(d_{i}^D = 0|h_{i2}) = \Pr(h_{i2} < \lambda_j + \nu_i | h_{i2})
\]

\[
= \Pr(\epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in J_i} \phi_k \beta_{ik} - \varphi_{ij} | h_{i2})
\]

\[
= \Pr(\epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in J_i} \phi_k \beta_{ik} - \varphi_{ij} | \epsilon_i).
\]

In order to derive an expression for \( \Pr(\epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in J_i} \phi_k \beta_{ik} - \varphi_{ij} | \epsilon_i) \) we must first understand the joint distribution of \( \epsilon_i - \nu_i \) and \( \epsilon_i \). Since these two random variables can be expressed as a linear combination of two independent normal random variables (namely, \( \epsilon_i \) and \( \nu_i \)), we know they are jointly normally distributed. Specifically,

\[
\begin{pmatrix}
\epsilon_i \\
\epsilon_i - \nu_i
\end{pmatrix}
\sim
N
\left(
\begin{pmatrix}
0 \\
0
\end{pmatrix},
\begin{pmatrix}
\sigma_{\epsilon}^2 & \sigma_{\epsilon}^2 \\
\sigma_{\epsilon}^2 & \sigma_{\nu}^2 + \sigma_{\epsilon}^2
\end{pmatrix}
\right)
\] (27)

The conditional distribution of a bivariate normal is also normal. Specifically,

\[
\epsilon_i - \nu_i | \epsilon_i \sim N(\epsilon_i, \sigma_{\nu}).
\] (28)

Hence, we have

\[
\Pr(d_{i}^D = 0|h_{i2}) = \Phi \left( \frac{\lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in J_i} \phi_k \beta_{ik} - \varphi_{ij} - \epsilon_i}{\sigma_{\nu}} \right)
\]

\[
= \Phi \left( \frac{\lambda_j - \alpha_j - \theta h_{i1} - \sum_{k \in J_i} \phi_k \beta_{ik} - \varphi_{ij} - h_{i2} + \alpha_j + \theta h_{i1} + \sum_{k \in J_i} \phi_k \beta_{ik} + \varphi_{ij}}{\sigma_{\nu}} \right)
\]

\[
= \Phi \left( \frac{\lambda_j - h_{i2}}{\sigma_{\nu}} \right).
\] (29)

Together this yields

\[
\Pr(d_{i}^D = 0 \cap d_{i}^U = 0|h_{i2}) f(h_{i2}) = 
\Phi \left( \frac{\lambda_j - h_{i2}}{\sigma_{\nu}} \right) (1 - \Phi(\gamma_0 + \gamma_1 h_{i2})) \phi \left( \frac{h_{i2} - \alpha_j - \theta h_{i1} - \sum \phi_k \beta_{ik} - \varphi_{ij}}{\sigma_{\epsilon}} \right).
\] (30)

Combining the three cases we considered above, the likelihood for any market is given by
\[ L = \Pi_i \left[ \Phi \left( \frac{\alpha_j + \alpha_h h_1 - \lambda_j}{\sqrt{\sigma^2 + \sigma^2_\nu}} \right) \right]^{d^P} \]

\[
\cdot \int \tilde{\phi}(\epsilon_i) \Phi(\gamma_0 + \gamma_1 (\alpha_j + \theta h_1 + \sum_{k \in I_i} \phi k \beta_{i k} + \varphi \beta_{ij} + \epsilon_i)) d\epsilon_i \Phi \left( \frac{\lambda_j - \alpha_j - \theta h_1 - \sum \phi k \beta_{i k} - \varphi \beta_{ij}}{\sqrt{\sigma^2 + \sigma^2_\nu}} \right)^{d^i_{UI}} \]

\[
\cdot \left[ \Phi \left( \frac{\lambda_j - h_{i2}}{\sigma_\nu} \right) (1 - \Phi(\gamma_0 + \gamma_1 h_{i2})) \right] \Phi \left( \frac{h_{i2} - \alpha_j + \theta h_1 - \sum \phi k \beta_{i k} - \varphi \beta_{ij}}{\sigma_\epsilon} \right)^{(1-d^i_{UI})(1-d^P)} \]

\[ \text{still in SNF} \]

For estimation purposes, we minimize the negative log-likelihood.

## F.2 Estimation

For computational reasons, we estimate the model in two steps by partitioning parameters into two groups: “national” parameters, \((\theta, \varphi, \sigma_\epsilon, \sigma_\nu)\), and market-level parameters, \((\alpha_j, \lambda_j, \phi_k, \gamma_0, \gamma_1)\). This procedure takes advantage of the fact that the market-level parameters only show up in the likelihood for the specific market, meaning that conditional on the national parameters (which show up in the likelihood for all markets), we can maximize the likelihood market-by-market which is computationally attractive since this is easily parallelizable.

The estimation process is as follows.

1. Specify grids for the “national” parameters \((\theta, \varphi, \sigma_\epsilon, \sigma_\nu)\)

For all combinations of points on these grids,

2. Maximize the likelihood with respect to \((\alpha_j, \lambda_j, \phi_k, \gamma_0, \gamma_1)\) within each market holding \((\theta, \varphi, \sigma_\epsilon, \sigma_\nu)\) fixed

After searching over all grid points,

3. Maximize the sum of the market-level likelihoods over the grid of “national” parameters.

Steps 1 through 4 are repeated as the grids specified in step 1 are fine tuned. Step 3 is accomplished by using a gradient based nonlinear solver within markets. The final grids for each parameter are

- \(\theta \in \{0.945, 0.9562, 0.9675, 0.9788, 0.99, 1.1\}\)
- \(\sigma_\epsilon \in \{0.05, 0.0563, 0.0625, 0.0688, 0.0750\}\)
- \(\sigma_\nu \in \{0.1, 0.15, 0.2, 0.25, 0.3\}\)
- \(\varphi \in \{-0.003, -0.0015, 0, 0.0015, 0.003\}\)
Another problem to tackle is estimating the following integral which enters the likelihood for patients who are discharged upstream, \( \int \tilde{\phi}(\epsilon_i) \Phi(\gamma_0 + \gamma_1 (\alpha_j + \theta h_{i1} + \sum_{k \in J_i} \phi_k \beta_{ik} + \varphi \beta_{ij} + \epsilon_i)) d\epsilon_i \). Due to runtime concerns, we estimate this integral using rectangular quadrature. We explore sensitivity to this method of estimating the integral by focusing on a large market and estimating the error that is introduced when estimating the integral via rectangular quadrature relative to Monte Carlo integration. We tested a variety of methods for choosing the location of nodes, the number of nodes, and the boundary of the integral (since the domain of integration is unbounded). Equally spaced nodes performed best and the percent error was not sensitive to the number of nodes. We tested limiting the boundary to within 4\( \sigma \), 5\( \sigma \), and 6\( \sigma \) of the mean. Results were essentially invariant to the choice of boundary this far out so in practice we limit the domain of integration to be within 4\( \sigma \) of the mean. The percent error was less than 0.01\% across all of the tests with equally spaced nodes.

F.3 Deriving the Marginal Density of \( \epsilon_i \) Donditional on \( d^P_i = 0 \)

In the likelihood derived in Appendix F.1, we saw that the probability density function of \( \epsilon_i \) conditional on \( d^P_i = 0 \) shows up in the likelihood for patients discharged upstream. This is a result of modeling the upstream decision conditional on \( d^P_i = 0 \). For estimation purposes, we need an explicit formula for \( \tilde{\phi}(\epsilon_i) \) since we estimate the integral by numerical quadrature. Specifically, we are interested in the following probability density function

\[
\tilde{\phi}(x) = \frac{d}{dx} \left[ \Pr (\epsilon_i \leq x | d^P_i = 0) \right] = \frac{d}{dx} \left[ \Pr (\epsilon_i \leq x | h_{i2} < \lambda_j + \nu_i) \right] = \frac{d}{dx} \left[ \Pr \left( \epsilon_i \leq x | \epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum \phi_k \beta_{ik} - \varphi \beta_{ij} \right) \right].
\]

To understand the functional form of \( \frac{d}{dx} \left[ \Pr (\epsilon_i \leq x | \epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum \phi_k \beta_{ik} - \varphi \beta_{ij}) \right] \) we must first revisit the joint distribution of \( \epsilon_i \) and \( \epsilon_i - \nu_i \). Recall from Section 1 that

\[
\begin{pmatrix} \epsilon_i \\ \epsilon_i - \nu_i \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + \sigma^2 \end{pmatrix} \right).
\]

Thus, the joint distribution of \( \epsilon_i, \epsilon_i - \nu_i | \epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum \phi_k \beta_{ik} - \varphi \beta_{ij} \) follows a bivariate truncated normal. Hence, \( \frac{d}{dx} \left[ \Pr (\epsilon_i \leq x | \epsilon_i - \nu_i < \lambda_j - \alpha_j - \theta h_{i1} - \sum \phi_k \beta_{ik} - \varphi \beta_{ij}) \right] \) is the marginal density of a bivariate truncated normal cumulative density function. Thus, we have that

\[
\tilde{\phi}(x) = \frac{1}{C} \phi \left( \frac{x}{\sigma_\epsilon} \right) \Phi \left( \frac{\lambda_j - \alpha_j - \theta h_{i1} - \sum \phi_k \beta_{ik} - \varphi \beta_{ij} - x}{\sigma_\nu} \right),
\]

where \( C \) is the total probability in the truncated distribution of \( (\epsilon_i, \epsilon_i - \nu_i) \) and is given by

\[
C = \int_{-\infty}^{\infty} \int_{-\infty}^{\lambda_j - \alpha_j - \theta h_{i1} - \sum \phi_k \beta_{ik} - \varphi \beta_{ij}} \exp \left\{ -\frac{\sigma^2 + \sigma^2}{2\sigma^2} \left[ \left( \frac{x}{\sigma_\epsilon} \right)^2 - 2\rho \left( \frac{x}{\sigma_\epsilon} \right) \left( \frac{y}{\sqrt{\sigma^2 + \sigma^2}} \right) - \left( \frac{y}{\sqrt{\sigma^2 + \sigma^2}} \right)^2 \right] \right\} dy dx.
\]
Appendix Figure A1: Distribution of SNFs per Sub-HRR

Notes: Figure shows the distribution of the number of SNFs per sub-HRR. Distribution is based on 655 sub-HRRs which are constructed as described in Appendix A.
Appendix Figure A2: Weekly Discharge Rule

Notes: Figure plots the results from estimating a weekly probit that allows for differential effect of health over time which relies on 12,610,109 patient-week observations. Specifically, we estimate $d_t(i, \text{comm}) = 1 \iff h^*_i = \rho_t h_{it} + \xi_i > \lambda$ where $d_t(i, \text{comm}) = 1$ if patient $i$ is discharged downstream during week $t$, $h^*_i$ is patient $i$’s latent health at discharge and $h_t$ and $h_{it}$ is health in week $t$. We let $\rho_t = \bar{\rho} + t\Delta \rho$. This figure shows $\rho_t h_{it}$ for weeks 1, 2, 3, and 4.
Notes: Figure correlates the coefficient on log-distance from the demand model across two specifications with different controls. The x-axis shows the coefficient from our main specification which only controls for the health index at admission. The y-axis shows an alternative specification that controls for the first ten principal components of the following controls: demographics (above median age, female, married, Hispanic, White, dual eligible), health measures at admission (walking, indwelling, falls, shortness of breath, depressed, delirium, vomiting, fever, dehydration, weight loss, and long-form ADL), and the top 20 CCS codes. Coefficients are shown for 648 of the 655 markets we observe in our data since 7 markets only have a single SNF.
Appendix Figure A4: Empirical Bayes Shrinkage

Notes: Figure shows a kernel density of value added as well as the analogous density after performing empirical bayes shrinkage.
Appendix Figure A5: Correlating Difference between 90th and 10th Percentile SNFs with Value Added Across Markets

Notes: Figure shows a scatter plot of the difference between the 90th and 10th percentile SNFs in terms of value added on the average value added within each of 648 markets. Seven markets are excluded because they only have a single SNF. The correlation is weighted by the number of patients in the market. The line is a linear fit weighted by the number of patients in each market.
Appendix Figure A6: Correlates of SNF Value Added

(A) Number of Beds

(B) SNF Occupancy

(C) For Profit Status

(D) Health Index at Admission

(E) Black Patient Share

(F) Medicaid Share

Notes: Figures show binscatters of value added on SNF characteristics using 20 bins for continuous SNF characteristics (discrete characteristics are shown as box plots). Number of beds, SNF occupancy, and for-profit status come from OSCAR data. The health index at admission, black patient share, and medicaid share are calculated using all 5,871,691 patient-stays in our full sample of Medicare patient-stays subject to the restrictions that the stay is the first of its episode, the patient is aged 65 or older and on Traditional Medicare, not on Medicare Advantage, not missing any of the 35 health measures used in our health index at baseline, not missing demographics, and the corresponding SNF has at least 50 stays.
Appendix Figure A7: Correlations of Value Added Estimates with Alternative Measures

(A) Quality Ratings

(B) Inspection Ratings

(C) Staffing Ratings

(D) Overall Ratings

(E) Share of Days in Facilities 90 days

(F) Home at 90 days
(G) Spending per day 90 days

(H) SNF Readmission 90 days

(I) Hospital Readmission 90 days

(J) Average Health

(K) Average Health Improvement

(L) Average Health Improvement + Select In
Notes: Figures show binscatters of alternative measures of SNF quality on SNF value added ($\alpha_j$) controlling for market fixed effects. The alternative measures are (A) CMS Quality Star Ratings, (B) CMS Inspection Star Ratings, (C) CMS Staffing Star Ratings, (D) CMS Overall Star Ratings, (E) share of days spent in facilities within 90 days of SNF admission, (F) the probability of being home 90 days after SNF admission, (G) spending per day within 90 days of SNF admission, (H) the probability of readmission to a SNF within 90 days of SNF admission, (I) the probability of hospital readmission within 90 days of SNF admission, (J) average health value added, (K) average health improvement value added, (L) average health improvement + select in value added, and (M) average health improvement + select out value added. Binscatters and correlations are weighted by the number of patients in the SNF. The y-axis of each binscatter starts from the mean of the variable minus 1.5 times the standard deviation to the mean plus 1.5 times the standard deviation.
**Appendix Table A1:** Summary Statistics for Different Geographic Units

<table>
<thead>
<tr>
<th>Geographic unit</th>
<th>HRR</th>
<th>HSA</th>
<th>Sub-HRRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of geographic units</td>
<td>306</td>
<td>2,991</td>
<td>655</td>
</tr>
<tr>
<td>Average SNFs</td>
<td>45.5</td>
<td>4.7</td>
<td>21.3</td>
</tr>
<tr>
<td>Average stays</td>
<td>19,544</td>
<td>2,000</td>
<td>8,952</td>
</tr>
<tr>
<td>Share with &gt; 20 SNFs</td>
<td>0.690</td>
<td>0.030</td>
<td>0.473</td>
</tr>
<tr>
<td>Share with &gt; 50 SNFs</td>
<td>0.294</td>
<td>0.004</td>
<td>0.023</td>
</tr>
<tr>
<td>Share of stays outside geog. unit</td>
<td>0.119</td>
<td>0.302</td>
<td>0.179</td>
</tr>
</tbody>
</table>

**Notes:** Table shows summary statistics for three different geographic units: HRRs, HSAs, and Sub-HRRs. Sub-HRRs – which are our definition of markets for SNF choice – are constructed from HRRs and HSAs as described in Appendix A.
### Appendix Table A2: Sample Restrictions

<table>
<thead>
<tr>
<th>Sample Restrictions</th>
<th>Stays</th>
<th>SNFs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample</td>
<td>24,078,049</td>
<td>16,355</td>
</tr>
<tr>
<td>Medicare and &gt;= 65</td>
<td>10,726,854</td>
<td>16,307</td>
</tr>
<tr>
<td>First stays</td>
<td>7,206,086</td>
<td>15,731</td>
</tr>
<tr>
<td>Demographics and health</td>
<td>6,547,959</td>
<td>15,638</td>
</tr>
<tr>
<td>CCS codes</td>
<td>6,031,753</td>
<td>15,624</td>
</tr>
<tr>
<td>5-day assessment between days 1 and 8</td>
<td>5,949,813</td>
<td>15,620</td>
</tr>
<tr>
<td>At least 50 episodes</td>
<td>5,905,261</td>
<td>13,867</td>
</tr>
</tbody>
</table>

**Notes:** Table shows the remaining number of patient stays and distinct SNFs after different sample restrictions. Full sample is all Medicare and Medicaid patient stays in in MDS 3.0 between October 01, 2011 to September 30, 2016. Medicare and ≥ 65 restrict to patients with Medicare who are at least 65 years old. Traditional Medicare restricts to patients who are always covered by Traditional Medicare and excludes patients coverage from Medicare Advantage in any month. First stays restricts to the first stay of each episode. CCS codes restricts to patients with a non-missing Clinical Classification Software (CCS) code from the prior hospital admission. We restrict to patient stays with a 5-day assessment between days 1 and 8, which aligns with the Medicare PPS assessment schedule inclusive of grace period. We exclude SNFs with fewer than 50 episodes.
### Appendix Table A3: Health Outcomes Used in the Regression Tree

<table>
<thead>
<tr>
<th>Physical</th>
<th>Restraints</th>
<th>Activities of daily living (ADL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has fallen (0-1)</td>
<td>Bed rail restraints (0-2)</td>
<td>Walking (0-4)</td>
</tr>
<tr>
<td>Vomiting (0-1)</td>
<td>Trunk rail restraints (0-2)</td>
<td>Hygiene (0-4)</td>
</tr>
<tr>
<td>Fever (0-1)</td>
<td>Limb rail restraints (0-2)</td>
<td>Dressing (0-4)</td>
</tr>
<tr>
<td>Dehydration (0-1)</td>
<td>Other bed restraints (0-2)</td>
<td>Locomotion (0-4)</td>
</tr>
<tr>
<td>Life prognosis &lt; 6 months (0-1)</td>
<td>Trunk chair restraints (0-2)</td>
<td>Transfer (0-4)</td>
</tr>
<tr>
<td>Anticoagulants (0-1)</td>
<td>Limb chair restraints (0-2)</td>
<td>Toilet (0-4)</td>
</tr>
<tr>
<td>Antibiotics (0-1)</td>
<td>Chair prevent rise restraint (0-2)</td>
<td></td>
</tr>
<tr>
<td>Diuertics (0-1)</td>
<td>Other chair restraints (0-2)</td>
<td>Bed-mobility (0-4)</td>
</tr>
<tr>
<td>Indwelling catheter (0-1)</td>
<td></td>
<td>Eating (0-4)</td>
</tr>
<tr>
<td>Weight loss (0-2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-breathness (0-3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary continence (0-4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bowel continence (0-4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed (0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antipsychotics (0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-anxiety meds (0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressants (0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypnotics (0-1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delirium (0-6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Table shows the 35 variables from the MDS 3.0 PPS Assessments that are used in construction of the health index. All variables take integer values and the range of the variable is shown in parentheses.
**Appendix Table A4: Alternative Discharge Models**

<table>
<thead>
<tr>
<th></th>
<th>Health in levels</th>
<th>Health in levels + health improvement</th>
<th>Weekly probit with health in levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.73</td>
<td>1.73</td>
<td>1.85</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$h_1$</td>
<td>4.19</td>
<td>4.16</td>
<td>5.25</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.018)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>$\Delta h_1$</td>
<td></td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.024)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,956,793</td>
<td>1,956,793</td>
<td>12,610,109</td>
</tr>
</tbody>
</table>

**Notes:** Tables shows the estimated parameters from three alternative discharge models specified in Appendix [C]. The first two specification are probits that use health in levels and health in levels with a health improvement term, respectively. These models rely on 1,956,793 patients a are assessed at 30 days. The third specification is a weekly probit that allows for differential effects of health over time which relies on 12,610,109 patient-week observations. Specifically, we estimate $d_t(i, \text{comm}) = 1 \iff h_{id}^* = \rho_t h_{it} + \xi_i > \lambda$ where $d_t(i, \text{comm}) = 1$ if patient $i$ is discharged downstream during week $t$, $h_{id}^*$ is patient $i$’s latent health at discharge and $h_t$ and $h_{it}$ is health in week $t$. We let $\rho_t = \bar{\rho} + t \Delta \rho$. 

74
**Appendix Table A5:** Health Index Fit for All Patients in a SNF at 30 days

<table>
<thead>
<tr>
<th>Fit for those who have a 30-day assessment (h as of day-30 assessment, N = 2,429,200):</th>
<th>Share of baseline sample</th>
<th>Mean h</th>
<th>Std. Dev. Of h</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients assessed at day 30</td>
<td>0.414</td>
<td>0.137</td>
<td>0.086</td>
</tr>
<tr>
<td>Discharged to community w/in 7 days</td>
<td>0.075</td>
<td>0.188</td>
<td>0.092</td>
</tr>
<tr>
<td>Still at the SNF 7 days after</td>
<td>0.311</td>
<td>0.127</td>
<td>0.081</td>
</tr>
<tr>
<td>Discharged elsewhere w/in 7 days</td>
<td>0.028</td>
<td>0.114</td>
<td>0.075</td>
</tr>
</tbody>
</table>

**Notes:** Table shows summary statistics for our health index at 30-days for all patients assessed at 30-days. This includes patients for whom we hot deck impute their health index at 30 days due to missing health measures. Details of this imputation procedure are presented in Appendix E.
**Appendix Table A6:** Discharge Destinations by 30 Day Assessment Availability

<table>
<thead>
<tr>
<th></th>
<th>Not missing health</th>
<th>Missing health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of patients in SNF at 30 days</td>
<td>0.797</td>
<td>0.203</td>
</tr>
<tr>
<td>Share discharged within 7 days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community (downstream)</td>
<td>0.109</td>
<td>0.474</td>
</tr>
<tr>
<td>ACH (upstream)</td>
<td>0.029</td>
<td>0.089</td>
</tr>
<tr>
<td>Hospice (upstream)</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>Death (upstream)</td>
<td>0.004</td>
<td>0.026</td>
</tr>
<tr>
<td>Elsewhere</td>
<td>0.010</td>
<td>0.034</td>
</tr>
<tr>
<td>Still in SNF</td>
<td>0.848</td>
<td>0.375</td>
</tr>
</tbody>
</table>

**Notes:** Table shows the discharge destinations of patients conditional on being assessed at 30-days (N = 2,429,200) for both patients who are missing and not missing the 30 day assessment.
**Appendix Table A7: Imputation Procedure**

<table>
<thead>
<tr>
<th>Step</th>
<th>Case 1 (N = 445,706)</th>
<th>Case 2 (N = 64,421)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Match on $\bar{h}<em>1$, $\bar{h}</em>{1.5}$, $d_2$</td>
<td>Match on $\bar{h}_1$, $d_2$</td>
</tr>
<tr>
<td></td>
<td>Match on $\bar{h}_{1.5}$, $d_2$, adjacent</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>$\bar{h}_{1.5}$</td>
<td>Match on $d_2$ and closest $\bar{h}_1$</td>
</tr>
<tr>
<td></td>
<td>Match on $\bar{h}_{1.5}$, $d_2$, adjacent</td>
<td>Match on $d_2$, but allow $d_2 = 0$</td>
</tr>
<tr>
<td>Step 3</td>
<td>$\bar{h}_1$</td>
<td>if $d_2=2$, closest $\bar{h}_1$</td>
</tr>
<tr>
<td>Step 4</td>
<td>Match on $d_2$, closest $\bar{h}<em>1$, $\bar{h}</em>{1.5}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Match on $d_2$, but allow $d_2 = 0$</td>
<td></td>
</tr>
<tr>
<td>Step 5</td>
<td>if $d_2=2$, closest $\bar{h}<em>1$, $\bar{h}</em>{1.5}$</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Table describes our hot deck imputation procedure which is described in Appendix E. Column 1 shows the procedure when we observe health at admission ($h_1$), health at 14 days ($h_{1.5}$) and discharge destination within 7 days of the 30 day assessment ($d_2$). Column 2 shows the procedure when we only observe $h_1$ and $d_2$. In the table, $\bar{h}_t$ corresponds to the discretized health. $d_t$ is a categorical variable that indicates the discharge destination of a patient within 7 days of the assessment at time $t$: still in the SNF (0), discharged downstream (1), or discharged upstream/elsewhere (2). $t = 1$ corresponds to the 5-day assessment, $t = 1.5$ corresponds to the 14-day assessment, and $t = 2$ corresponds to the 30-day assessment.
### Appendix Table A8: Summary Statistics by Imputation Cases

<table>
<thead>
<tr>
<th>Sample</th>
<th>Not missing $h_2$</th>
<th>Missing $h_2$</th>
<th>Only missing $h_2$</th>
<th>Missing $h_{1.5}$ and $h_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsample</td>
<td>d$_2$ = 0</td>
<td>d$_2$ = 1</td>
<td>d$_2$ = 2</td>
<td>d$_2$ = 0</td>
</tr>
<tr>
<td>Statistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,920,768</td>
<td>510,127</td>
<td>144,812</td>
<td>231,977</td>
</tr>
<tr>
<td>N Imputed</td>
<td>508,442</td>
<td>144,807</td>
<td>230,455</td>
<td>68,891</td>
</tr>
<tr>
<td>Share Imputed</td>
<td>0.9967</td>
<td>1.0000</td>
<td>0.9934</td>
<td>0.9996</td>
</tr>
<tr>
<td>Discretized $h_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5.75</td>
<td>6.72</td>
<td>5.87</td>
<td>8.05</td>
</tr>
<tr>
<td>SD</td>
<td>3.25</td>
<td>3.64</td>
<td>3.43</td>
<td>3.52</td>
</tr>
<tr>
<td>Number of candidate comparison observations chosen from</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>8.96</td>
<td>14.52</td>
<td>3.36</td>
<td>2.06</td>
</tr>
<tr>
<td>SD</td>
<td>24.91</td>
<td>31.88</td>
<td>6.12</td>
<td>2.19</td>
</tr>
</tbody>
</table>

**Notes:** Table shows summary statistics by imputation cases. Column (1) shows the non-imputed sample, while Column (2) shows the sample for whom $h_2$, health at 30 days, is imputed. Columns (3) and (4) break up the summary statistics in column (2) for two different cases: missing only $h_2$, and also missing $h_{1.5}$, health at 14 days; within each of these cases we further break down summary statistics by whether and where they were discharged within the next 7 days, which is described by d$_2$, a categorical variable that indicates the discharge destination of a patient within 7 days of the 30-day assessment: still in the SNF (0), discharged downstream (1), or discharged upstream/elsewhere (2).