ECONOMETRICA

PROVIDER INCENTIVES AND HEALTHCARE COSTS: EVIDENCE FROM LONG-TERM CARE HOSPITALS

LIRAN EINAV
Department of Economics, Stanford University and NBER

AMY FINKELSTEIN
Department of Economics, MIT and NBER

NEALE MAHONEY
Booth School of Business, University of Chicago and NBER

PROVIDER INCENTIVES AND HEALTHCARE COSTS: EVIDENCE FROM LONG-TERM CARE HOSPITALS

LIRAN EINAV
Department of Economics, Stanford University and NBER

AMY FINKELSTEIN
Department of Economics, MIT and NBER

NEALE MAHONEY
Booth School of Business, University of Chicago and NBER

We study the design of provider incentives in the post-acute care setting—a high-stakes but under-studied segment of the healthcare system. We focus on long-term care hospitals (LTCHs) and the large (approximately $13,500) jump in Medicare payments they receive when a patient's stay reaches a threshold number of days. Discharges increase substantially after the threshold, with the marginal discharged patient in relatively better health. Despite the large financial incentives and behavioral response in a high mortality population, we are unable to detect any compelling evidence of an impact on patient mortality. To assess provider behavior under counterfactual payment schedules, we estimate a simple dynamic discrete choice model of LTCH discharge decisions. When we conservatively limit ourselves to alternative contracts that hold the LTCH harmless, we find that an alternative contract can generate Medicare savings of about $2,100 per admission, or about 5% of total payments. More aggressive payment reforms can generate substantially greater savings, but the accompanying reduction in LTCH profits has potential out-of-sample consequences. Our results highlight how improved financial incentives may be able to reduce healthcare spending, without negative consequences for industry profits or patient health.

KEYWORDS: Healthcare, post-acute care, financial incentives, nonlinear contracts.

1. INTRODUCTION

HEALTHCARE SPENDING is one of the largest fiscal challenges facing the U.S. federal government. Within the healthcare system, post-acute care (PAC) is an under-studied sector, with large stakes for both spending and patient health. PAC refers to formal care provided to help patients recover from an acute care event such as a surgery. Medicare spending on PAC is substantial, about $60 billion in 2013, or about 20% more than the much-studied Medicare Part D program. Over 40% of hospitalized Medicare patients are discharged to PAC, and 13% of Medicare deaths involve a PAC stay in the prior 30 days. PAC spending is growing faster than overall Medicare spending and accounts for almost
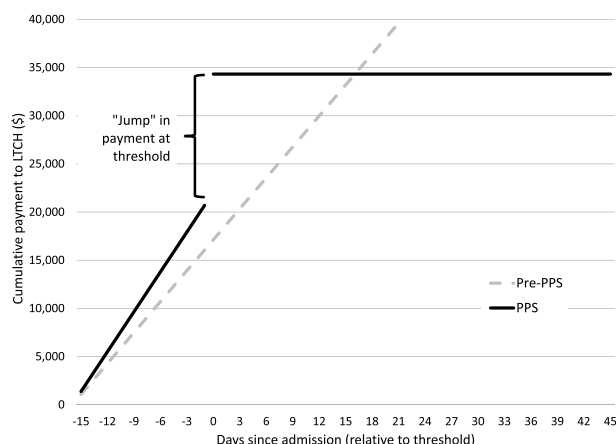
FIGURE 1.—LTCH payment schedules before and after PPS. Figure presents the payment schedule in both the pre-PPS and PPS periods. Sample pools admissions that are associated with different short stay outlier (SSO) thresholds, and *x*-axis is normalized by counting days relative to the threshold. The linear payment schedule begins with the first day of admission, and the *y*-axis is normalized to zero for day −16.

three-quarters of the unexplained geographic variation in Medicare spending (Newhouse, Garber, Graham, McCoy, Mancher, and Kibria (2013)).[1]

In this paper, we study the impact of provider financial incentives in determining patient flows and government spending in the Medicare PAC system. The PAC setting is attractive for several reasons. First, given its fiscal importance, understanding the effects of financial incentives is a natural area for inquiry. Second, the institutional environment—involving multiple interlocking and potentially substitutable settings that operate under different reimbursement regimes—suggests that financial incentives may have first-order consequences. Third, inefficiencies in the PAC sector have potentially important implications for public health, given that PAC patients are disproportionately high risk and might be more vulnerable to inefficiencies in the delivery of care.

Our analysis focuses on patients whose point of entry into the PAC system is a long-term care hospital (LTCH).[2] Medicare spending on LTCHs was about $5.5 billion in 2013, or slightly under 10% of Medicare PAC spending (MedPAC (2015a)). We focus on LTCH patients because of the sharp variation in provider incentives at this type of facility. This is illustrated in Figure 1: providers are reimbursed a daily amount (of approximately $1,300 on average) up to a threshold number of days, at which point there is a large (approximately $13,500 on average) jump in payments for keeping a patient an additional day beyond the threshold, but no payments for any days beyond it. We investigate the effects of this jump in payments using detailed Medicare claims data on the universe of LTCH stays over the 2007–2012 period, when this nonlinear payment schedule was in effect, as well as the 2000–2002 period, when LTCHs were instead reimbursed under a linear (i.e., constant per diem) payment schedule.

We start by briefly presenting descriptive evidence on the effect of the jump in payments on discharge behavior. While some of these results have been previously documented, we

---

[1]These statistics are taken from MedPAC (2004, 2015a), and MedPAC (2015b), with the exception of the statistic on deaths which we calculate using the data described in Section 2.

[2]The acronym LTCH is typically pronounced "el-tack," presumably reflecting the fact that LTCHs are sometimes referred to as long-term acute care hospitals (LTACs), which is pronounced in this manner.

present them to motivate our model of LTCH behavior. Discharges respond strongly to the payment increase, with the share of stays discharged increasing from 2% to 9% at precisely the day of the jump. The marginal patient discharged at the threshold appears to be much healthier than patients discharged beforehand: at the threshold, patients are disproportionately more likely to be discharged to a less intensive PAC facility or home ("downstream") than to an acute care hospital ("upstream"), and they have substantially lower post-discharge mortality than patients discharged on earlier days.

A natural question raised by this evidence is whether distortions in the timing of discharge have an impact on patient health. Given the high baseline mortality rate for LTCH patients (30% die within 90 days of LTCH admission), if the distortions are harmful, it seems plausible that we could detect an effect. Empirical analysis is challenging, however, because unlike discharge behavior, mortality effects may not appear right "at" the threshold. This challenge notwithstanding, we find no compelling evidence of mortality effects from the distortions in discharge behavior. There is no evidence of a change in the level or the slope of the mortality hazard in the vicinity of the threshold. We also find no indication of a mortality impact when we analyze the effect of small over-time changes in the day at which the jump in payments occurs. Of course, these results do not allow us to comprehensively rule out a mortality effect—we cannot, for instance, rule out an effect for every type of patient or at each and every hospital; and these results do not speak to adverse health effects that would not manifest in higher mortality rates. However, at minimum, they provide no "smoking gun" evidence of patient harm, and suggest that the marginal patients are able to receive similar care—at least in terms of mortality impact—whether they are located in LTCHs or in an alternative setting, which empirically is usually a less intensive PAC institution, such as a Skilled Nursing Facility (SNF).

Motivated by this descriptive evidence, we specify and estimate a dynamic model of LTCH behavior. The purpose of our model is to analyze how providers respond to the payment schedule on days further from the threshold, and to assess how treatment patterns and Medicare payments would be affected by counterfactual payment schedules. In our model, patients are characterized by their health, which evolves stochastically over time. LTCHs face a (daily) decision of whether to retain the patient or discharge her to another facility. The LTCH's objective function includes both net revenue (Medicare payments net of costs) and other, non-monetary considerations, such as patient outcomes. If the patient is discharged from the LTCH, the provider receives no subsequent net revenue, but internalizes potential consequences of the patient being treated in an alternative location. If the LTCH keeps the patient, it receives net revenue that depends on Medicare's payment schedule, while also accounting for the non-monetary outcomes associated with the patient being treated in the LTCH and the option value of making a similar discharge decision the following day. The provider therefore faces a standard dynamic discrete choice problem.

We estimate the model by simulated method of moments to match the observed discharge and mortality patterns under the linear and nonlinear payment schedules. We then use the estimated model to investigate the effects of alternative contracts that—like the observed contract—have a daily reimbursement rate up to a cap but that—unlike the observed contract—do not have a jump in payments at a threshold day. We find, for example, that if we were to lower the fixed payment to eliminate the jump in payments at the threshold, we would reduce total payments per admission for the episode of care by 25% on average, or about $13,000 per admission. However, such a payment schedule substantially reduces LTCH revenue and estimated profits, and therefore may have out-of-sample impacts on LTCH behavior that our estimates would not capture, such as inducing LTCH exit or lower service quality.

We therefore also engage in a more conservative set of counterfactuals in which we restrict attention to alternative contracts that would hold the LTCH harmless if their behavior did not change. Specifically, we consider the set of contracts that hold LTCH profits constant under their observed discharge schedule. Thus, if we apply this schedule and it triggers a "behavioral" response by LTCHs, they must be better off. Using our estimated model, we are able to identify a broad set of "win-win" payment schedules that reduce Medicare payments and, by construction, leave LTCHs (weakly) better off.[3] The contract that generates the largest savings reduces Medicare payments for the episode of care by 4.5%, and increases LTCH profits by 5.1%.

Our paper relates to a large literature examining how healthcare spending responds to financial incentives. Given the importance of healthcare spending in the economy and in public sector budgets, the existence of this large literature is not surprising. What is surprising—and arguably unfortunate from this perspective—is that the vast majority of this literature (including much of our own work) has concentrated on the impact of consumer financial incentives, such as deductibles and co-payments, while paying relatively less attention to the impact of provider financial incentives.[4] Existing work on provider-side incentives has focused on descriptive evidence that providers do, indeed, respond to incentives, with much of the evidence coming from the introduction of the Inpatient Prospective Payment System in 1983 (Cutler (1995), Cutler and Zeckhauser (2000)). More recently, Clemens and Gottlieb (2014) and Ho and Pakes (2014) provided a rare look at the behavioral response of physicians to financial incentives.

The relative lack of research on the provider side presumably reflects the difficulties in finding clean variation in incentives. Perhaps not surprisingly, the sharp incentives created by the LTCH payment schedule have already received some attention in academic (Kim, Kleerup, Ganz, Ponce, Lorenz, and Needleman (2015)), popular (Weaver, Mathews, and McGinty (2015)), and policy (MedPAC (2016)) spheres. Our descriptive work on discharges around the threshold is quite similar to this prior work, while our descriptive analysis of the health of the marginal dischargee and of mortality effects is new.

Our paper is most closely related to Eliason et al. (forthcoming) who—in concurrent independent work—also studied the impact of the LTCH payment schedule on discharge behavior descriptively and through the lens of a dynamic model. Our findings and those of Eliason et al. are very much in concert. Both papers present evidence that LTCHs' discharge decisions strongly respond to the sharp financial incentives at the threshold, and each paper develops a dynamic model to simulate the impact of alternative payment policies, the results of which (when comparable) are also very similar. Our study places a greater emphasis on the impact on patient outcomes and examines a somewhat different set of counterfactual payment policies, but restricts attention to the average response. In contrast, Eliason et al. allowed for and placed greater emphasis on the heterogeneity in the behavioral response across LTCHs and patient demographics.

Finally, from a more conceptual perspective, our paper is related to a growing literature that seeks to interpret descriptive evidence of the behavioral responses to nonlinear payment schedules ("bunching") through the lens of economic models that allow for assessments of behavior under counterfactual schedules (e.g., Chetty, Friedman, Olsen, and

---

[3]Given the lack of compelling evidence of mortality effects at the threshold, it seems reasonable to assume that mortality is unlikely to be impacted much under these "LTCH held harmless" alternative contracts.

[4]The majority of healthcare spending, however, is accounted for by a small share of high-cost individuals whose spending is largely in the "catastrophic" range where deductibles and co-payments no longer bind; for example, 5% of the population account for 50% of healthcare expenditures (Cohen and Yu (2012)). It seems likely that for such patients, consumer cost-sharing may have little impact relative to provider-side incentives.

Pistaferri (2011), Einav, Finkelstein, and Schrimpf (2015), Manoli and Weber (2016), Bajari, Hong, Park, and Town (2017), Dalton, Gowrisankaran, and Town (2017)).

The rest of the paper proceeds as follows. Section 2 provides some background on the PAC sector, LTCHs, and our data. In Section 3, we describe the discharge and mortality patterns around the jump in payments. Section 4 motivates the need for dynamics, presents the model, and discusses estimation and identification. Section 5 presents the estimation results and the impact of counterfactual payment policies. Section 6 concludes.

2. SETTING AND DATA

2.1. *Post-Acute Care in the United States*

Post-acute care (PAC) is the term for rehabilitation and palliative services provided to patients recovering from an acute care hospital stay. In the United States, the Center for Medicaid and Medicare Services (CMS) associates PAC with three types of facilities—long-term care hospitals (LTCHs), skilled nursing facilities (SNFs), and inpatient rehabilitation facilities (IRFs)—as well as care at home provided by home health agencies (HHAs) (MedPAC (2015b)). In 2013, Medicare paid $60 billion to PAC providers, approximately 16% of the $368 billion paid that year in Traditional Medicare (TM) claims; PAC facilities constitute about 70% of total PAC spending, with the remaining 30% associated with HHAs (MedPAC (2015a)).

In recent years, the geographic variation and growth rate of spending on PAC have raised concerns about the efficiency of the sector. From 2001 to 2013, Medicare spending on PAC grew at an annual rate of 6.1%, two percentage points higher than the rate of spending growth for TM as a whole (The Boards of Trustees for Medicare (2002, 2014), MedPAC (2015a)). A recent Institute of Medicine report found that, despite accounting for only 16% of spending, PAC contributed to a striking 73% of the unexplained geographic variation in spending, suggesting that there may be substantial inefficiencies in the sector (Newhouse et al. (2013)).

It is useful to think about patients as generally flowing "downstream" through the healthcare system. Upon experiencing an acute health event, they go to a regular Acute Care Hospital (ACH); from there they may be sent to a PAC facility to recover, and eventually go home once they are sufficiently healthy and can function independently. Some ACH patients "skip" the PAC stay and return home directly from the ACH, and some patients relapse and move "upstream" from a PAC facility back to an ACH.

The top panel of Figure 2 gives a sense of transitions among ACHs, PAC facilities (LTCHs, SNFs, and IRFs), home (including HHAs), and death (including hospice). (Throughout the rest of the paper, we use the term PAC facilities to refer to LTCHs, SNFs, and IRFs, because these are facilities that provide in-house care, in contrast to HHAs, which provide care at the patient's home.) In our data, described below, 26% of patients who are discharged from an ACH receive follow-up care from a PAC facility.[5] From these PAC facilities, 60% of patients continue to flow home, where they may still receive treatment from an HHA, while 34% are discharged back to an ACH. The remaining 6% are discharged to a hospice or due to death.

Just like the natural flow of patients into and out of the PAC system, there is also a general ordering of care within it. LTCHs provide the most intensive care, SNFs and IRFs

---

[5]In analysis that includes HHAs in the calculation, the share of ACH patients who are discharged to PAC rises to 42% (MedPAC (2015b)).
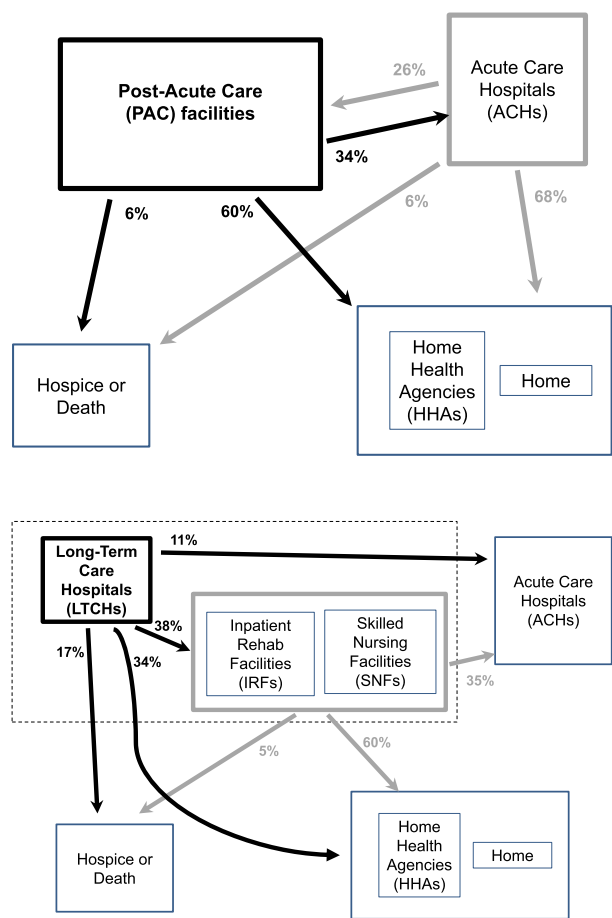
FIGURE 2.—Patient flow into and out of post-acute care. Top panel shows patient flow from acute care hospitals (ACHs) to the different destinations: post-acute care (PAC) facilities; home and home health agencies; and death or hospice. Post-acute care facilities include Long-Term Care Hospitals (LTCHs), Skilled Nursing Facilities (SNFs), and Inpatient Rehabilitation Facilities (IRFs). Bottom panel shows how the patient flow pattern is different, within PAC, between Long-Term Care Hospitals (LTCHs) and other PAC facilities (SNFs and IRFs). All numbers are calculated using the universe of Traditional Medicare admissions during the PPS period (October 2007 to July 2012). Numbers are shares of total discharges from each type of facility, excluding a small share of discharges (never greater than 5%) that are more difficult to classify. See Appendix A for more details.

provide less intensive care, and HHAs provide the least intensive bundle of medical services. Severity of Illness categories are a commonly used measure of intensity of care, and are constructed using the patient's age, diagnoses, procedures, and comorbidities. The share of patients in the highest severity of illness category declines from 43% at LTCHs, to approximately 12% at SNFs and IRFs, to 4% at HHAs (AHA (2010)). Medicare payments per day follow the same declining pattern.

Our point of entry into the PAC landscape is through admission to an LTCH. Virtually all LTCH admissions are from an ACH. The bottom panel of Figure 2 looks at patient flows from LTCHs. About 11% of LTCH patients are discharged back to an ACH, 38% are discharged to another PAC facility (SNF or IRF), and 34% are discharged home, where they may continue to receive care from an HHA. The remaining 17% are dis-

charged to a hospice (4%) or die within the LTCH (13%). In contrast, once in a SNF or IRF, patients almost never get discharged to an LTCH, die much less frequently (5%), and much more often (60%) return directly home.

Despite the interlocking nature of the PAC system, the way that Medicare reimburses post-acute care varies substantially by the setting. Historically, all providers were paid according to an administrative estimate of their costs. Since the early 2000s, however, Medicare has shifted to paying PAC providers under separate prospective payment systems that vary based on the type of provider. Loosely, HHAs are paid per 60-day episode-of-care, SNFs are paid a fixed rate per day, and IRFs and LTCHs are paid a fixed amount per admission (like ACHs). We provide more details on LTCH payments in Section 3.

The fact that each type of facility is paid under a different system has raised concerns. From a public health perspective, there is concern that the separate payment systems do not give providers enough incentive to coordinate care across different facilities. From a budgetary perspective, there is concern that providers may shuffle patients across facilities with the aim of increasing Medicare payments. These concerns have spurred various proposals for payment reform, including a recent bill which proposes providing a "bundled payment" to a single PAC coordinator, and letting this coordinator internalize the costs and benefits associated with the sequence of admissions and discharges for the entire episode of care (H.R.1458: BACPAC Act of 2015).

2.2. *Long-Term Care Hospitals*

Our primary focus is on patients whose point of entry into the PAC system is a long-term care hospital (LTCH). The demarcation "LTCH" describes how the provider gets paid by Medicare. It is a regulatory concept, rather than a medical one. For a hospital to get paid as an LTCH, it must have an average inpatient length of stay of 25 days or more. Naturally, there are many ways to meet this requirement; from a medical standpoint, the question of what an LTCH is or does is not well-defined.

The LTCH category of hospitals was created to solve a potential "side effect" of the 1982 Tax Equity and Fiscal Responsibility Act (TEFRA), which established the prospective payment system (PPS) for acute care hospitals. Under the new PPS, hospitals were paid per discharge, and not based on their costs, as a way to provide incentives for hospitals to be efficient in their treatment decisions. Regulators who were designing the PPS realized that there was a small number of hospitals that had long average length-of-stays (LOS) and would not be financially viable under the fixed-price PPS. LTCHs were thus created as a carve-out from PPS for hospitals that had an average LOS of at least 25 days. At that point in time, there were 40 hospitals that qualified as LTCHs—mainly former tuberculosis and chronic disease hospitals in the Boston, New York City, and Philadelphia metropolitan areas. LTCH payments were based on costs measured in 1982, roughly in the spirit of the pre-1982 payment system, and adjusted for inflation in subsequent years. See Liu, Baseggio, Wissoker, Maxwell, Haley, and Long (2001) for more on the background of the LTCH sector.

Over the last 30 years, and perhaps because of the LTCH exemption from PPS, there was rapid growth in the LTCH sector. Because new entrants did not have cost data for 1982, payments for new entrants were determined by costs in their initial years of operation. This encouraged new entrants to be inefficient when they first opened and earn profits by increasing their efficiency over time (Liu et al. (2001)). From the initial 40 hospitals first designated as LTCHs in 1982, there are now over 400 LTCHs in the country.

Geographic penetration of LTCHs is extremely varied. There are only a few LTCHs in the west of the country, and three states (Massachusetts, Texas, and Louisiana) account

for a third of all LTCHs (Liu et al. (2001)). In places where there are LTCHs, these hospitals are an important part of Medicare's PAC landscape. For instance, in hospital service areas (HSAs) with at least one LTCH, we calculate that LTCHs account for 13% of Medicare PAC facility days and 28% of Medicare PAC facility spending; nationwide, payments to LTCHs account for 12% of Medicare PAC facility spending.[6]

LTCHs are much more likely to be for-profit than other medical providers. According to 2008 data from the American Hospital Association (AHA), 72% of LTCHs are for-profit (versus 17% for ACHs), 22% are non-profit, and 6% are government run. The LTCH market is dominated by two for-profit companies, Kindred Health Systems and Select Medical, which run about 40% of LTCHs, according to the AHA data. Company reports indicate that LTCHs are highly profitable. For their business segments that include LTCHs, Kindred's profits have hovered between 22% and 29% of revenue and Select's profits have ranged between 16% and 22% of revenue.[7]

Approximately half of LTCHs are known as Hospitals-within-Hospitals (HwHs), meaning that they are physically located within the building or campus of an ACH but have a separate governing body and medical staff. Regardless of their location, LTCHs tend to have strong relationships with a single ACH (MedPAC (2004)). Because of concerns over close relationships between LTCHs and their partner ACHs, in 2005 CMS established a policy known as the "25-percent rule" that creates disincentives for admitting more than 25% of patients from a single facility; however, Congress has delayed the full implementation of the law.[8]

2.3. *Data*

Our main analysis focuses on patients who are admitted to an LTCH and follows them throughout their entire healthcare episode. Our primary data source is the Medicare Provider and Analysis Review (MedPAR) data, spanning the years 2000–2012. The data set contains claim-level information on discharges from ACHs, LTCHs, SNFs, and IRFs. Each record is a unique stay for which a claim was submitted, and the data contain information on procedures, admission and discharge dates, admission sources and discharge destinations, hospital charges, and Medicare payments. The MedPAR data also provide us with basic demographic information such as the age, sex, and race of the beneficiary, and information about the patient's diagnoses.

We supplement this primary source with several ancillary data sources. First, we use Medicare's beneficiary summary file to approximate the (quite small) post-LTCH discharge payments to hospices and HHAs, as well as post-LTCH discharge hospice days; Appendix A provides more details. Second, we use Medicare's beneficiary files to determine whether the beneficiary is dually eligible for Medicare and Medicaid and the date of death. A key advantage of these data is that they allow us to observe death *regardless of whether and where the patient is receiving care.* Third, we use the Medicare chronic conditions file to measure whether the individual has any of 27 chronic conditions in the

---

[6]Statistics calculated using the 2007–2012 MedPAR data described below.

[7]Profits are defined as EBITA (earnings before interest, taxes, and amortization). Kindred's profits are based on 2009 to 2015 company reports. Prior to 2009, Kindred did not separate out their reporting of LTCH profits from the much larger SNF category. Select's profits are based on company reports from 2004 to 2015.

[8]There is also a regulation known as the "5-percent rule" that addresses the incentive for HwHs to "ping-pong" patients between the ACH and LTCH. If more than 5% of patients who are discharged from an LTCH to an ACH are readmitted to the LTCH, the LTCH will be compensated as if the patient had a single LTCH stay (42 CFR 412.532).

calendar year prior to the LTCH stay. Finally, we use data from the American Hospital Association (AHA) survey over the same period to determine whether an LTCH is for-profit, non-profit, or government owned, and whether it is co-located with an ACH.

Our analysis focuses on the current Medicare payment schedule for LTCHs, known as LTCH-PPS. We analyze the time periods before and after full implementation of LTCH-PPS, which was phased in over a five-year period starting on October 1, 2002. We define the pre-PPS period as discharges that occurred from January 1, 2000 to September 30, 2002. For this period, we measure post-discharge payments, days, and mortality through March 31, 2003, which is six months after the last LTCH discharge. We exclude the October 2002 to September 2007 phase-in period because provider behavior during this period potentially reflects the combination of changing financial incentives and learning about the new incentive structure, complicating the interpretation of the data. We define the PPS period as discharges that occurred from October 1, 2007 to July 31, 2012, and analyze post-discharge payments, days, and mortality through December 31, 2012, which is similarly six months after the last LTCH discharge.

Table I shows summary statistics on ACH, LTCH, and SNF/IRF admissions in the pre-PPS and PPS periods.[9] Since an observation is an admission, some patients (16%) show up multiple times in the data. LTCH patients are, on average, slightly younger than ACH patients and much younger than SNF/IRF patients. LTCH patients are also almost twice as likely to be black and about one-third more likely to be eligible for Medicaid, relative to ACH and SNF/IRF patients. These differences are fairly stable over time. In terms of health, LTCH patients appear less healthy than those in an ACH or SNF/IRF. LTCH patients have more chronic conditions prior to the stay and higher mortality. For example, about 15% of LTCH patients die within 30 days of admission and 30% die within 90 days; these mortality rates are about 50% larger than mortality rates for SNF/IRF patients and about twice as large as those for ACH patients.

In terms of the intensity of medical care, LTCH stays are closer to ACH stays than stays at an SNF/IRF. The majority of LTCH and ACH patients receive at least one medical procedure versus about 2% of patients who visit an SNF/IRF. The most common LTCH procedures (cardiac catheterization and blood transfusion) are also more similar to those that occur at an ACH, relative to occupational and physical therapy, which are the most common procedures in SNF/IRF. Length of stay at an LTCH, however, is (by design) much more similar to that of a SNF/IRF. The average stay at an ACH is 5 days, while it is just over 25 days in LTCH and SNF/IRF.

The bottom rows of Table I show statistics on Medicare and out-of-pocket payments. Medicare payments in the PPS period average $2,074 per day at an ACH, $1,391 per day at an LTCH, and $507 per day at a SNF/IRF. However, because LTCH stays are much longer than ACH stays, per-admission Medicare payments at LTCHs average over $35,000, which is three times greater than per-admission ACH and SNF/IRF payments. Out-of-pocket payments at ACHs and LTCHs arise from Medicare's Part A deductible ($1,156 in 2012) and from co-insurance payments that apply when the patient has more than 60 hospital days in the benefit period ($289 per day in 2012). Because patients have no out-of-pocket exposure between the deductible and their 60th hospital day, out-of-pocket payments are a modest 7.7% of Medicare payments at ACHs and 5.4% at LTCHs in the PPS period. SNFs, on the other hand, have a separate co-insurance schedule with

---

[9]We group SNF and IRF admissions together for convenience, as both represent post-acute care that is "less intense" than an LTCH and because IRFs only account for a small (6.4%) fraction of these admissions.

TABLE I

SUMMARY STATISTICS

	Pre-PPS (Jan 2000–Sep 2002)			PPS (Oct 2007–Jul 2012)		
	ACH	LTCH	SNF/IRF	ACH	LTCH	SNF/IRF
Number of stays (000s)	29,362	219	5,187	47,940	587	11,237
Panel A. Patient attributes						
Average age	74.5	73.8	80.2	73.4	71.6	79.1
Fraction male	0.43	0.44	0.34	0.44	0.49	0.37
Fraction white	0.84	0.74	0.87	0.82	0.73	0.85
Fraction black	0.11	0.20	0.10	0.13	0.20	0.11
Fraction aged 65+	0.86	0.83	0.94	0.81	0.75	0.91
Fraction dual eligible	0.24	0.31	0.27	0.27	0.38	0.27
Panel B. Patient health indicators						
Number of chronic conditions[a]	3.8	6.4	4.4	4.9	7.8	5.6
30-day mortality since admission	0.082	0.142	0.112	0.078	0.158	0.088
90-day mortality since admission	0.142	0.274	0.218	0.139	0.306	0.186
Fraction home within 90 days[b]	0.807	0.558	0.497	0.793	0.460	0.554
Three most common DRGs:[c]	Joint Repl. (3.9%) Septicemia (2.8%) Dig. Disorders (2.1%)				Ventilator (10.7%) Resp. Failure (8.3%) Septicemia (5.7%)	Rehab w/ CC (17.1%) Rehab w/o CC (10.2%) Ungroupable (3%)

(*Continues*)

TABLE I—*Continued*

	Pre-PPS (Jan 2000–Sep 2002)			PPS (Oct 2007–Jul 2012)		
	ACH	LTCH	SNF/IRF	ACH	LTCH	SNF/IRF
Panel C. Procedures during stay						
Length of stay[d]	5.6	26.6	24.0	5.2	25.3	25.6
Fraction with no procedures	0.43	0.61	0.95	0.40	0.28	0.98
Number of procedures (cond. on any)	2.5	2.4	2.0	2.6	2.8	2.1
Three most common procedures:	Transfusion (6.2%) Arteriography (5.5%) Cardiac cath. (5.2%)	Cath (7.5%) Transfusion (5.5%) Occ. Therapy (5.0%)	Phys. Therapy (2.6%) Occ. Therapy (2.4%) Transfusion (0.3%)	Transfusion (9.9%) Cath. (6.5%) Dialysis (4.5%)	Cath (19.7%) Transfusion (17.8%) Ventilation (14.4%)	Occ. Therapy (1.3%) Phys. Therapy (1.2%) Transfusion (0.2%)
Panel D. Payments and cost (2012 $)						
Total Medicare payments per stay	9,415	28,351	9,860	10,816	35,216	12,953
Medicare payments per day	1,672	1,068	412	2,074	1,391	507
Out-of-pocket payments	772	2,336	1,618	836	1,907	1,951
Out-of-pocket payments per day	137	88	68	160	75	76
Total reported costs	–	28,351	–	–	36,092	–
Reported cost per day	–	1,068	–	–	1,426	–

[a]Number of chronic conditions is measured in the calendar year prior to the stay.

[b]Reports fraction home at least once during the 90 days after admission, where "home" means alive and not in a facility (ACH, LTCH, SNF/IRF, or hospice).

[c]DRG groupings changed between the pre-period and post-period, so for simplicity we report this only for the post-period.

[d]Length of stay is censored at 100 days for SNFs, since after that Medicare does not pay and therefore further days are not observed. This applies to about 2% of stays.

TABLE II

POST-DISCHARGE OUTCOMES[a]

	Pre-PPS (Jan 2000–Sep 2002)			PPS (Oct 2007–July 2012)		
	Overall	Upstream	Downstream	Overall	Upstream	Downstream
Number of discharges (000s)	188.7	41.8	147.0	509.7	80.3	429.4
Post-discharge 30-day mortality	11.2	24.9	7.4	14.2	47.6	8.0
Post-discharge 90-day mortality	20.2	37.5	15.3	24.3	60.0	17.6
Post-discharge payments[b]	13,100	31,405	7,901	22,808	35,775	20,382
Post-discharge facility days[b]	17.1	32.8	12.6	26.1	33.0	24.8

[a]Table presents summary statistics on post-discharge costs and facility days using the baseline sample of LTCH stays described in Table I, excluding discharges due to death.

[b]Post-discharge payments and post-discharge days refer to the entire post-discharge episode of care, which we define as beginning at the day of discharge and ending when there are two consecutive days with no payments from either an ACH, SNF/IRF, or LTCH.

payments of $144.50 per day in 2012 for stays in excess of 20 days, and a much higher out-of-pocket share.

Our analysis encompasses not only the experience of the patient in the LTCH (i.e., length of stay and payments) but also their post-discharge experience. We define a post-discharge episode of care as the spell of continuous days with a Medicare payment to an ACH, SNF/IRF, or LTCH; the episode ends if there are two days or more without any Medicare payments being made to any of these institutions. For each post-discharge episode, we report 30-day mortality, 90-day mortality, post-discharge Medicare payments, and post-discharge facility days (i.e., days in an ACH, SNF/IRF, LTCH, or hospice). Table II shows summary statistics on post-discharge outcomes. Focusing on the PPS period, about one-quarter of LTCH patients die within 90 days of discharge. Average length of stay in the post-discharge episode of care is 26 days, which is similar to the average time in the LTCH (see Table I). Average post-discharge Medicare payments is $22,808, about 60% of Medicare payments to the LTCH (see Table I).

In some of our analyses below, we find it useful to classify live discharges from the LTCH as either "upstream" or "downstream" based on their discharge destination. Upstream discharges represent patients in worse health than downstream destinations. Specifically, we group LTCH discharges to hospice or ACH as upstream and we group discharges to SNF/IRF, home (with or without home healthcare), and other as downstream.[10] Table II shows that most (about 85%) of LTCH discharges are downstream, and that patients initially discharged downstream have substantially lower post-discharge mortality, length of stay, and payments.

3. LTCH PAYMENTS, DISCHARGE PATTERNS, AND OUTCOMES

In this section, we present descriptive analysis on LTCHs' response to financial incentives. The analysis motivates several of the key choices for our model of LTCH discharges, which we present in Section 4.

---

[10]Table A.I shows with more granularity the discharge destinations within upstream and downstream. In the PPS period, 76% of patients discharged upstream are sent to ACH (versus hospice); of patients discharged downstream, about half are sent to SNF/IRF and another 44% are discharged to home or home health care.

3.1. *LTCH Payments*

We start by describing how LTCH payments vary with the patient's length of stay, an object we refer to as the LTCH budget set or payment schedule. Appendix B provides more details. Figure 1 summarizes the payment schedules in the pre-PPS and PPS periods.

Prior to October 1, 2002, LTCHs were paid their (estimated) daily cost, generating a linear relationship between the length of the hospital stay and payments. As described earlier, this "cost plus" reimbursement of LTCHs was seen as potentially encouraging inefficient entry into the LTCH market. Because of this and other concerns, the 1997 Balanced Budget Act (BBA) and 1999 Balanced Budget Refinement Act (BBRA) implemented a PPS for LTCHs. LTCH-PPS was phased in over a 5-year period starting on October 1, 2002 and was fully implemented by October 1, 2007. At a broad level, LTCH-PPS is designed to operate like the PPS for acute care hospitals (IP-PPS), under which hospitals are paid a lump sum that is based on the patient's diagnosis (diagnosis-related group, or DRG) and does not vary with the patient's length of stay.

Much like LTCHs were originally created to address a potential problem with the introduction of PPS for ACHs, the features of the LTCH-PPS payment schedule can similarly be thought of as attempting to address a potential problem arising from the introduction of PPS for LTCHs. In particular, in designing LTCH-PPS, officials were concerned that LTCHs might discharge patients after a small number of days but still receive large lump-sum payments intended for longer hospital stays. To address this concern, they created a short stay outlier (SSO) threshold. If stays were shorter than the SSO threshold, payments would be based on the pre-PPS cost-based reimbursement schedule and LTCHs would not receive a large lump sum. However, while reducing the incentive to cycle patients in and out of the LTCH, the SSO system creates potentially problematic incentives at the SSO threshold. At the day when payments switch from per day reimbursement to the lump-sum prospective payment amount, Medicare payments for keeping a patient an additional day "jump" by a large amount.

Figure 1 graphs the average payment schedules in the pre-PPS and PPS periods, pooling across LTCH facilities and DRGs. The y-axis shows cumulative Medicare payments, inflation-adjusted to 2012 dollars. The x-axis shows the length of the stay relative to the SSO threshold, which we normalize to be day 0. The SSO threshold is defined as five-sixths the geometric mean length of stay for that DRG in the previous year and therefore varies by DRG (and also, to a lesser extent, by year). The average threshold is at 22.6 days; the modal threshold (accounting for 22.7% of PPS stays) is 20 days; the range is 14 to 56 days, but 99% of the sample has a SSO threshold between 16 and 39 days. As a result, in this and subsequent figures, we present results relative to the SSO threshold so that we can pool analyses across DRGs.[11] Because the SSO threshold is undefined in the pre-PPS period, we assign pre-PPS stays the threshold for their DRG from the first year of the PPS period, 2007.

Under the pre-PPS system, average payments scale linearly with the length of stay at a rate of \$1,071 per day. Under the PPS system, average payments increase linearly by \$1,380 per day to the left of the SSO threshold, jump by \$13,625 at the SSO threshold, and remain constant thereafter. The jump in payments is large: it is equal to 55% of the

---

[11]We start the x-axis range at -15 days because nearly all SSO thresholds occur after 16 days. If we extended the x-axis range to -16, for example, there would be a change in the composition of DRGs between days -16 and -15 due to the entry of new DRGs into the sample. We end the x-axis range at $+45$ days because there are relatively few patients (2.1%) who are kept at the LTCH more than 45 days beyond the SSO threshold.

cumulative payment amount on the day prior to the threshold, or equivalent to about 10 days of payments at the pre-threshold daily rate.

This sharp jump in payments was presumably not the intention of the policymakers who designed the LTCH-PPS, but it arises naturally from the interaction of two sensible policies. As is standard in fixed price contracts, the LTCH-PPS payments were likely set to approximate average costs per stay. As noted, payments on a cost-plus basis up to the SSO threshold were introduced to avoid paying LTCHs large lump-sum amounts for relatively short stays. The (approximate) average cost for stays longer than the threshold naturally introduces a jump in payments in the transition from a per day payment regime to a per-stay regime, creating potentially problematic incentives. Particularly concerning is where the threshold was set: we estimate that under the pre-PPS payment scheme, 44% of stays would have been below the subsequent short stay outlier threshold, which is a large fraction for a policy that is at least ostensibly designed to target "outlier" events.

In Section 5, we explore the impact of alternative, counterfactual payment schedules. To eliminate the jump in payments, our counterfactuals alter the payment prior to the SSO threshold (so that it does not approximate per day costs), alter the fixed PPS amount (so that it does not approximate average costs), or alter both segments of the payment schedule.

3.2. *Discharge Patterns*

To motivate our model of LTCH behavior, we present three main descriptive results on discharge patterns from the LTCH around the threshold; some have been previously documented, while others are, to the best of our knowledge, new.

First, there is a large spike in discharges at precisely the day of the jump in payments, indicating a strong response to financial incentives. This finding has been noted by a number of previous studies (Weaver et al. (2015), Kim et al. (2015), MedPAC (2016)). Specifically, the top left panel of Figure 3 shows the aggregate pattern of discharges by length of stay in the pre-PPS and PPS periods. A discharge occurs when the patient is transferred to another facility, sent home, or dies at the LTCH. The y-axis shows discharges as a share of the total number of stays at the LTCH. The x-axis plots the length of stay relative to the DRG-specific SSO threshold, defined in the same manner as in Figure 1. In the PPS period, there is a sharp increase in discharges at the SSO threshold, with the share of discharges increasing from about 2% to 9% per day. Discharge rates remain elevated over the subsequent 7–10 days before reverting to baseline. In the pre-PPS period, there is no evidence of any bunching at the SSO threshold; differences in the pre-threshold discharge rate may reflect changes in patient health or other secular trends between the periods. Importantly, there is not a sharp decrease in discharges immediately before the SSO threshold under PPS; as we discuss in more detail below, this motivates our decision to write down a dynamic model of LTCH behavior (where LTCHs respond well in advance of the jump in payments) rather than a myopic model (where LTCHs only respond immediately before the jump).

Second, the marginal patients discharged at the threshold are in relatively better health: they are disproportionately discharged downstream and have lower post-discharge mortality rates than patients discharged at other times; Eliason, Grieco, McDevitt, and Roberts (forthcoming) have also documented that marginal patients are disproportionately discharged downstream. The rest of the panels of Figure 3 decompose the discharge pattern by the location of discharge: downstream, upstream, and death. They show increases at the threshold in discharges both upstream and downstream, but the proportional increase is substantially larger on the downstream margin. Moreover, because the
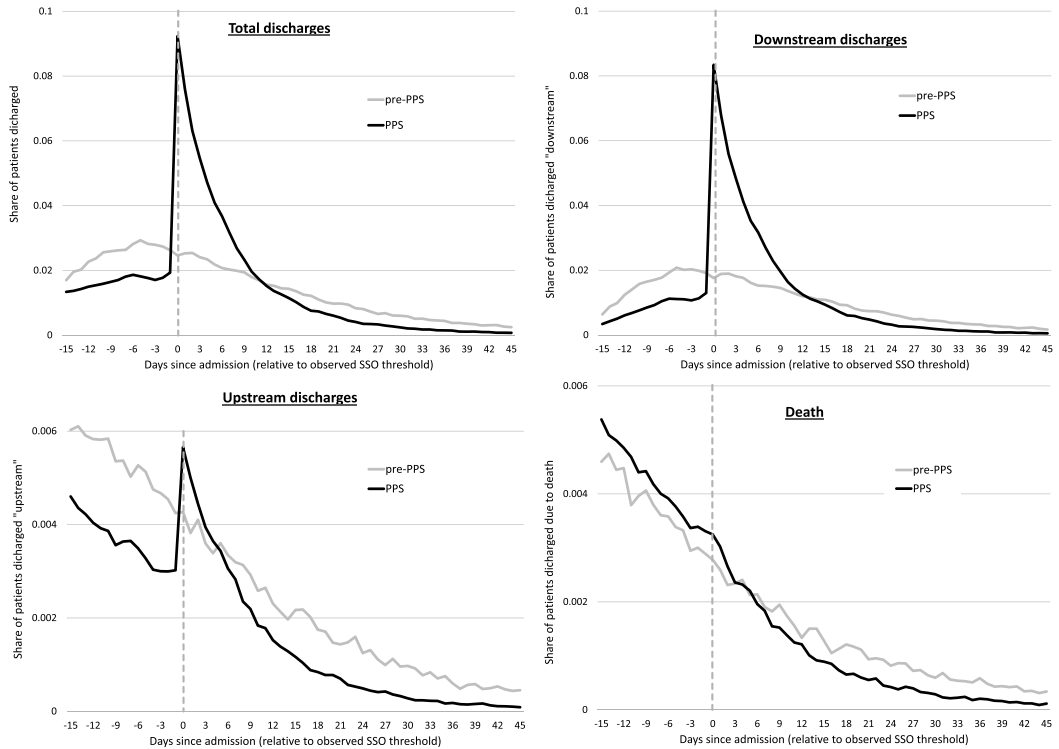
FIGURE 3.—Discharge patterns by length of stay. Figure presents the distribution of the time of discharge relative to the SSO threshold. That is, each line shows the number of discharges on a given (relative) day divided by the total number of LTCH admissions. Sample pools admissions that are associated with different SSO thresholds, and x-axis is normalized by counting days relative to the threshold. The top left panel presents the distribution for all discharges, the top right and bottom left panel present the same information separately for downstream (SNF, IRF, LTCH, home health, home, or other) and upstream (ACH or hospice) discharges, and the bottom right panel presents discharges due to death occurring within the LTCH.

pre-threshold discharge rate is much higher downstream, the sharp change in the aggregate discharge rate at the threshold (top left panel) is almost entirely driven by downstream discharges. We defer our discussion of the right bottom panel on mortality to the subsection below.[12]

Third, among patients discharged downstream, the marginal patients discharged at the threshold are relatively sicker, with higher post-discharge payments than pre-threshold dischargees. Figure 4 illustrates this, plotting Medicare payments for the episode of care that occurs after the LTCH discharge, by length of stay at the LTCH. We show these post-discharge payments separately for patients discharged upstream and downstream and view them as a proxy for the patient's health at the time of discharge. For patients discharged downstream, there is a sharp increase in post-discharge payments at the

___

[12]In addition, Figure B.2 plots the 30-day post-discharge mortality rate, defined as death within 30 days of a (live) discharge, by length of stay. The graph shows a sharp drop in post-discharge mortality at the SSO threshold, again suggesting that the patients who are discharged at the threshold are healthier than the patients who are discharged immediately beforehand. Of course, the decline in mortality not only reflects changes in the composition of patients discharged at the threshold, but could in principle reflect a treatment effect of discharge on health. We address this in the next section.
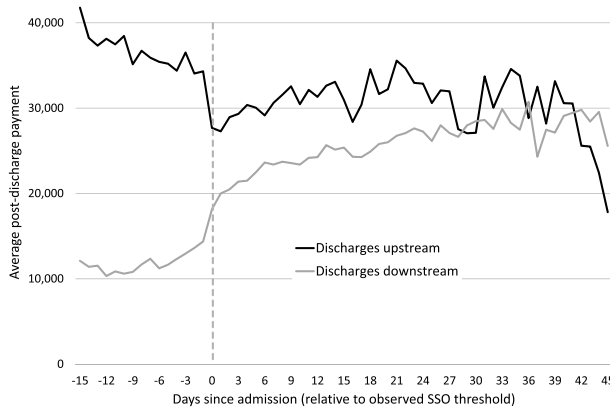
FIGURE 4.—Post-discharge payments. Figure presents the average post-discharge payments for patients discharged alive, by discharge day and discharge destination (upstream vs. downstream, as defined in Figure 3). We define a post-discharge episode as ongoing until there is a break of at least two days that does not involve a facility stay; see text for more details.

threshold, with average post-discharge payments increasing from approximately $10,000 to $20,000. There is a small change in the opposite direction for patients initially discharged upstream. For longer lengths of stay, the figure becomes noisy due to the small number of discharges.

Figure 4 suggests a simple model of LTCH behavior, which motivates the model we present in Section 4. Prior to the threshold, retaining patients is profitable, and only the healthiest patients are discharged to SNF/IRF or to their home and only the sickest patients are discharged to an ACH or a hospice. After the threshold, on the downstream margin, LTCHs work "down the distribution" and discharge less healthy patients, increasing post-discharge payments on average. Similarly, on the upstream margin, LTCHs work "up the distribution," discharging patients who are in better health, and decreasing post-discharge payments on average. The marginal patient discharged downstream at the threshold is therefore sicker than the average patient discharged downstream prior to the threshold, while the marginal patient discharged upstream is slightly healthier than the average patient discharged upstream in prior days. As we discuss more in Section 4, Figure 4 also suggests the need for a dynamic model—in which health evolves over time and LTCHs make daily discharge decisions based on the patient's contemporaneous health—rather than a static model in which the hospital commits to a pre-specified length of stay at the time of LTCH admission.

3.3. *(Lack of) Mortality Effects*

A natural question raised by the discharge patterns is whether the distortions in the timing of discharges have an impact on patient health and in particular on mortality. Since the 90-day mortality rate of LTCH patients is approximately 30% (Table I), if these distortions are harmful to health, it seems plausible that we might be able to pick up an effect with our data.

Empirical identification of mortality effects from the distortion in patient location at the threshold is challenging, however. Health evolves according to a stochastic process, with sicker patients having a higher probability of death. Distortions to the location of care might impact the level of someone's health, generating an on-impact effect on the

probability of death analogous to the on-impact effect on discharges we detected. However, distortions to the location might also affect the stochastic process for health, which would be associated with a longer-run change in mortality rate, but might not have an immediate mortality effect. We therefore attempt to examine not only whether there is an immediate impact on mortality at the threshold, but whether we can detect any longer-run changes.

The bottom right panel of Figure 3—which plots daily mortality rates within the LTCH by patient length of stay—shows that mortality rates among patients in the LTCH are declining over the course of the LTCH stay with little difference around the SSO threshold. However, interpretation is complicated by selection. As the sickest patients die, the remaining patient pool is healthier, which presumably contributes to the downward mortality trend. And since LTCHs are differentially discharging healthier patients at the SSO threshold, the composition of patients who remain at the LTCH after the threshold is different, making it tricky to disentangle any potential treatment effects on mortality at the threshold from changes in the selection of patients remaining at the LTCH at the threshold.

To circumvent this issue, we take advantage of the fact that, as we noted in Section 2.3, our data also allow us to track mortality outcomes for patients even after their LTCH discharge. Figure 5 therefore analyzes mortality patterns in the days post LTCH-admission, *unconditional on the patient's current location*. Conceptually, our mortality analysis follows the logic of a reduced form regression, where the mortality hazard is the outcome, discharge patterns are the endogenous variable, and the SSO threshold is the instrument. In particular, since we know there is a sharp jump in discharge patterns at the threshold (analogous to a large first stage), if there is a change in the level or slope of the mortality hazard at the threshold (i.e., nonzero reduced form), we can infer that the distortion in discharge location has an impact on mortality.

The top panel of Figure 5—which shows daily mortality rates by days since LTCH admission—is thus similar to the bottom right panel of Figure 3, but uses the full set of LTCH patients (unconditional on their location) rather than only those who have yet to be discharged. As before, "natural selection" leads mortality rates to decline over time, but we now can interpret more cleanly the mortality pattern around the SSO threshold. The plot shows no obvious evidence of a change in the level of mortality hazard in the vicinity of the threshold during the PPS period. In Appendix C, we examine this mortality pattern more formally using a regression discontinuity design and similarly fail to reject the null of a smooth mortality hazard around the SSO threshold.[13] These findings are consistent with no mortality effect but do not allow us to rule out a gradual effect that would not appear sharply in the data.

If distortions in the location of care affected the stochastic process for health, we might not observe an immediate effect, but would see a change in mortality over a longer time horizon. The bottom panel of Figure 5 attempts to look for a more gradual effect by plotting a 30-day mortality rate (again unconditional on the patient's current location), by days since LTCH admission, where the 30-day mortality hazard measures the share of patients who are alive on a given day but die during the subsequent 30 days. The plot once again shows no effect around the threshold, suggesting that there are no gradual effects of the distortion in location on mortality. In Appendix C, we present a regression discontinuity analysis that more formally confirms this result.

---

[13]Our baseline estimate (shown in column (1) of Table F.I) allows us to rule out with 95% confidence a daily mortality increase of more than 0.05 percentage points and a daily mortality decline of more than 0.04 percentage points (off of a base of 0.6 percent).
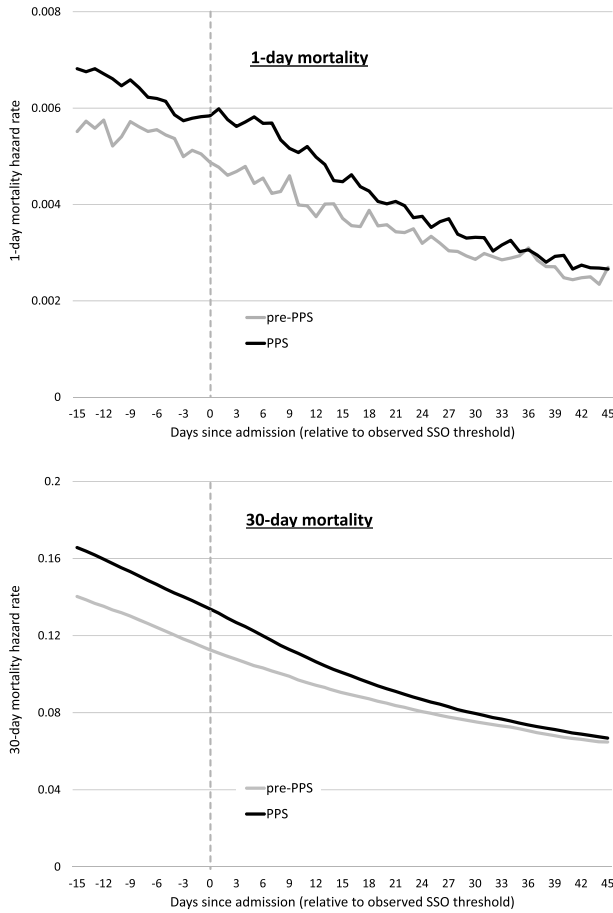
FIGURE 5.—Mortality patterns by days since LTCH admission. Figure presents post-LTCH-admission mortality-hazard rates by day. Mortality includes any mortality, whether it occurs within the LTCH or after discharge. Each panel presents hazard rates for different subsequent horizons: same day (top) and 30-day forward (bottom).

Figure 5 thus suggests little evidence of a quantitatively large effect on mortality that is created by the sharp changes in discharge behavior at the SSO threshold. To supplement this analysis, we also test for mortality effects using variation over time in the location of the SSO threshold within DRGs; Eliason et al. (forthcoming) similarly exploited this variation to examine how changes within DRG in the SSO threshold affect discharge behavior. Recall that the SSO threshold is determined as five-sixths of the geometric mean length of stay in the prior year. During our 2007–2012 sample period, about 80% of (stay weighted) DRGs experience at least one change in the SSO threshold, typically a shift of a single day. We use these changes in the SSO threshold—which occur for different DRGs in different years, roughly evenly distributed across the sample years—to examine the mortality effects of length of stay in a difference-in-differences framework. In particular, we collapse our data to the DRG-year level and estimate regressions of the form

$$y_{dt} = \alpha_s \text{SSO}_{dt} + \tau_t + \kappa_d + \varepsilon_{dt}, \tag{1}$$
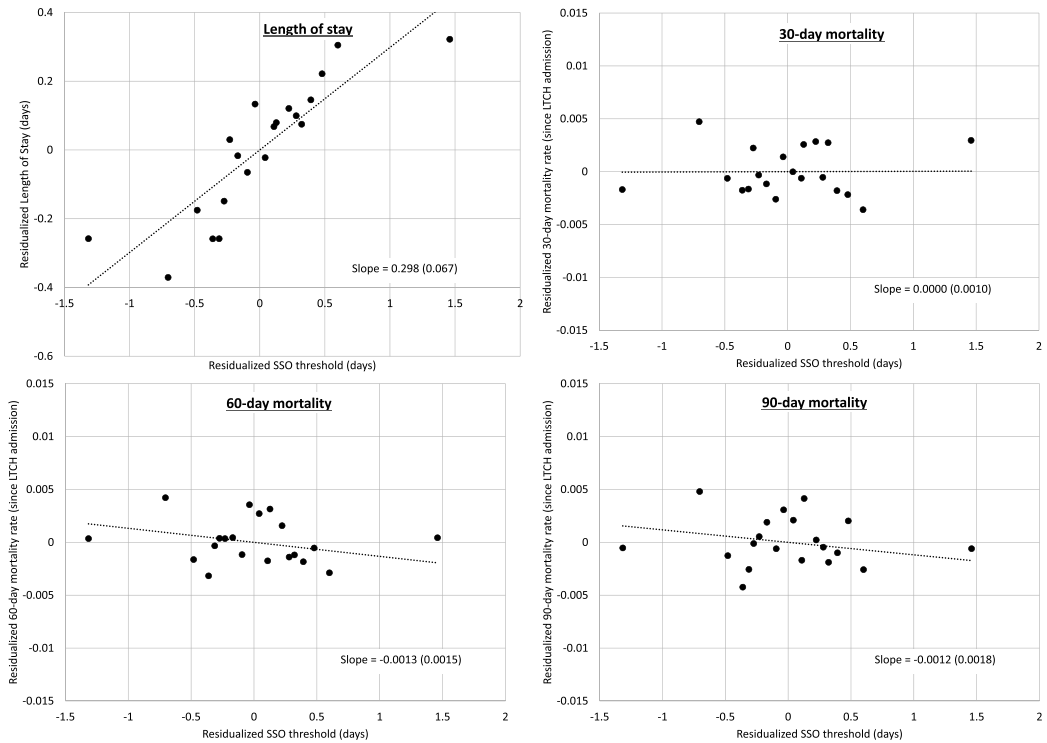
FIGURE 6.—The effect of changes in the SSO threshold on mortality. Figure shows residualized binned scatter plots (as in, for example, Chetty, Friedman, and Rockoff (2014)). The vertical axis shows the outcome variable net of year and DRG fixed effects, and the horizontal axis shows the SSO threshold net of year and DRG fixed effects. The panels show scatter plots where we aggregate the data by ventiles of the horizontal axis variable (SSO threshold). In the top left panel, the outcome variable is length of stay. In the remaining panels, the outcome variable is the 30-, 60-, and 90-day mortality rates (unconditional on location of care). The plots also display the best fit line from estimating equation (1), along with the estimated slope coefficient and heteroscedasticity-robust standard errors in parentheses.

where y_{dt} is the average outcome for DRG d in year t, SSO_{dt} is the SSO threshold associated with DRG d in year t, τ_t and κ_d are year and DRG fixed effects, respectively, and ε_{dt} is the error term. We estimate a first-stage regression that relates changes in the SSO threshold within DRGs to changes in the average length of stay within DRGs. We also estimate reduced form regressions that relate changes in the SSO threshold to changes in mortality within 30, 60, and 90 days of LTCH admission, and IV regressions that relate length of stay to mortality, instrumenting for length of stay with the SSO threshold.

Figure 6 displays the results. In each panel, the horizontal axis shows the SSO threshold net of year and DRG fixed effects. The vertical axis shows various outcome variables, also net of year and DRG fixed effects. The graphs show scatter plots of the data, aggregated by ventiles of the horizontal axis variable; they also show the slope coefficient α_s estimated by equation (1).

The top left panel shows that there is a strong first-stage relationship, with a one-day increase in the SSO threshold raising the average length of stay by 0.3 days (standard er-

TABLE III

UsING SSO THRESHOLD VARIATION TO ESTIMATE MORTALITY EFFECTS[a]

	Mean	FS	RF Est.	IV (pp)	IV (pct.)	95% CI (pct.)
30-day mortality	0.158	0.298	0.0000	0.0001	0.0007	(−0.035, 0.036)
		(0.067)	(0.001)	(0.003)	(0.018)	
60-day mortality	0.253	0.298	−0.0013	−0.0045	−0.0176	(−0.047, 0.012)
		(0.067)	(0.001)	(0.004)	(0.015)	
90-day mortality	0.306	0.298	−0.0012	−0.0040	−0.0129	(−0.043, 0.018)
		(0.067)	(0.002)	(0.005)	(0.016)	

[a]Table shows 2SLS estimates of the effect of length of stay on mortality, instrumenting for length of stay with over-time changes in the SSO threshold. The first column shows the mean mortality rate over 30-, 60-, and 90-day time horizons (unconditional on location of care). The second column shows the first-stage effect of the SSO threshold on length of stay (see equation (1)); the third column shows the reduced form effect of the SSO threshold on mortality from a linear regression with year and DRG fixed effects (see equation (1)). The fourth column shows the 2SLS estimate, where the first stage is a regression of length of stay on the SSO threshold and year and DRG fixed effects (shown in column 2), and the second stage is a regression of mortality on length of stay and year and DRG fixed effects. The final two columns show the 2SLS estimate (and 95% confidence interval) as a percentage of the mean mortality rate. Standard errors and confidence intervals are all heteroscedasticity robust.

ror $= 0.07$).[14] The three other panels of Figure 6 show the relationship between mortality and the SSO threshold. Table III shows the corresponding IV estimates of the impact of length of stay on mortality, where we use the change in the SSO threshold as an instrument for length of stay. The estimated effects of length of stay on mortality are negative but statistically insignificant, with the point estimates ranging from a 0.45 percentage point decline in 60-day mortality to a 0.01 percentage point change in 30-day mortality. Because baseline mortality rates are high, we can reject fairly small proportional effects. For example, relative to a 90-day mortality rate of 30.6%, our estimates allow us to rule out mortality declines greater than 4.3% or mortality increases greater than 1.8% with 95% confidence interval.

Overall, while these results provide no "smoking gun" evidence of patient harm, they do not allow us to comprehensively rule out negative health effects. And even if an average mortality impact can be ruled out, it may mask important heterogeneity, so we would still not be able to rule out mortality effects for every type of patient or at each and every hospital. In addition, we cannot rule out adverse health effects that would not manifest in higher mortality rates. While it may be tempting to analyze the effect of LTCH length of stay on other health-related outcomes, non-mortality health outcomes are tracked and measured differentially based on location of care, and are thus likely to be mechanically related to the length of the LTCH stay.[15]

Still, we view the mortality analysis as suggestive that the marginal patients affected by the PPS payment schedule are likely able to receive similar care—as measured by mortality—whether they are located in an LTCH or in an alternative setting, which empirically is usually a SNF. Two other pieces of evidence are consistent with this interpretation. First, we showed earlier that the patients who are most affected by the SSO threshold

[14]LTCHs might respond to the change in SSO threshold by admitting more or fewer patients, which would complicate the analysis, especially if there was a change in the composition of admitted patients. However, we find no evidence that the number or mix of admissions vary in response to changes in the SSO threshold (not reported).

[15]For instance, hospital-acquired infections are measured at facilities, but not at home. All else equal, a patient who has a longer hospital stay is more likely to acquire a hospital-acquired infection even if their health would have been the same if their hospital stay were shorter.

are disproportionately healthy, and thus potentially less sensitive to variation in the location of care. Second, using a different empirical design that studies the impact of LTCH entry into regional healthcare markets, we find that LTCH entry leads to substantial substitution from SNFs to LTCHs, but no mortality impact, again suggesting that marginal patients can receive appropriate care at both types of facilities (Einav, Finkelstein, and Mahoney (2018)).

4. QUANTIFYING THE IMPORTANCE OF FINANCIAL INCENTIVES

The results in the last section provide descriptive evidence of the response of LTCHs to the sharp financial incentives associated with the SSO threshold. These patterns motivate the discharge model that we now specify in order to assess how these patterns would change in response to counterfactual financial contracts that do not exhibit such sharp incentives.

4.1. *The Importance of Dynamics: Intuition and Motivation*

It is natural to think of a hospital discharge decision as a dynamic discrete choice problem. Every day, the LTCH assesses the patient's health and decides whether to retain and treat the patient at the LTCH or discharge the patient to another location, where the patient might receive different medical treatment. To develop intuition, we assume for now that the hospital cares only about maximizing profits. We will relax this in our baseline model below.

The LTCH decision is asymmetric and resembles an optimal stopping problem. If the patient is discharged, the hospital has no subsequent decision to make, as it loses control over the patient. However, if the patient is retained for an extra day, the hospital obtains the flow costs and benefits associated with treating the patient for an extra day, as well as the costs and benefits associated with the option value of making the optimal decision the next day.

This dynamic option value is particularly important in light of the sharp jump in payments at the SSO threshold. To gain intuition, consider an overly simplified setting in which the LTCH's cost of treating each patient is c per day, and its revenues are given by $p \approx c$ for each day prior to the SSO threshold, $P \approx 10p \gg p$ at the day of the threshold, and zero thereafter. Assume also that, with some relatively low probability, in any given day there is an exogenous probability the LTCH is forced to discharge the patient (e.g., due to mortality or a relapse) or forced to retain her (e.g., because a family member is not available to transport her home).

From the LTCH's perspective, there are three qualitatively different periods. First, consider the period after the SSO threshold: the patient generates cost and no revenues, so the hospital has no incentive to retain the patient unless it has to. Indeed, as we saw in Figure 3, hospitals discharge their patients fairly rapidly after the SSO threshold has passed. Second, on the day at which the SSO threshold hits, the hospital has a very strong (static) incentive to retain the patient, and thus would hold on to the patient unless forced to discharge her. Finally, prior to the SSO threshold, the hospital does not have strong *static* incentives to retain or discharge the patient (recall, we assume $p \approx c$ in this example), yet it faces *dynamic* incentives to keep the patient until the SSO threshold in order to obtain the large payment P.

How strong are these dynamic incentives to retain patients prior to the SSO threshold? Let $V_{\text{SSO}} \gg 0$ denote the financial value associated with LTCH patients who make

it to the SSO threshold; in other words, V_{SSO} is the financial reward $P \approx 10p$ minus the (significantly lower) expected cost associated with the 8.3 days (on average) that patients remain in the hospital after the SSO threshold. Prior to the SSO threshold, the dynamic incentives to retain the patient are $V_{SSO} \cdot \Pr(\text{LOS} \geq \text{SSO} | \text{LOS} \geq t)$, the financial benefit from reaching the SSO threshold multiplied by the probability the patient reaches the threshold. Obviously, the probability is increasing with t, so the incentives to retain a patient are higher as the SSO threshold gets closer. However, with negligible discounting due to time, and a relatively low probability of exogenously losing the patient (approximately 2% in the baseline sample), the probability term is fairly large, and the dynamic incentives prior to the SSO threshold are not substantially lower than the static incentives at the SSO threshold.

It is useful to contrast this simple framework with a myopic model of LTCH behavior in which dynamic considerations are ignored. In a myopic model, LTCHs would experience a sharp increase in the financial incentives to retain a patient between the SSO day and the day that immediately precedes it, leading to a sharp decline in the discharge rate. As can be seen in Figure 3, the data provide no evidence of such a pattern, with daily discharge rates at the SSO day and the days that precede it being essentially the same, which is consistent with the (overly) simplified dynamic incentives sketched above. Below, we will show more formally that a myopic model does not fit well the discharge patterns we observe.

We can also contrast our dynamic model with a completely static model in which the hospital commits to a specific length of stay at the time of LTCH admission. This type of static model is not a good descriptive model of our environment; LTCHs maintain flexibility and make discharge decisions on an ongoing daily basis. The key drawback of committing to a specified future discharge date is that the hospital "ties its hands" and is not able to respond to future information. Such information is not consequential in the setting described above but becomes important once we enrich the simple framework and allow the health of LTCH patients to evolve stochastically, as we do in our baseline model below. As we showed in Figure 4, there is a clear relationship between the average health of discharged patients (as proxied by their post-discharge costs) and the timing and destination of discharge. Once health is heterogeneous, and evolves stochastically over time, the health status at the time of admissions is not very predictive of health status at the time of discharge, preventing a static model from matching this relationship. Again, we will show this more formally below.

4.2. *A Model of Dynamic Discrete Discharge Choice*

Our baseline model, which we present here, builds on the intuition above, relaxing some of the assumptions and allowing for heterogeneous, stochastically evolving patient health.

Consider a patient i who is admitted at day $t = 0$ to LTCH l. We index patient i's health at the time of admission by $h_{i,0}$, and assume that $h_{i,t}$ (conditional on patient i staying at LTCH l) evolves stochastically from day to day. Specifically, we assume that $h_{i,t}$ follows a monotone Markov process, such that $h_{i,t} \sim F(\cdot|h_{i,t-1})$ with $F(\cdot|h)$ stochastically increasing in h.[16] We use higher values of h to indicate better health, so the daily mortality hazard $m(h)$ is strictly decreasing in h.

---

[16]In sensitivity analysis reported in Appendix F, we examine the robustness of our findings by allowing the health process to vary with time since admission.

Hospital l's flow (daily) monetary profits from patient i (whose health is given by h) during the tth day since admission is given by

$$\pi(h, t) = p(t) - c, \qquad (2)$$

where $p(t)$ is the hospital's revenue, which depends on CMS's reimbursement schedule for patient i, and c is the hospital's daily cost of treating each patient. An important assumption in this specification is that daily costs c are constant and do not vary with the patient's health.[17]

Our focus is on the hospital's discharge decision. Following our descriptive analysis, we consider two alternative destinations for patient i, downstream and upstream, so that every day the hospital makes a choice between keeping the patient overnight in the LTCH (l), discharging her downstream (d), or discharging her upstream (u). The hospital's non-monetary payoffs every day are given by

$$u^j(h) = v^j(h) + \sigma_\epsilon \epsilon_{ijt} \quad \text{for } j = l, d, u, \qquad (3)$$

where $j = l, d, u$ is the location in which the patient stays that day, $v^j(h)$ captures hospital l's value from having the patient staying at location j (which can be viewed as the part of the patient's utility that is internalized by the hospital), and ϵ_{ijt} is an error term, which is distributed i.i.d. type I extreme value and scaled by the parameter σ_ϵ. The error term presumably captures idiosyncratic considerations associated with the patient and/or the hospital. Moreover, because hospital l loses control over the patient upon discharge, it will be convenient to denote by $V^j(h)$ the present value to hospital l of the patient's utility from being discharged to destination $j = d, u$.

This setting lends itself to a simple dynamic programming problem, which can be represented by the following Bellman equation:

$$V(h, t) = \pi(h, t) + \delta\big(1 - m(h)\big) E \left(\max \left\{ \begin{array}{l} u^l(h) + \int V\big(h', t+1\big)\, dF\big(h'|h\big), \\[2mm] u^d(h) + \int V^d\big(h'\big)\, dG^d\big(h'|h\big), \\[2mm] u^u(h) + \int V^u\big(h'\big)\, dG^u\big(h'|h\big) \end{array} \right\} \right), \qquad (4)$$

where δ is the LTCH's (daily) discount factor. The two state variables are the health of the patient (h) and the number of days since LTCH admission (t). While we did not find a mortality effect in our descriptive analysis, by allowing the health process outside the LTCH to evolve according to $G^d(\cdot|h)$ and $G^u(\cdot|h)$, instead of $F(\cdot|h)$ within the LTCH, our model allows patient health to evolve differentially across alternative locations of care.

It is convenient to benchmark $v^j(h)$ against the LTCH value from having the patient stay at the LTCH. That is, we normalize $v^l(h) = 0$ for all h, and normalize $V^j(h)$ accordingly. Applying these adjustments and using the well-known expression for the logit's

---

[17]While one may be worried that sicker patients are more costly, we view the homogeneous cost assumption as a reasonable approximation for several reasons. First, large components of LTCH cost structure are unlikely to vary much with the health status of the patient occupying the bed. These health-invariant costs include the equipment and personnel associated with the bed and the shadow cost of capacity constraints. Second, in the implementation below, we will examine empirical patterns across DRGs; any health-related variation in costs across DRGs will therefore be captured as long as the DRG-specific payment rates reflect this variation (as they are designed to). We have also verified that the quantitative implications of allowing cost to vary with health are relatively minor in the context of our counterfactual exercises.

inclusive value, we can write the problem as

$$V(h, t) = p(t) - c + \delta(1 - m(h))\sigma_\epsilon \ln \left\{ \begin{array}{l} \exp\left(\int V(h', t+1) \, dF(h'|h) \right) \\ + \exp(V^d(h)) + \exp(V^u(h)) \end{array} \right\}. \qquad (5)$$

Finally, we note that the state variable t only affects the problem through the hospital revenue function $p(t)$, and $p(t) = 0$ for all $t > $ SSO, so the problem becomes stationary after the SSO threshold, and the solution is simply a fixed point of

$$V^{t>\text{SSO}}(h) = -c + \delta(1 - m(h))\sigma_\epsilon \ln \left\{ \begin{array}{l} \exp\left(\int V^{t>\text{SSO}}(h') \, dF(h'|h) \right) \\ + \exp(V^d(h)) + \exp(V^u(h)) \end{array} \right\}. \qquad (6)$$

We can therefore solve the dynamic problem by first solving for the fixed point associated with the post-SSO stationary part of the problem given by equation (6), and then iterating backwards until $t = 0$ using equation (5).

4.3. *Parameterization and Estimation*

Parameterization

We make several additional assumptions before we take the model to the data. The first is about the health process. Given that mortality is monotone in h, it is convenient to normalize the health index by mortality risk. We do so by assuming that h is defined by its associated mortality hazard using the following relationship:

$$m(h) = 1 - \Phi(h), \qquad (7)$$

where $\Phi(\cdot)$ is the standard normal CDF. We note that h is an index so the above is simply a normalization, which endows h with a cardinal measure. Equipped with this normalization, we then make parametric assumptions about the initial health distribution (as of LTCH admission, $t = 0$), and about how the health process evolves over time within the LTCH. Specifically, we assume that $h_{i,0}$ is drawn from $N(\mu_0, \sigma_0^2)$ and that $F(\cdot|h_{i,t-1})$ follows an AR(1) process:

$$h_{i,t} = \mu + \rho h_{i,t-1} + \varepsilon_{i,t}, \quad \text{where } \varepsilon_{i,t} \sim N(0, \sigma^2). \qquad (8)$$

In our baseline specification, we allow the health process inside the LTCH to be different in the pre-PPS and PPS periods, in order to accommodate potential differences in the LTCH patient mix; these may result from the growth of the LTCH sector, time trends in medical technology and practice, or directly from the change in financial incentives. Note that we do not need to model the health process outside the LTCH since any effect of the post-discharge location on health would affect the discharge decisions through its (unidentified) effect on the continuation values, $V^j(h)$, which we do not model explicitly.

Second, we approximate $V^u(h)$ and $V^d(h)$ using a linear function in h, so that

$$V^j(h) = v_{0j} + v_{1j}h \quad \text{for } j = d, u. \qquad (9)$$

Recall from Section 3 that healthier patients (higher h), who are associated with lower mortality, are discharged to d, while sicker patients (lower h), associated with higher

mortality, are discharged to u. It is therefore natural to expect $v_{1d} > 0$ and $v_{1u} < 0$. That is, all else equal, discharge destination d becomes more attractive as health gets better (h is higher) and destination u becomes more attractive as patients' health worsens (h is lower). As explained below, one of the intercept terms v_{0d} and v_{0u} needs to be normalized, so we set $v_{0u} = 0$.

Third, as we discussed in Section 2, LTCHs are part of an interlocking post-acute care system, with changes in LTCH incentives potentially affecting Medicare spending throughout the patients' entire episode of care. In particular, Figure 4 shows sharp changes in both upstream and downstream post-discharge payments at the SSO threshold, indicating a relationship between patients' health at discharge and total Medicare spending. To account for such spillovers in our counterfactuals, we model the relationship between health at discharge and post-discharge payments as

$$P^j(h) = \exp(\zeta_{0j} + \zeta_{1j}h) \quad \text{for } j = d, u, \tag{10}$$

where $P^j(h)$ represents the post-discharge payments for a patient initially discharged to location $j = d, u$ with health status h at the time of discharge. We allow this relationship to vary by whether the patient is initially discharged upstream or downstream and use an exponential specification so that predicted post-discharge payments are strictly positive.

Fourth, as is typical in these types of models, we set (rather than estimate) the daily discount factor to $\delta = 0.95^{1/365}$.[18] Finally, to account for changes in cost over time, we deflate the costs parameter c in the pre-PPS period by the ratio of reported administrative costs across periods, which we estimate to be 0.75. Thus, overall we are left with 19 parameters to estimate: five parameters ($\mu_0, \sigma_0, \mu, \sigma, \rho$) that are associated with the health distribution and the way it evolves over time in the pre-PPS period, five corresponding parameters in the PPS period, the cost parameter (c), four parameters ($v_{1u}, v_{0d}, v_{1d}, \sigma_\epsilon$) associated with the relative value of patients at facilities u and d, and four parameters associated with post-discharge payments in the PPS period ($\zeta_{0d}, \zeta_{1d}, \zeta_{0u}, \zeta_{1u}$).

Estimation

An important decision is how to treat heterogeneity across patients, observable health conditions, and LTCH hospitals. In our baseline specification, we abstract from such heterogeneity and instead model the "average" discharge decision as it pertains to the "average" LTCH patient and the "average" payment schedule. That is, we pool all payment schedules observed in the data, separately for the pre-PPS and PPS periods, measure each day in the schedule relative to the DRG-specific SSO threshold in the PPS period (which is normalized to zero at the SSO threshold), and construct the average payment schedule for each day, as shown in Figure 1. We analogously pool the discharge patterns, separately for the pre-PPS and PPS periods, in a 61-day window around the SSO threshold (from day -15 to day 45 as shown in Figure 3). We then estimate our model in an attempt to match these average patterns.

An advantage of this approach of focusing on the average patterns, rather than the heterogeneous patterns, is that it only requires us to solve the dynamic problem once (for each period, pre-PPS and PPS), which is computationally attractive. Given that our primary focus is on the aggregate effect of financial incentives across the entire LTCH

---

[18]While having a discount factor so close to 1 might generally create convergence issues, in our Bellman equation the probability of survival enters jointly with the discount factor so that the "effective discount factor" $\delta(1 - m(h))$ is not too close to 1.

sector, abstracting from the heterogeneity across patients and hospitals is less likely to be consequential. Heterogeneity in the response is also the focus of the related exercise reported by Eliason et al. (forthcoming).

We estimate the model using simulated method of moments, to match the daily mortality and discharge patterns presented in Figure 3, as well as post-discharge payment moments that are based on Figure 4. Specifically, we use two sets of moments. First, we use 183 moments for the pre-PPS payment schedule, reflecting the daily discharge and mortality risks within the 61-day window around the SSO threshold. One set of moments is associated with discharge rates to d, another with discharge rates to u, and a third with mortality rates. We then construct another set of 183 corresponding moments for the PPS period. Because much of the identification is driven by the sharp change to discharge incentives at the SSO threshold, we assign greater weights to moments that are closer to day zero (the SSO threshold) by making the weights decrease by a constant amount (1/61) per day away from the threshold. The second set of moments uses the data on post-discharge payments to allow us to capture spillover effects in our counterfactuals. Specifically, we average post-discharge payments for each discharge destination (d or u), separately for discharges before and after the SSO threshold. We then match them to the model prediction regarding the health status distribution of discharged patients, thus allowing us to link health and post-discharge payments.[19]

Generating the model predictions requires us to solve the dynamic problem described in the previous section for each set of parameters. To ease the computation, we approximate the health process $F(\cdot|h_{i,t-1})$ with a discrete health space that evolves according to a Markov transition matrix (Tauchen (1986)).[20] This eases the solution of the dynamic problem, and at the same time allows us to read the discharge probabilities directly off the solution, without any need to integrate (presumably by simulation) over potentially intractable integration regions.

4.4. *Identification*

The model specified above is closely related to a large family of dynamic discrete choice models that have been studied extensively in the literature. Pakes (1986) is an early application of such models, and Arcidiacono and Ellickson (2011) provides a recent review. In order to understand the challenge of nonparametric identification of the model, it is probably easiest to discuss the relationship between our model and the econometric framework studied in Hu and Shum (2012) and in Hu, Shum, Tan, and Xiao (2015). Our model presented above would map very closely to this framework if we observed a variable that corresponds one-to-one to latent health $h_{i,t}$.

In our setting, the mortality rate is observed and plays a conceptually similar role to such a variable, but because mortality is a binary variable it unfortunately does not have a one-to-one relationship with the continuous health index, and therefore our model cannot be nonparametrically identified. Instead, we rely on the parametric assumptions associated with the initial health distribution and the AR(1) process for the evolution of health in order to identify the rest of the model. Conditional on identifying the parameters that

---

[19]Because there is no variation in payments in the pre-PPS period, we do not have the variation in dischargees' health that we need to identify the pre-PPS post-discharge payments model. Since we focus on the PPS period in our counterfactuals, and therefore do not need these estimates, we do not estimate pre-PPS post-discharge payments.

[20]In particular, we approximate the health distribution with a grid of 250 evenly spaced values that span a range of $+/-$ three standard deviations around the mean of the steady-state health process.

govern initial health and the way it evolves over time, identification of the other parameters is reasonably straightforward.

In Appendix E, we provide many more details about the (parametric) identification of the model. To provide intuition, it is useful to build up from the case when patient health is held fixed at a given value of h and there is only a downstream margin of discharge. In this case, discharges are characterized by three parameters: the scale parameter on the logit errors σ_ϵ, the cost parameter c, and the value of discharging the patient downstream v_d.

We can use the jump in payments at the SSO threshold to separately identify the scale parameter σ_ϵ from the cost c and value of discharge v_d parameters. Since we normalize the coefficients on net profits ($p(t) - c$) to 1, the scale parameter σ_ϵ can be thought of as the inverse "profits sensitivity" of the LTCH. Prior to the SSO threshold, increasing σ_ϵ reduces the option value of retaining the patient until the jump in payments because the LTCH places less weight on the financial value of the jump in payments, thus raising the value of discharging the patient. After the SSO threshold, profits are negative, so increasing σ_ϵ makes retaining the patient not as bad, thus lowering the value of discharging the patient. As a result, σ_ϵ can be thought of as modulating the change in incentives at the SSO threshold, with a higher value for σ_ϵ resulting in a smaller change in discharges at the jump in payments conditional on patient health.

In contrast, increasing the costs parameter c and increasing the value of downstream discharge v_d raises the value of discharging the patient both before and after the SSO threshold. Increasing the costs parameter c reduces the net profits from retaining a patient at the LTCH, with an impact that is proportional to the patient's expected length of stay. Thus, a higher value for c raises the incentive to discharge the patient in every period, and especially in the first few days when (holding h fixed) the expected length of stay is longest. Increasing the value of downstream discharge v_d raises the incentives to discharge the patient, but unlike the effect of costs, the effect is fairly constant over time. Thus, c and v_d are separately identified because of their differential impact on the first few days of the LTCH stay.

In summary, the scale parameter σ_ϵ is separately identified from the costs c and v_d because it modulates the size of the shift in incentives to discharge at the SSO threshold, while the costs and v_d parameters are mostly affected by the level of discharge rates. They are separately identified from each other because of the differential movement in the first few days of the LTCH stay. The intuition for identification on the upstream margin is similar. Identification can be achieved from the PPS moments alone, but given that we restrict these parameters to be time-invariant, it is also aided by variation in discharge patterns between the pre-PPS and PPS periods.

If health status h were observed, we could make the argument above conditional on health, and thus identify each object as a nonparametric function of h. In practice, h is unobserved, but identifying the health process is conceptually easy given our assumptions. If there were no discharges, which is roughly the case during the first week of the LTCH stay, the only attrition from the sample would be due to mortality. With only five parameters that determine the initial health distribution and how it evolves from day to day, mortality rates over five days are sufficient to identify the health process parameters, separately in the pre-PPS and PPS periods. Once the unobserved health distribution is identified, we can integrate over h and apply a similar intuition to the one we described above for the homogeneous h case. Moreover, once the health process is identified, the cross-sectional distribution of h varies over time in "known" ways, so we can also identify how the key parameters—in particular the V's—vary as a function of h.

Finally, the parameters of the post-discharge payments model ($\zeta_{0d}, \zeta_{1d}, \zeta_{0u}, \zeta_{1u}$) are identified by the sharp change in health of patients discharged on different sides of the

SSO threshold. These parameters essentially determine the mapping from the model predictions regarding the average health status of patients discharged before and after the SSO threshold to the corresponding average of observed post-discharge payments, which also exhibit a sharp change around the SSO threshold (Figure 4).

Obviously, as is typically the case, the intuition for identification requires us to have substantial variation in the data. In practice, some of the variation is not as large, and statistical power issues require us to impose more parametric structure, so the estimable model is not as flexible—especially in terms of the extent to which parameters vary with h—as the identifiable structure would be.

5. RESULTS

5.1. *Parameter Estimates and Model Fit*

Table IV presents the parameter estimates. Our point estimate of c is $1,109, implying that LTCH's actual costs are 77.8% of the average pre-SSO daily reimbursement rate. The v_{1u}, v_{0d}, and v_{1d} parameters capture the value the LTCH places on the patient's utility from being discharged to u or d relative to remaining at the LTCH, as well as any potential effect on patient health evolution in the discharge location (relative to remaining at the LTCH). The estimates imply that LTCHs are indifferent between u and d for a patient with $h = 2.0$, which is a fairly low health level. For instance, $h = 2.0$ is the 4.3th percentile of the steady-state PPS health (normal) distribution (which has mean of 5.9 and standard deviation 2.3) and corresponds to a daily mortality hazard of 2.2%. Consistent with our description of patients flowing downstream as their health improves, d is relatively better for healthier patients and u is better for sicker patients. The magnitude of the slope parameter v_{1d} is about one-fifth as large (in absolute value) as the slope parameter v_{1u}, which indicates that a given change in financial incentives will have a much larger effect on discharges on the downstream margin. These estimates are consistent with the descriptive evidence that shows a substantially larger response on the downstream margin at the SSO threshold.

TABLE IV

PARAMETER ESTIMATES[a]

Parameter		Std. Error		Parameter		Std. Error
Health process during pre-PPS:				Preferences:		
μ_0	11.31	0.139		c (000s)	1.11	0.018
σ_0	4.31	0.060		v_{1u} (000s)	−31.12	1.880
μ	0.39	0.023		v_{0d} (000s)	−76.22	2.462
ρ	0.99	0.002		v_{1d} (000s)	6.63	0.189
σ	2.45	0.002		σ_e (00s)	8.14	0.217
Health process during PPS:				Post-discharge payments:		
μ_0	5.37	0.134		ζ_{0u}	10.09	0.038
σ_0	1.87	0.060		ζ_{1u}	−0.77	0.049
μ	3.90	0.091		ζ_{0d}	12.65	0.039
ρ	0.34	0.006		ζ_{1d}	−0.28	0.009
σ	2.14	0.062				

[a]Table presents parameter estimates in our baseline specification. Standard errors are computed using the asymptotic GMM formula, where the variance-covariance matrix is computed using the bootstrap method, sampling with replacement from the set of LTCH admissions.

Relative to the slope parameters v_{1u} and v_{1d}, the scale parameter σ_ϵ on the logit error is fairly small. The estimates imply that a tenth of a standard deviation increase in the error term increases the value of discharging a patient to a given location by \$104, the product of $\sigma_\epsilon = 814$ and a tenth of a standard deviation of the logit error ($\pi/\sqrt{6}$). In contrast, a tenth of a standard deviation increase in steady-state health index (which has a standard deviation of 2.3) raises the value of discharging a patient downstream by \$1,510 ($= 0.1 \times 2.4 \times 6,629$) and lowers the value of discharging a patient upstream by \$7,087 ($= 0.1 \times 2.4 \times 31,123$), indicating that health status is capturing most of the unobserved heterogeneity in discharge behavior.

The ζ parameters capture the relationship between health at discharge and post-discharge payments in the PPS period. Consistent with our interpretation of Figure 4, the estimates indicate that post-discharge payments are declining as the patient gets healthier, with semi-elasticities of $\zeta_{1u} = -0.77$ and $\zeta_{1d} = -0.28$ on the upstream and downstream margins, respectively.

We are cautious not to over-interpret the change between the pre-PPS and PPS periods in the health process parameters. Because they are the only parameters that are allowed to vary across the time periods, they not only capture differences in the health of admitted patients but may also reflect other factors that vary over time, such as changes in medical technology or the administrative capacity of providers.

The model fits the data reasonably well. Figure D.1 presents our discharge moments and the simulated moments from the estimated model. The model does a very good job fitting the spike in discharges to u and d in the PPS period. This is particularly encouraging because this variation is a key source for identification. The model fit for the mortality patterns in the pre-PPS and PPS periods is good over the initial days, but less good at longer time horizons. This is likely due to our fairly parsimonious parameterization of the health process. The model fit is also poorer for discharges to u in the pre-PPS period.

In Appendix D, we write down and estimate two non-dynamic alternatives to our baseline model, which we discussed informally in Section 4.1 above: a myopic model and a completely static model. We show that the fit of our baseline, dynamic model is substantially better than the fit of either of the alternative, non-dynamic models.

5.2. Illustrating the Key Tradeoffs

Figure 7 provides some intuition for how the model operates, plotting the policy function at the estimated parameters. Healthy patients (higher h) are discharged to d, while sick patients (lower h) are discharged to u. Consistent with the descriptive evidence in Figure 4, LTCHs work "down the distribution" at the jump and lower their discharge threshold on the d margin, and conversely work "up the distribution" on the u margin and increase the discharge threshold. The larger shift on the d margin relative to the u margin relates directly to our discussion above on the magnitude of the slope parameter estimates in Table IV (v_{1d} and v_{1u}). The relatively small outward shift in the policy function just before the SSO threshold is consistent with the descriptive results which show limited evidence on "missing mass" immediately to the left of the SSO threshold. As noted above, such "missing mass" would be expected in a myopic model, which would produce a sharp decline in the discharge rate between the SSO day and the day that immediately precedes it.

Figure 8 unpacks the mechanism that gives rise to this policy function, providing intuition for the model's predictions. In the top panel, we display the LTCH's choice-specific payoffs (i.e., continuation values) as a function of patient health and the number of days
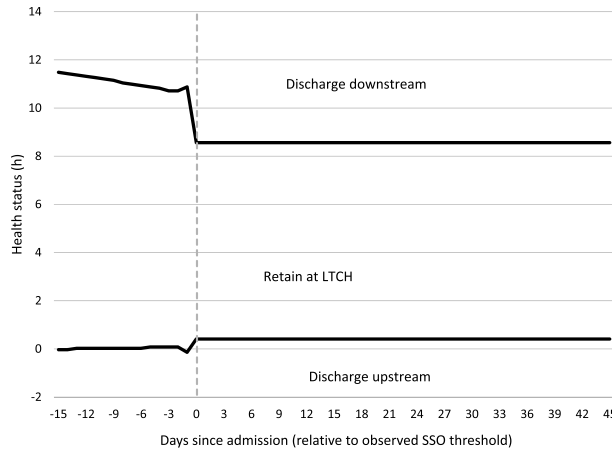
FIGURE 7.—Optimal discharge policy. Figure shows the policy function implied by the estimated model. The top black line approximates the health level above which a patient is discharged to d, and the bottom black line approximates the health level below which a patient is discharged to u. Higher h denotes better health (lower mortality). Recall that the policy function is not a deterministic function of h; given the ϵ's in the LTCH's flow payoff function (see equation (3)), h is related to discharge stochastically. The policy lines in the above figure are drawn so that at that given level of h, 50% of the patients are discharged to d (top line) and u (bottom line).

until the SSO threshold, from day -15 through day 0. The dashed line on the left shows the value of discharging a patient upstream, and the dashed line on the right shows the value of discharging a patient downstream; these are linear in patient health (by assumption) and do not vary with the patient's length of stay (by design). The solid lines show the continuation value from retaining the patient at the LTCH; these are nonlinear in patient health and vary over days $t = -15, \ldots, 0$, after which the problem becomes stationary and the value is the same as at day 0 for all subsequent t.

In the bottom panel, we show the probability of retaining the patient at the LTCH until the SSO threshold as a function of patient health for days $t = -15, \ldots, -1$. Prior to the SSO threshold, the value of retaining the patient is primarily driven by the probability of keeping the patient through the SSO threshold and obtaining the large lump-sum payment. As the bottom panel shows, even well before the SSO threshold most patients (except for the very sick patients with low h) are kept through the SSO threshold. Since there are only small increases in this probability from day to day, there is almost no change in the cutoff points at which patients are optimally discharged upstream or downstream; this explains why the policy function shown in Figure 7 is virtually flat prior to the threshold. At the SSO threshold, the dynamic incentives disappear, and the continuation value of retaining the patient drops substantially (difference between day -1 and day 0 in top panel), making discharge more appealing, and shrinking the range of health status in which it is optimal to retain patients, as shown in Figure 7.

5.3. *The Effects of Counterfactual Financial Incentives*

We use our model to simulate discharge patterns and Medicare payments under a variety of counterfactual payment schedules. Throughout this section, we assume that the initial distribution of health of admitted patients stays the same but that the subsequent
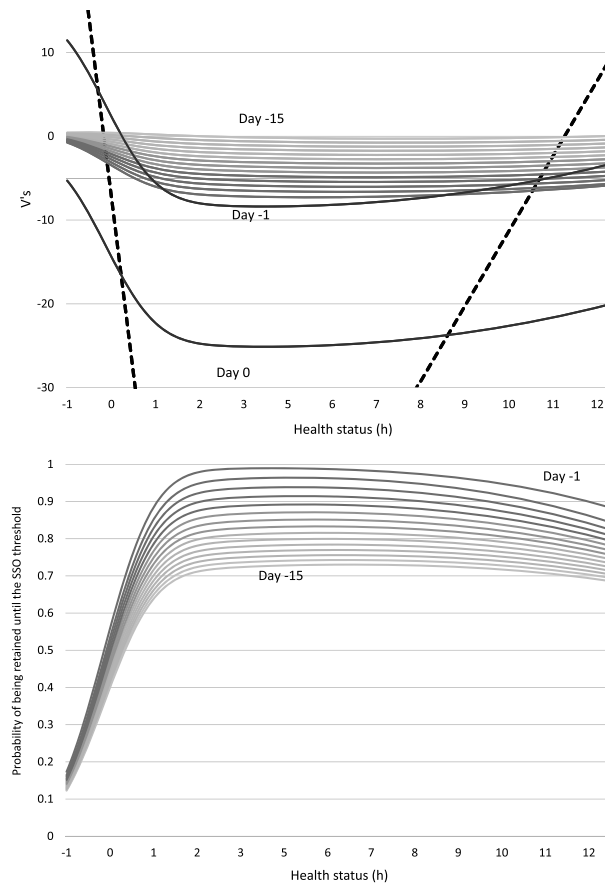
FIGURE 8.—Choice-specific continuation values as a function of the state variables. Top panel presents choice-specific continuation values as a function of the state variables: health status of the patient (on the horizontal axis) and the number of days until the SSO threshold (shown in separate lines) from day −15 through day 0. The dashed lines are the continuation values from discharging the patient upstream (left dashed line) and downstream (right dashed line), and these (by design) do not change with time to the SSO threshold. The solid lines are the continuation values from retaining the patient at the LTCH, and these do vary over time. They are monotone in days; within a day the pattern of continuation values by health status changes at day −1 (the day before the SSO threshold) when the large payment is guaranteed, and continues with a similar pattern (but much lower level) of continuation values on the threshold day (day 0). Continuation values after day 0 are identical to those shown for day 0 given the stationary nature of the problem after the threshold. The bottom panel of the figure presents the probability of the patient being retained at the LTCH until the SSO threshold (conditional on the optimal discharge policy).

discharge decisions reflect the incentives provided by the counterfactual payment schedules.

We limit our attention to alternative schedules that maintain the current practice of a cap on payments after a certain number of days. We do this both because it respects the current policy approach toward LTCH payments and because an "uncapped" schedule would lead to a small number of long stays, which is outside of the variation in our data. Specifically, we consider three main types of counterfactuals: payment schedules that remove the jump while holding the threshold day constant, payment schedules that eliminate the jump (and allow the threshold day to vary) while holding LTCHs harmless,

and cost-based reimbursement at a constant per diem, capped at 60 days. Throughout, we assume that these counterfactual payment schedules do not affect patient mortality, consistent with both the in-sample evidence presented here and the out-of-sample evidence of the impact of LTCH entry into a market in Einav, Finkelstein, and Mahoney (2018).

Removing the Jump

We start by considering two simple modifications of the baseline payment schedule that eliminate the jump in payments at the SSO threshold, but, like the baseline PPS payment schedule, provide no payments on the margin for stays in excess of the SSO threshold. Figure 9 plots these counterfactual payment schedules and the baseline schedule for comparison. The top panel shows a counterfactual we call "higher rate per day," which eliminates the jump by increasing the per diem payment from \$1,380 to \$2,145 prior to the SSO threshold but holds the post-threshold payment fixed. The bottom panel shows a counterfactual schedule we call "lower cap," which eliminates the jump in payments at the SSO threshold by reducing the PPS payment from \$34,319 to \$22,074 but holds the pre-
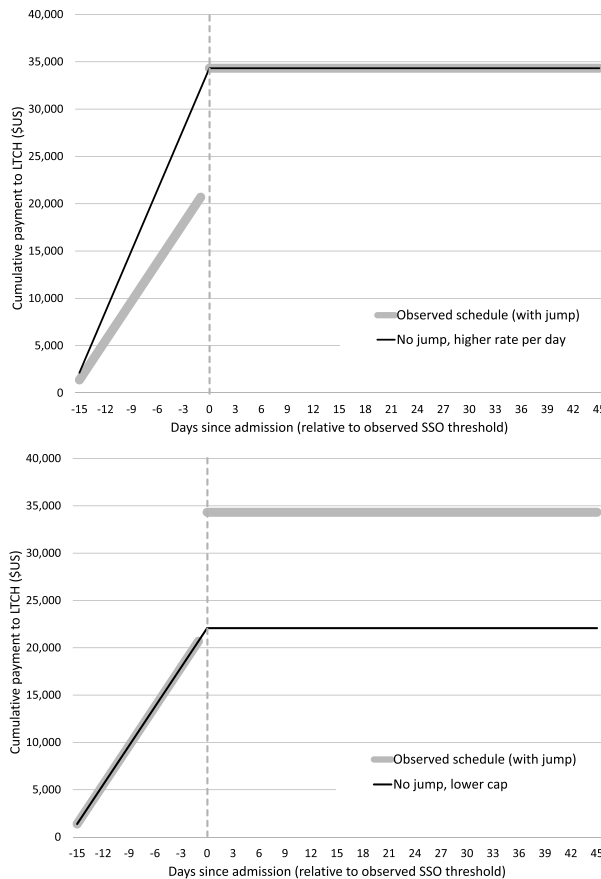

FIGURE 9.—Counterfactual payment schedules. Figure shows the observed (PPS) payment schedule (thick gray line in both panels) and the first two counterfactual payment schedules we consider (black line in each panel). Both counterfactual schedules eliminate the jump in payments at the SSO threshold, but do this in different ways.
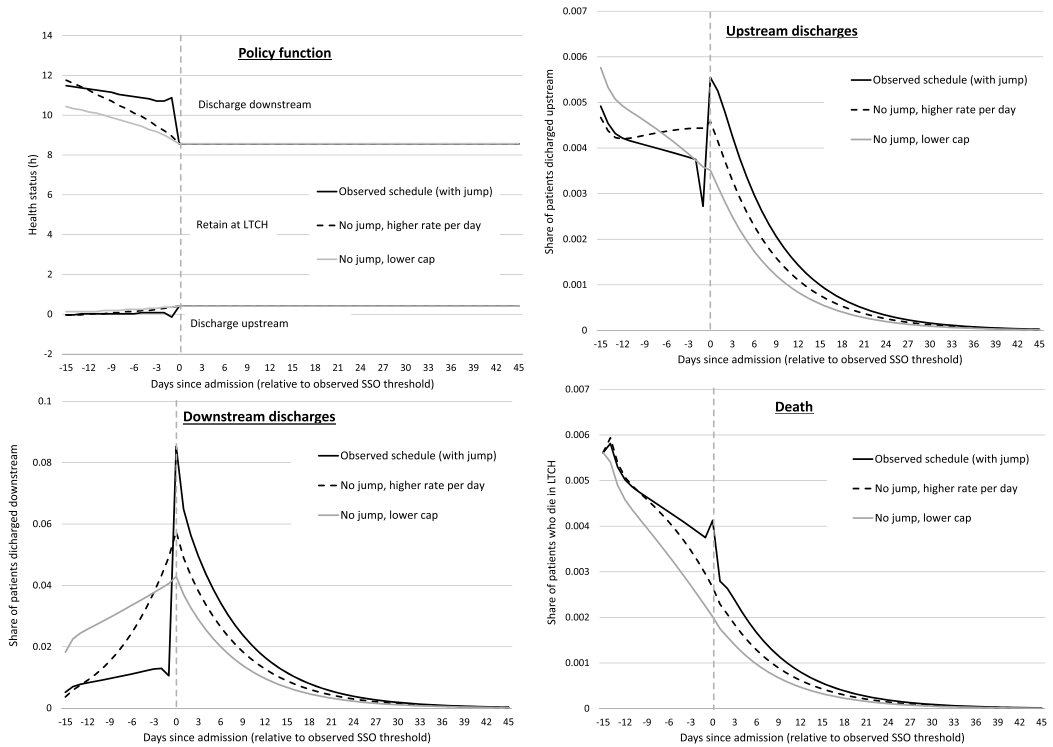
FIGURE 10.—Counterfactual policy functions and discharge patterns. Top left panel shows the implied discharge policy function from the two "no jump" counterfactual payment schedules described in the main text and illustrated in Figure 9. The discharge policy function associated with the observed contract design is shown in gray and is the same as the one reported in Figure 7. The three other panels show discharges (upstream, downstream, and to death) under these two counterfactual payment schedules. The solid black line reports results that are based on our parameter estimates (reported in Table IV) and the observed payment schedule, and each other line reports the results under a different counterfactual payment schedule.

SSO per diem payment fixed. The "higher rate per day" contract is weakly more generous than the baseline schedule, while the "lower cap" contract is weakly less generous.

We use our model to simulate discharge patterns and Medicare payments under these two counterfactuals. Figure 10 shows the policy functions under each payment schedule and the corresponding discharge patterns. Table V summarizes the impact of each of these payment schedules on Medicare payments to LTCHs and to other facilities; Appendix F assesses the sensitivity of these results to alternative specifications.

The black dashed line in the top left panel of Figure 10 shows the policy function under the "higher rate per day" counterfactual. During the first few days, the policy function is similar to that under the observed schedule. However, as the length of stay increases, the elimination of the jump reduces the incentive to retain patients, and the policy function shifts inwards on the upstream and downstream margins. The black dashed lines in the other panels of Figure 10 show discharge patterns under this counterfactual. Mirroring the changes in the policy function, the "higher rate per day" counterfactual increases discharges in the 10 days prior to the SSO threshold, relative to the observed schedule. As shown in column (2) of Table V, these higher discharge rates reduce average length

TABLE V

DISCHARGES AND PAYMENTS FROM COUNTERFACTUAL PAYMENT SCHEDULES[a]

	Observed Schedule (1)	Higher Rate per Day (2)	Lower Cap (3)	Lowest Medicare Payment Within "LTCH Preferred" Schedules (4)	Linear Schedule at Estimated Cost (5)	Linear Schedule at "Opportunity" (SNF) Cost (6)
LTCH payments:						
Total payments	27,953	28,316	16,024	26,313	35,546	7,392
Total profits	6,518	8,991	−407	6,876	−971	−8,865
Average LOS[b]	19.3	17.4	14.8	17.5	32.9	14.7
Payment per day	1,446	1,625	1,082	1,501	1,080	504
Discharges upstream:						
Total payments	3,813	3,551	3,223	3,583	5,911	3,240
Share of discharges	0.11	0.11	0.10	0.11	0.17	0.10
Payment per discharge	33,669	33,315	32,561	33,312	35,513	32,494
Discharges downstream:						
Total payments	16,031	16,009	16,153	15,795	10,394	15,247
Share of discharges	0.79	0.80	0.83	0.80	0.66	0.83
Payment per discharge	20,367	19,913	19,548	19,674	15,648	18,436
Total Medicare payments	47,796	47,877	35,399	45,691	51,851	25,879
Counterfactual payment schedule						
Pre-SSO per diem	1,380	2,145	1,380	1,931	1,109	507
Cap amount	34,319	34,319	22,074	32,830	66,540	30,420

[a]Table presents results from the counterfactual payment schedules. Column (1) reports results that are based on our parameter estimates (reported in Table IV) and the observed payment schedule, and each other column reports the results from predicted discharge patterns under a different counterfactual payment schedule. The counterfactual payment schedules we consider are described in the main text.

[b]Length of stay is measured from day −15. To make it comparable to the summary statistics reported in Table I, all numbers should be increased by 7.5 days (because the average SSO threshold across admissions in our sample is 22.5 days).

of stay by 1.9 days or 7.1%.[21] Eliason et al. (forthcoming) reported a similar exercise, and obtain qualitatively similar results.[22]

Despite the significantly higher per day payments prior to the jump ($2,145 versus $1,380), Medicare payments to LTCHs increase by a very small amount of $362 or 1%. The small increase is due to a large behavioral response to the incentives. Holding discharge patterns fixed, LTCHs would get paid about 50% more per day for stays below the SSO threshold, but in response to the elimination of the jump, patients are now discharged earlier, so overall payments are lower. Holding discharge patterns fixed, we calculate that the mechanical effect of this counterfactual is a $1,748 increase in Medicare payments to the LTCH, implying that the behavioral response to the removal of the jump

---

[21]Length of stay is measured from day −15. Because the average SSO threshold in our sample is 22.5 days, values for length of stay need to be increased by 7.5 (=22.5 − 15) days to match the summary statistics.

[22]Specifically, Eliason et al. (forthcoming) reported on a "per diem counterfactual," which is very similar to our "higher rate per day" exercise. They found a similar length of stay effect (1.25 days shorter, relative to our 1.9 estimate) and a modest effect on total Medicare cost. It is important to note that the comparison is imperfect: the counterfactual is not exactly the same, the model and the observation periods are slightly different, and they focused on the most common DRGs while we use all LTCH admissions.

reduces Medicare payments to the LTCH by $1,394 per admission. LTCH profits per admission rise by $2,441 or 35% relative to the observed schedule.

We next explore the effects of the "lower cap" payment schedule. The gray line in the top left panel of Figure 10 shows the policy function under this counterfactual. The elimination of the jump in payments shifts the policy function inwards during the entire pre-threshold period, relative to that under the observed schedule. The gray lines in the other panels of Figure 10 show that discharges correspondingly rise, with the daily share of discharges to d increasing four-fold and the share of discharges to u increasing more modestly over most of the pre-threshold period. As shown in column (3) of Table V, average length of stay is reduced by 4.5 days, and payments to the LTCH are reduced by $11,967 or 43%. The mechanical effect (holding observed discharge patterns fixed) of the "lower cap" payment schedule is a reduction in payments of $8,851 or about 74% of the overall reduction, with the remaining 26% due to the behavioral response. LTCH profits per admission fall by $7,030 per admission and are estimated to be negative, a point we return to below.

The remaining rows of Table V consider the impact of these counterfactual payment schedules on Medicare payments throughout the rest of the episode of care. For these counterfactuals, the spillovers on post-discharge payments are small. For the "higher rate per day" counterfactual, post-discharge payments for patients discharged to u and d are affected by a few hundred dollars. For the "lower cap" counterfactuals, the decrease in post-discharge payments is larger but still small compared to decrease in LTCH payments. Combining the effects on LTCH payments and post-discharge payments, our estimates indicate that the "higher rate per day" has virtually no effect on total Medicare payments ($74 increase) and the "lower cap" reduces total Medicare payments by a substantial $12,456 or 26%.

While interesting, neither of the above counterfactuals is fully satisfactory. While the "lower cap" counterfactual suggests that alternative payment schedules could substantially reduce Medicare payments, the large decrease in LTCH revenue (and in estimated profits) might have unintended consequences. For instance, under this payment schedule, LTCHs might cut back on socially valuable fixed investments or even exit the market. In contrast, the "higher rate per day" counterfactual, while clearly making LTCHs better off, has virtually no effect on Medicare payments. Yet, these two exercises suggest that there might be "intermediate" contracts that generate cost savings without the risk of unintended consequences. We explore such counterfactuals below.

"Win-Win" Payment Schedules

With the above considerations in mind, we now consider a set of counterfactuals that hold LTCH revenue fixed under the observed discharge patterns. Faced with these contracts, if LTCHs do not change their behavior, they will have identical revenue, costs, and profits as they would under the observed payment schedule. If LTCHs change their behavior, by revealed preferences, they must be (weakly) better off. Therefore, by design, these contracts should not have a negative impact on LTCHs.

In the same spirit as the previous counterfactuals, we consider contracts that pay a constant per diem amount up to a threshold length of stay, at which point the payments are capped and per diem payments drop to zero (obviously, with no jump). We consider contracts where the payment is capped at thresholds in a $+/- 10$ day range on either side of the current SSO threshold day. Since the generosity of the contract is strictly increasing in the per diem rate, for a given length of stay at which payments are capped, there is a
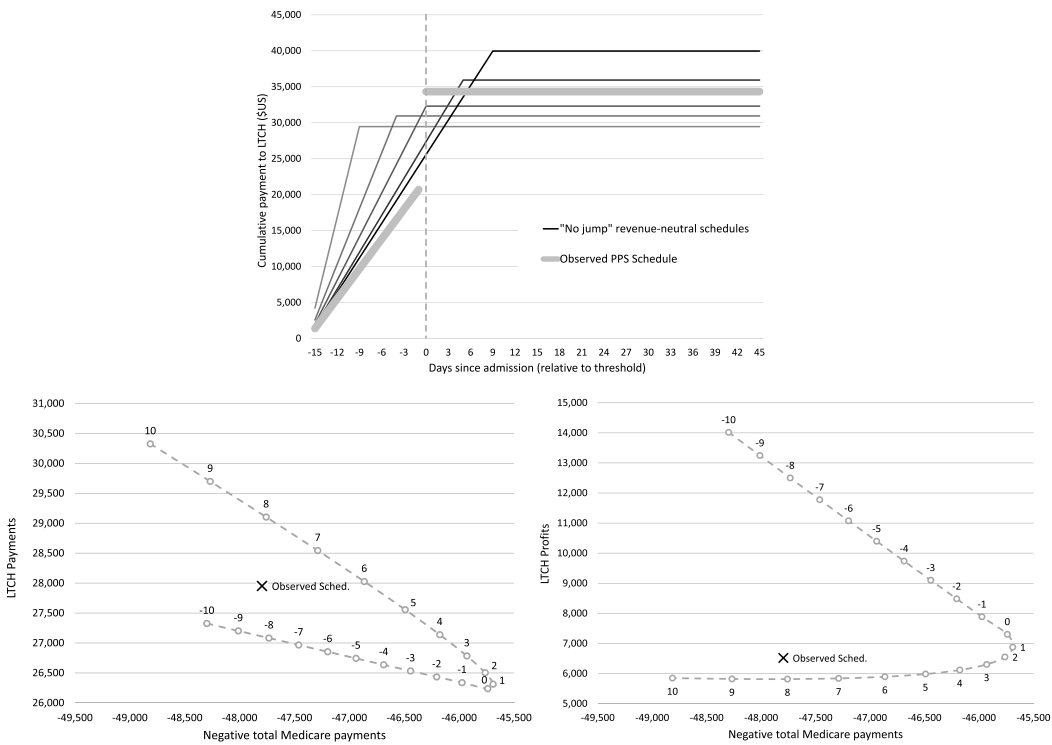
FIGURE 11.—"Win-win" payment schedules. The top panel shows some examples of the 21 potential "win-win" contracts we consider. All contracts pay a constant amount up to a threshold length of stay, where they are capped (so that per diem rate drops to zero) with no jump at the threshold. We consider threshold days ranging from $+/-$ 10 days of the current threshold, with the unique payment schedule defined for each threshold day as the one that would hold payments to the LTCH (i.e., LTCH revenue) fixed if they did not change their discharge behavior under the observed contracts. The bottom panels show outcomes (given the LTCH's counterfactual behavior) under these various potential "win-win" payment schedules shown in the top panel. For each schedule (represented by a dot which is labeled with the day the payment schedule switches from a per day rate to a cap), the bottom left panel shows LTCH payments per admission against (the negative of) total Medicare payments (including estimated post-discharge payments) for the episode of care; and the bottom right panel shows LTCH profits per admission against total Medicare payments.

unique payment schedule that holds LTCH revenue fixed under the baseline discharge patterns.

The top panel of Figure 11 plots a set of these contracts, where to avoid overcrowding the figure, we show payment schedules for only a subset of the 21 contracts considered. The bottom left panel of Figure 11 plots LTCH payments against total Medicare payments (including estimated post-discharge payments) for each of the contracts we consider, while the bottom right panel plots estimated LTCH profits against total Medicare payments. For comparison, both plots also show outcomes under the observed payment schedule.

The figures indicate that there is a broad set of "win-win" payment schedules that reduce total Medicare payments for the episode of care while leaving LTCHs (weakly) better off. Every counterfactual contract with a threshold between -8 and 8 days reduces Medicare spending, although there is substantial heterogeneity in the reduction. LTCH revenues increase for every contract with a threshold of 6 to 10 days and decline for contracts with a threshold of -10 to 5 days. Because LTCHs value both profits and patient

utility, LTCH profits under counterfactuals do not necessarily increase. LTCH profits are higher than their baseline level for contracts with a threshold -10 to 2 days and are lower for thresholds of 3 to 10 days. Counterfactuals that decrease profits do not lower them by a substantial amount.

The counterfactual with payment threshold of 1 day more than the current SSO threshold results in the largest reduction in Medicare spending and is a natural contract to focus on. Column (4) of Table V shows outcomes for this contract. Under this payment schedule, Medicare payments to LTCHs are reduced by $1,655 or 5.9%. Accounting for Medicare payments across the entire episode of care leads to somewhat higher savings of $2,127, or 4.4% of total episode payments. Despite the reduced payments, LTCH profits rise by $315 per stay or 4.5%: the decline in LTCH revenue is offset by lower costs, as length of stay is almost 2 days (7%) shorter.

Cost-Based Payment Schedules

The last two columns of Table V report results from two additional counterfactuals that explore cost-based reimbursement at a constant per diem rate. The first counterfactual pays LTCHs a constant per diem of their estimated costs, which is $1,091 per day. The second counterfactual pays LTCHs a constant per diem of $507, which is the average per day payment to SNFs during the post-discharge period. We think about this counterfactual as a form of "reference based pricing" where Medicare pays LTCHs the opportunity cost to Medicare of the patient—that is, the amount Medicare would have incurred for the patient at a location that provides fairly similar care (at least for the marginal patient). To avoid extrapolating too far outside of our data, for both of these counterfactuals we cap payments after 60 days, which leads LTCHs to discharge virtually all of their patients within 90 days of the current SSO threshold.

Paying LTCHs their estimated costs leads to a substantial $6,026 increase in payments to LTCHs and a smaller $2,509 overall increase in total Medicare payments. Payments increase because LTCHs retain patients for longer time periods rather than discharging them to SNFs where Medicare payments would be lower, with average length of stay increasing from 19.2 to 32.0 days. Paying LTCHs the average per diem for SNFs leads to a massive decrease in LTCH payments and total Medicare payments, accompanied by a sharp reduction in length of stay. Of course, concerns about unintended consequences, which we discussed in the context of the "lower cap" schedule, are also relevant here.

6. CONCLUSIONS

In this paper, we examined the impact of provider financial incentives in post-acute care (PAC), a setting with large stakes both for the government budget and patient health that has received scant attention in the academic literature. Within the context of PAC, we examined the impact of a jump in Medicare payments to long-term care hospitals (LTCHs) that occurs after a pre-specified length of stay, when reimbursement shifts from a per diem rate to a lump-sum payment.

The descriptive evidence showed a large response by LTCHs to the jump in payments. At the threshold, there is a large spike in discharges. The marginal patient affected by the payment threshold is relatively healthy. We are unable to detect any impact on patient mortality at the threshold, even in this high-mortality population.

This descriptive evidence motivated our specification of a stylized dynamic model of LTCH discharge behavior. We estimated the model and used it to examine the implications of alternative payment schedules, including "win-win" contracts that hold LTCHs

(and presumably their patients) harmless, while reducing Medicare payments. The contract with the largest Medicare savings reduced total Medicare payments by nearly 5% while increasing LTCH profits by a similar percentage.

We also considered more aggressive payment schedules that resulted in substantially higher Medicare savings but raised the possibility of unintended consequences due to a large reduction in LTCH profits. In particular, in our model we take admission to the LTCH as given and focus on the impact of counterfactual policies for this fixed set of patients. However, the large reduction in profits brought about in our more aggressive counterfactuals may affect which patients are admitted to an LTCH and might have even broader effects on the market, for instance through LTCH entry and exit decisions. We consider an important area for further work to move "up" the healthcare pathway and model the ACH's decision of whether to discharge a patient to an LTCH or another PAC provider.

More broadly, our results indicate how economic models and data can be combined to better inform contract design. A small dose of common sense is sufficient to see that the sharp jump in payments at the threshold is inefficient, and that some alternative payment schedule should be better for both the Medicare payer and the LTCH. Data and descriptive evidence, however, were important to demonstrate that the behavioral response by the LTCH to the current payment schedule was quantitatively meaningful, and an economic model—parameterized and estimated on the data—was necessary to identify specific "win-win" contracts that could create opportunities for both LTCHs and Medicare to gain. While naturally our results are specific to our particular setting, we hope that this type of approach can inform future work examining the impact of providers' financial incentives not only for the directly affected provider but throughout the healthcare system.

APPENDIX A: POST-DISCHARGE PAYMENTS AND DAYS

Our starting point of the analysis is an admission to an LTCH. We can observe all discharge destinations from an LTCH. Table A.I shows the share of discharges to different locations. In the PPS period, 13% of patients die during their LTCH stay, another 14% are discharged upstream (with the vast majority going to inpatient care), and 73% are discharged downstream (with approximately half of these patients going to an SNF/IRF and 45% going home, where they may receive care from an HHA).

We define a post-discharge "episode of care" as the spell of almost continuous days following discharge from an LTCH with Medicare payments to an ACH, SNF/IRF, or LTCH. In particular, the episode ends if there are at least two days without Medicare payments to these institutions.

Although in the MedPAR data we can observe all discharge destinations, we can only observe post-discharge Medicare payments and days for ACH and for PAC facilities (SNF, IRF, LTCH), but not for home health visits or hospice. To address the fact that we do not observe payments or days at HHAs or hospices, we supplement the MedPAR data with annual spending and utilization from the Beneficiary Summary File (BSF) Cost & Use file. For every stay in the MedPAR data, we observe whether the patient was discharged to an HHA or hospice at some point in the episode of care. For patients who were discharged to an HHA or hospice, we impute the patient's payments and days using the annual BSF data. In practice, HHA and hospice payments are quite small as a share of the total. For example, we estimate that of individuals with an LTCH discharge, LTCH and SNF/IRF payments constitute over 90% of total PAC payments, with home health accounting for only 8%.

TABLE A.I

DISCHARGE DESTINATIONS[a]

	Pre-PPS	PPS
Death	**0.138**	**0.132**
Upstream	**0.191**	**0.137**
Inpatient	0.981	0.761
Hospice	0.019	0.239
Downstream	**0.672**	**0.731**
LTCH	0.001	0.006
SNF	0.240	0.434
IRF	0.006	0.066
Home health	0.214	0.302
Home	0.453	0.139
Other	0.085	0.053
Total	1.00	1.00
Number of Obs.	218,857	587,385

[a]Table shows the percentage of discharges to death, upstream, and downstream in pre-PPS and PPS periods. The upstream and downstream discharges are further decomposed.

Since these annual amounts include some payments and days that occur before or after the "episode of care," our imputation likely leads us to overestimate post-discharge Medicare payments and days. However, we think that the approach provides a reasonable approximation. Table A.II shows that our estimates of post-discharge payments and facility days are not affected much if we instead impute 0 costs for HHA and hospice, and 0 days for hospice.

APPENDIX B: LTCH PAYMENT SYSTEMS

Prior to fiscal year 2003 (i.e., October 2002), CMS reimbursed LTCHs on a cost-based system. At the start of fiscal year 2003, CMS began transitioning LTCHs to a prospective payment system (PPS). The PPS, which was fully phased in by the start of fiscal year 2008 (i.e., October 2007), is the focus of our study. This appendix describes it in more detail, drawing heavily on Kim et al. (2015), 3M Health Information Systems (2015), and MedPAC (2014).

TABLE A.II

POST-DISCHARGE OUTCOMES[a]

	Pre-PPS (Jan 2000–Sep 2002)			PPS (Oct 2007–July 2012)		
	Overall	Upstream	Downstream	Overall	Upstream	Downstream
Number of discharges (000s)	188.7	41.8	147.0	509.7	80.3	429.4
Post-discharge payments (upper bound)	13,100	31,405	7,901	22,808	35,775	20,382
Post-discharge payments (lower bound)	12,106	30,712	6,821	20,144	33,446	17,655
Post-discharge facility days (upper bound)	17.1	32.8	12.6	26.1	33.0	24.8
Post-discharge facility days (lower bound)	16.9	32.1	12.6	24.7	26.0	24.4

[a]Table presents upper and lower bounds for our imputation of post-discharge payments and days using the baseline sample of LTCH stays described in Table I, excluding discharges due to death. Appendix A provides more detail. The upper bound is used for our empirical analysis.

B.1. *LTCH PPS Rules*

In contrast to the cost-based system, which had reimbursed hospitals based on the estimated cost of each patient's case, the PPS outlined a fixed reimbursement amount for each patient, based on the patient's DRG. These DRG-based lump-sum payments were meant to reflect the typical resources consumed by each type of patient. However, in order to discourage short stays in hospitals which were meant to provide long-term care, the PPS includes a short stay outlier (SSO) threshold, with reduced payments below the full DRG payment for LTCH patients who are discharged before a DRG-specific threshold.

Full DRG Payment

The full DRG payment is computed as

$$\text{Full DRG Payment} = \text{Adjusted Federal Rate} \times \text{DRG Relative Weight,} \tag{B.1}$$

where

$$\begin{aligned}
&\text{Adjusted Federal Rate} \\
&= (\text{Unadj. Federal Rate} \times \text{Labor-Related Share} \times \text{Wage Index}) \\
&\quad + (\text{Unadj. Federal Rate} \times \text{Nonlabor-Related Share}).
\end{aligned} \tag{B.2}$$

This payment structure is similar to the much-studied Inpatient PPS used for (regular) acute care hospitals (ACHs) that was introduced in 1983, but differs in two ways. First, although the DRGs are defined in the same way for the LTCH and Inpatient PPS, the relative weights associated with DRGs have different values in LTCH-PPS. Second, the LTCH-PPS unadjusted federal rate is larger than the corresponding Inpatient PPS value. The result is that LTCH-PPS payments are substantially greater than Inpatient PPS payments for the same DRG, presumably to reflect the greater costs at an LTCH relative to an ACH.[23]

Short Stay Outlier (SSO) Payment

If an LTCH stay has a length of stay (LOS) shorter than or equal to five-sixths of the geometric average length of stay (ALOS) for the DRG, it is paid as a short stay outlier. We call the smallest integer greater than five-sixths of the geometric ALOS the SSO threshold. The SSO threshold is constant within a DRG-PPS Rate Year (with the exception of 2009).

A short stay outlier is paid the *lowest* of the following:
1. Full DRG Payment.
2. 120% of the DRG per diem amount times the length of stay, where the DRG per diem amount is defined as the ratio of the full DRG payment to the geometric average of

---

[23]Also, like Inpatient PPS, LTCH PPS offers a High Cost Outlier (HCO) payment for particularly costly stays. Specifically, an LTCH can receive a HCO payment if the cost of the case exceeds the HCO Threshold. The HCO payment is made in addition to the regular payment amount. Importantly, for our purposes, HCO payments can be made regardless of whether the LTCH stay is considered an SSO outlier or eligible for the full DRG payment. We therefore exclude HCO payments from our analysis and model. About 9% of LTCH stays in our baseline sample have HCO payments, and the median HCO payment in our baseline sample is $12,428.

the LOS for that DRG. This option is roughly equivalent to a linear interpolation of the full DRG payment between Day 0 and the SSO threshold.[24]

3. 100% of the cost of the case, which is computed as total charges multiplied by the facility-specific cost-to-charge ratio.

4. A blend of the Inpatient PPS amount (used at ACH) and 120% of the DRG per diem amount. Note that this option converges to option 2 as LOS increases.

B.2. *Empirical Payment Schedules*

PPS Payment Schedule

We use a commercial software offered by the company 3M (the product is called "Core Grouping Software" (CGS)) to compute counterfactual Medicare payments for each post-PPS period stay.[25] Specifically, for each stay in the PPS period, we compute the PPS payment for the actual discharge day and each possible counterfactual discharge day. The inputs into this calculation are the admission date, estimated hospital charges, principal and secondary diagnoses, procedures, discharge status, age, and sex of the patient. For counterfactual lengths of stay, we assume that hospital charges scale linearly with the observed length of stay.

With this information, the software produces the DRG code, the SSO threshold day, and the total Medicare payment for each length of stay. To validate the software, we compare the predicted DRG against the DRG we observe in the data, and the predicted payment against the observed payment for the observed length of stay. The predicted DRG matches the observed value in 99.9% stays and the predicted Medicare payment is within one dollar of the observed Medicare payment in 90% of stays.

Figure B.1 illustrates the resultant, estimated payment schedules for both the pre-PPS and PPS periods. Note that this figure differs slightly from Figure 1 in the paper, which depicts a stylized model of the post-period payment schedules in which the pre-threshold payments are constant per diem. In practice, the pre-threshold payments appear to be slightly bowed downwards; we abstract from this in Figure 1 which we use in our model estimates, where we use the average payment per day for stays discharged before the threshold to construct the slope of the payment schedule prior to the threshold.

What features of the payment rule created the jump in payments at the SSO threshold? Recall that right of the SSO threshold, short stay outlier rules do not apply and the payment is just the full DRG payment, which means the cumulative payment schedule is always a flat line to the right of the threshold. To the left of the SSO threshold, each stay is paid the minimum of four alternative payments; the shape of the payment schedules therefore depends on which of the four alternatives is binding. If options 1, 2, or 4 were binding, we would not see a jump at the threshold. Therefore, we conclude that cost of the case must be binding in most cases because we observe a jump on average. Note that the cost of case being binding is necessary rather than the sufficient condition for creating a jump in the payment schedule; the costs could theoretically be such that the payment schedule only has a kink at the SSO threshold rather than a jump. In practice, however, the cost of the case is on average lower than the other options, and we see a jump at the threshold.

---

[24]To see this, note that 120% of the DRG per diem amount times the length of stay is approximately equal to 120% × (Full DRG payment)/((6/5)SSO Threshold) × LOS, which is equal to (Full DRG payment)/(SSO Threshold) × LOS.

[25]For more information about this software, see: http://solutions.3m.com/wps/portal/3M/en_US/ Health-Information-Systems/HIS/Products-and-Services/Products-List-A-Z/Core-Grouping-Software/.
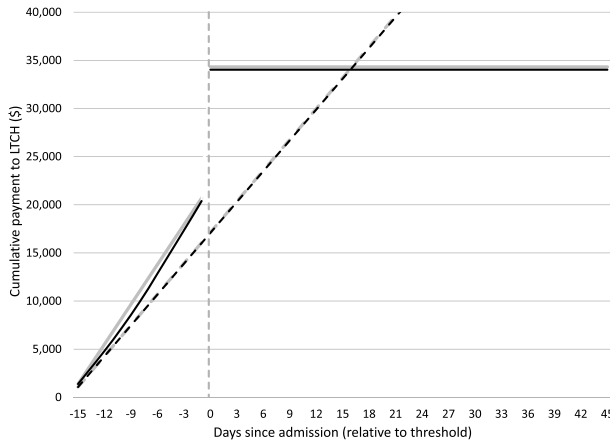
FIGURE B.1.—Empirical versus approximated payment schedules. Figure presents the payment schedules used in the paper (gray lines, which are the same as Figure 1 in the main text) against the observed payments (black lines). Appendix B provides more detail about the (slight) differences.

Pre-PPS Payment Schedule

In the pre-PPS period, LTCHs were paid their (estimated) costs, up to facility-specific per day limit (MedPAC (2014)). For most facilities, this limit was binding. For these facilities, we calculate the LTCH payment schedule as the per day limit multiplied by the length of the stay. For a small number of facilities, the payment limit does not appear to bind. For these facilities, we assume that reported costs are linear in the patient's length of stay, and we calculate the payment schedule as the (imputed) per day cost multiplied by the length of stay. When we analyze discharge patterns in the pre-PPS period, we assign each stay the SSO threshold it would have had in the first year of the PPS period, based on the DRG assignments made using the CGS software described above.
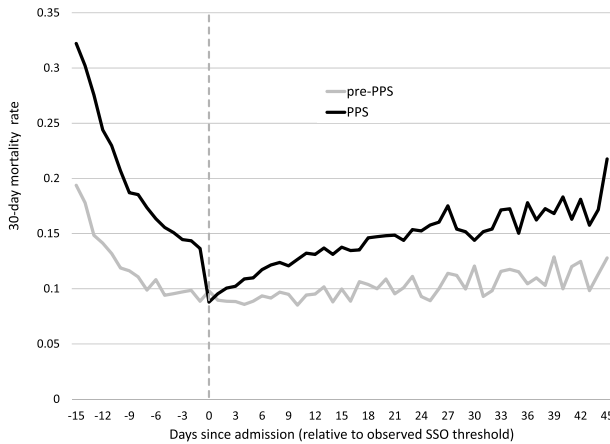

FIGURE B.2.—Post-discharge mortality rates. Figure presents the (forward looking) 30-day mortality rate after a (live) discharge, as a function of the day of discharge.

APPENDIX C: MORTALITY ANALYSIS

We formally test for a mortality effect using a regression discontinuity (RD) design. Let i index individuals and t index days relative to the SSO threshold. Let y_{it} be a mortality indicator. For our analysis of the 1-day mortality hazard, y_{it} takes a value of 1 if the individual dies on day t and takes on a value of 0 if the individual is alive. For the 30-day mortality analysis, y_{it} takes on a value of 1 if the individual dies in the subsequent 30 days and takes on a value of 0 if the individual does not die over this period. Individuals who have already died are excluded from the analysis.

In our baseline RD specification, we allow for a linear trend in the running variable t and permit this linear trend to vary on different sides of the SSO threshold:

$$y_{it} = \alpha_0 + \alpha_1 t + 1_{t \geq 0}(\beta_0 + \beta_1 t) + \varepsilon_{it}. \tag{C.1}$$

The coefficient of interest β_0 captures the change in mortality at the SSO threshold, conditional on the linear controls. To confirm the robustness of our findings, we also estimate a specification with a quadratic time trend that, as before, is also allowed to vary on different sides of the SSO threshold:

$$y_{it} = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + 1_{t \geq 0}(\beta_0 + \beta_1 t + \beta_2 t^2) + \varepsilon_{it}. \tag{C.2}$$

In both specifications, we restrict our analysis to observations close to the threshold, focusing on bandwidths of 3, 5, and 10 days within the threshold. We cluster our standard errors at the DRG level, which allows for correlation in the health process not only within an individual over time but also within the set of individuals who have the same DRG and therefore may exhibit correlated mortality profiles. We focus the mortality analysis on the post-PPS period.

Table C.I shows the parameter estimates. Panel A reports the effect on the level of the 1-day mortality hazard (the β_0 coefficient). Column (1) of Panel A, which shows our baseline specification with a linear time trend and a 3-day bandwidth, indicates that 1-day mortality increases by less than 0.01 percentage point at the threshold. This estimate is tiny in absolute magnitude, small relative to the baseline daily mortality rate of 0.6%, and is statistically indistinguishable from zero. Columns (2) to (6) show that this finding is robust to alternative bandwidths and a quadratic time trend.

We use two approaches to examine more gradual effects of the threshold on mortality patterns. In Panel B, we report the effect on the slope of the 1-day mortality hazard (the β_1 coefficient) from the linear specification (equation (C.1)). If distortions in the location of care have an effect on the evolution of health, we might expect a change in the slope of the mortality hazard at the threshold, even if there is not an on-impact effect on the level. The point estimates are small, statistically insignificant, and robust to alternative bandwidths. In Panel C, we show effects on the level of the 30-day mortality hazard (the β_0 coefficient), which is also designed to measure more gradual effects. Column (1) of Panel C, which again shows the baseline specification with a linear time trend and a 3-day bandwidth, indicates an economically tiny and statistically insignificant 0.005 percentage point decline in 30-day mortality at the threshold (relative to a baseline 30-day mortality rate of 13.4%). As before, the effect is robust to alternative bandwidths and a quadratic time trend.

To complement the regression tables, in Figure C.1 we show standard RD plots for the 1-day and 30-day mortality effects. The dots show the underlying data. The solid lines show local linear regressions, constructed using a 3-day bandwidth and a uniform kernel

TABLE C.I

REGRESSION DISCONTINUITY EFFECT ON MORTALITY[a]

	Linear			Quadratic		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Effect on 1-day mortality hazard						
Post Threshold Indicator	0.00007	0.00021	0.00039	0.00003	−0.00013	0.00018
	(0.00023)	(0.00017)	(0.00013)	(0.00048)	(0.00030)	(0.00019)
	[0.763]	[0.213]	[0.003]	[0.954]	[0.671]	[0.336]
Bandwidth	3	5	10	3	5	10
No. of Obs.	3,617,134	5,685,012	10,864,654	3,617,134	5,685,012	10,864,654
Panel B. Effect on 1-day mortality hazard slope						
Post Threshold Indicator	−0.00013	0.00004	0.00002			
	(0.00008)	(0.00004)	(0.00002)			
	[0.115]	[0.264]	[0.149]			
Bandwidth	3	5	10			
No. of Obs.	3,617,134	5,685,012	10,864,654			
Panel C. Effect on 30-day mortality hazard						
Post Threshold Indicator	−0.00005	−0.00023	0.00013	0.00013	−0.00007	−0.00016
	(0.00016)	(0.00016)	(0.00022)	(0.00026)	(0.00021)	(0.00021)
	[0.778]	[0.156]	[0.548]	[0.608]	[0.751]	[0.425]
Bandwidth	3	5	10	3	5	10
No. of Obs.	3,617,134	5,685,012	10,864,654	3,617,134	5,685,012	10,864,654

[a]Table shows results from the regression discontinuity mortality analysis described in Appendix C. Columns (1)–(3) use a linear functional form (see equation (C.1)) before and after the SSO threshold, while columns (4)–(6) use a quadratic functional form (see equation (C.2)). Panels A and C report the estimate of the β_0 coefficient, which captures the jump in mortality rate at the SSO threshold in the PPS-period; Panel B reports the estimate of the β_1 coefficient, which captures the change in the slope of the mortality rate at the threshold in the PPS-period. Each column restricts the analysis to a different bandwidth number of days before and after the SSO threshold. The 1-day mortality hazard is defined as the share of individuals alive at a given day who die by the next day; the 30-day mortality hazard is similarly defined as the share of individuals alive at a given day who die in the next 30 days. All mortality rates are calculated unconditional on patient's location. Standard errors, clustered at the DRG level, are in regular brackets and *p*-values in square brackets.

so that they correspond to our baseline specification (Table C.I, column (1)) where we estimate a linear regression on a window of $+/-$ 3 days around the discontinuity. The dashed lines show 95% confidence intervals, constructed by bootstrapping with replacement over DRGs. The plots visually confirm the regression evidence which showed no jump in either the 1-day or 30-day mortality hazard at the threshold.

Because our standard errors in the regression discontinuity analysis rely on difficult-to-test assumptions about the correlation structure of the error term, we assess the robustness of our statistical inference using permutation inference (Rosenbaum (1984, 2002), Abadie, Athey, Imbens, and Wooldridge (2014)). Specifically, we estimate equation (C.1) with a bandwidth of 3, replacing the dummy variable for being to the right of the SSO threshold with a dummy variable for being to the right of placebo thresholds defined at $t = -12$ and $t = 42$ in the pre- and post-PPS periods. That is, we estimate an RD effect for a placebo threshold at each day starting 3 days after the start of our sample and ending 3 days before the end (to allow for a 3-day bandwidth); we also exclude days -3 to 3 in the post-PPS period since these days might be contaminated by a potential treatment effect.

Figure C.2 plots the actual effect and the distribution of placebo estimates for the 1-day and 30-day mortality hazards. The plots show that the actual change in mortality at the SSO threshold is not particularly large relative to the typical day-to-day variation in the
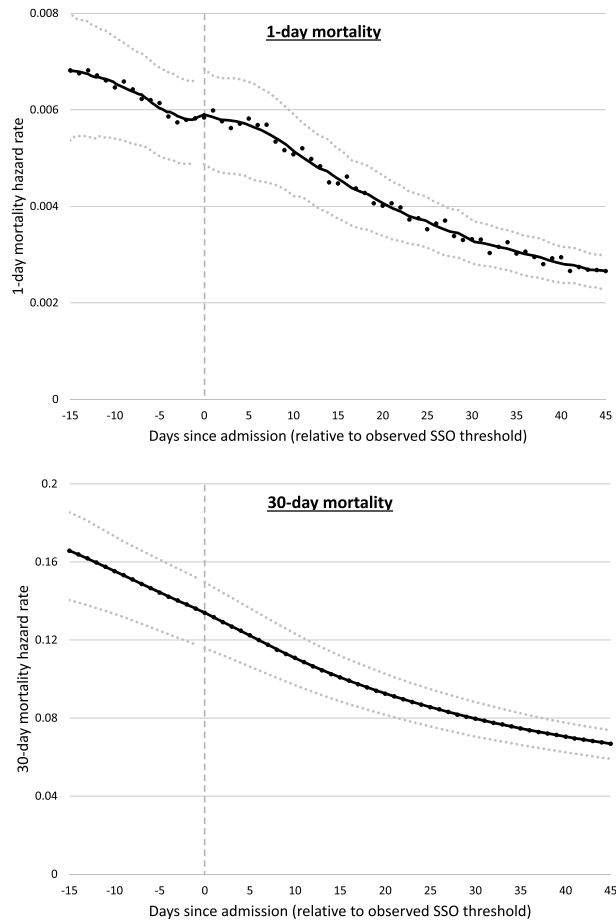
FIGURE C.1.—RD plots of mortality hazard by days since LTCH admission. Figure shows RD plots of the effects on mortality by days since admission in the PPS period. Mortality includes any mortality, whether it occurs within the LTCH or after discharge. Days since admission is normalized by counting days relative to the SSO threshold. The top panel shows the 1-day mortality hazard, defined as the fraction of living individuals who die in the next day; the bottom panel shows the 30-day mortality hazard, defined as the fraction of living individuals who die in the next 30 days. The dots show the underlying data averaged by day. The solid lines show local linear regressions, constructed using a 3-day bandwidth and a uniform kernel. The dashed lines show 95% confidence intervals, constructed by bootstrapping with replacement over DRGs.

mortality hazard. The distributions of placebo estimates imply a *p*-value of 0.796 for the 1-day mortality hazard and a *p*-value of 0.757 for the 30-day mortality hazard.

APPENDIX D: NON-DYNAMIC MODELS

In Sections 4.1 and 5.1, we briefly discussed two non-dynamic alternative models of LTCH behavior: (i) a myopic model in which the LTCHs make discharge decisions on a daily basis but do not internalize the dynamic implications of their decisions and (ii) a completely static model in which LTCHs commit to discharge decisions at the time of LTCH admission. Below, we present these models in more detail and argue these models perform poorly relative to our baseline dynamic model.
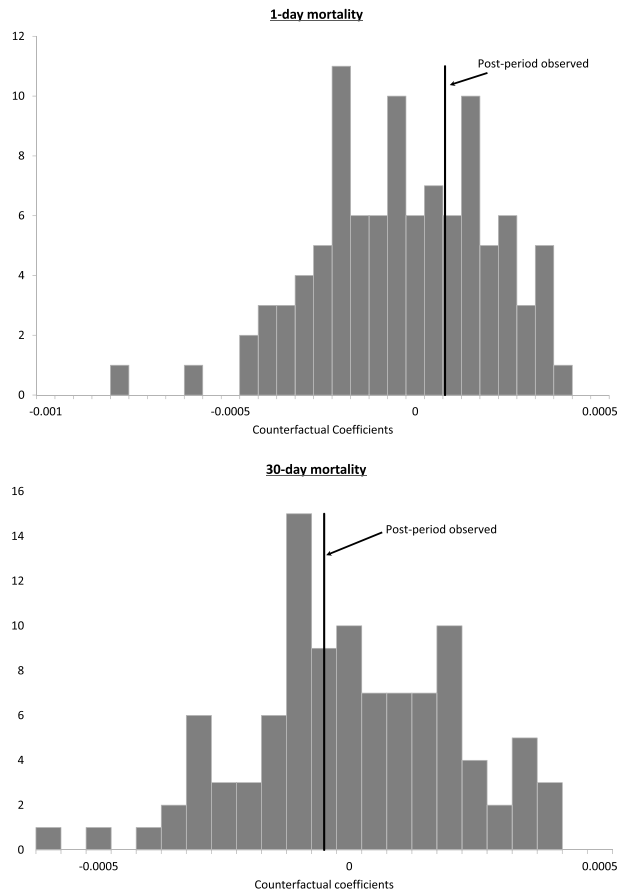
FIGURE C.2.—Perturbation tests for the estimated mortality effect. Figure shows perturbation tests for the mortality effect described in Appendix C. The top panel reports the estimated 1-day mortality effect and the bottom panel reports the estimated 30-day mortality effect from estimating equation (A3) with a bandwidth of 3, but replacing the dummy variable for being to the right of the SSO threshold with a dummy variable for being to the right of a placebo threshold; see Appendix C for more details. The figure also shows where the actual estimated effect falls within this range of placebo estimates.

D.1. *Myopic Model*

The first non-dynamic model we consider is one in which LTCHs are myopic and thus do not internalize the effects of their behavior on future periods. The objective function for the LTCH is to choose a location optimally ignoring any dynamic consideration. That is, as in the baseline model, we assume that hospital l's value from discharging a patient to location j is given by

$$u^j(h) = v^j(h) + \sigma_\epsilon \epsilon_{ijt} \quad \text{for } j = d, u, \tag{D.1}$$

and hospital l's value from retaining the patient is given by

$$u^l(h) = p(t) - c + v^l(h) + \sigma_\epsilon \epsilon_{ilt}. \tag{D.2}$$

Unlike the baseline model, the LTCH does not take into account any dynamic implication, such as the option value associated with discharging the patient later.

TABLE D.I

PARAMETER ESTIMATES FROM THE MYOPIC VERSION OF THE MODEL[a]

	Parameter		Parameter
Health process during pre-PPS:		Preferences:	
μ_0	7.53	c (000s)	1.24
σ_0	2.77	ν_{1u} (000s)	−380.45
μ	0.76	ν_{0d} (000s)	−4.48
ρ	0.94	ν_{1d} (000s)	0.21
σ	1.96	σ_ε (00s)	5.14
Health process during PPS:			
μ_0	7.67		
σ_0	2.86		
μ	1.84		
ρ	0.79		
σ	2.29		

[a]Table presents point estimates of the parameters for the myopic version of the model, which we describe in Appendix D.1.

Normalizing $v^l(h) = 0$ (as we do in the baseline model) yields the discharge probabilities

$$\Pr(j|h, t) = \frac{\exp[v^j(h)/\sigma_\epsilon]}{\exp[(p(t) - c)/\sigma_\epsilon] + \exp[v^u(h)/\sigma_\epsilon] + \exp[v^d(h)/\sigma_\epsilon]}. \quad \text{(D.3)}$$

We parameterize the health process in the same way as in the baseline model, and (as in the baseline model) assume that $v^j(h)$ is linear in h. We estimate the model by GMM to match the standard pre- and post-PPO period moments.

Figure D.2 shows the fit of the myopic model and Table D.I shows the parameter estimates. Not surprisingly, the myopic model struggles to fit the discharge patterns around the jump in payments. On the upstream margin, shown in the top left panel of Figure D.2, the myopic model predicts that LTCHs are relatively unresponsive to the change in incentives. While actual discharges nearly double at the SSO threshold, the myopic model predicts almost no jump in behavior. The reason is intuitive: To fit the relatively high level of discharges on the day before the jump in payments ($t = -1$), myopic LTCHs are estimated to be insensitive to price. But this results in the model being unable to match the jump at the SSO threshold. The dynamic model can better fit the data because it allows for a "second dimension" of LTCH behavior. Because LTCHs are dynamically building up a stock of patients who are marginal to the change in the incentives, they can exhibit a jump in discharges at the SSO threshold ($t = 0$) while still being inelastic enough to discharge sick patients on the day prior to the jump ($t = -1$).

Conversely, on the downstream margin, shown in the middle left panel of Figure D.2, LTCHs are too responsive to the jump in payments. While the myopic model matches the level of discharges after the jump in payments ($t = 0$), it predicts that the LTCH will discharge nobody on the day before the payment increase ($t = -1$), which undershoots the actual discharge rate of approximately 1%. In a sense, the myopic model is making the "opposite" mistake relative to the upstream margin. To fit the sharp jump at the threshold, the myopic LTCHs are estimated to be extremely price sensitive. But this price sensitivity makes the model unable to match the positive share of discharges at the SSO threshold, where keeping the patient an additional day would result in a large payday. Because forward-looking LTCHs build up a stock of marginal patients, the dynamic model is able
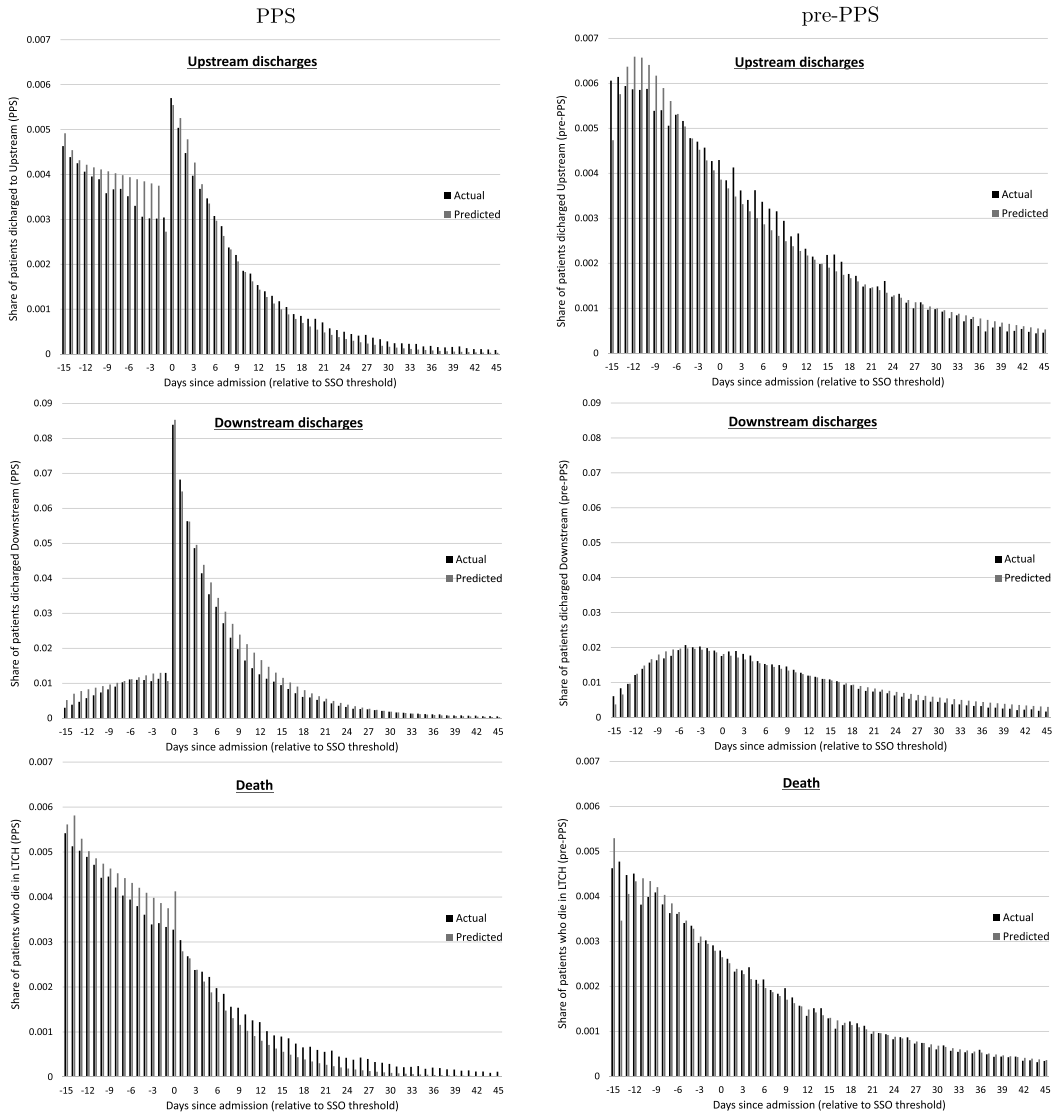
FIGURE D.1.—Model fit of the baseline model. Figure shows the moments we use for estimation, and how the model is able to fit them. Black bars in each panel represent the actual moments from the data, and the gray bars represent the predicted moments from the model estimates. The left three panels represent the PPS period, and the right three panels represent the pre-PPS period. The top panels show discharge rates upstream, the middle panels show discharge rates downstream, and the bottom panels show mortality rates (within the LTCH).

to match the jump in discharges while at the same time also discharging some patients on the day prior to the jump ($t = -1$).

D.2. *Static Model*

The second non-dynamic model we consider is a completely static model in which, at the time of admission, the LTCH commits to discharge the patient to a given location (ei-
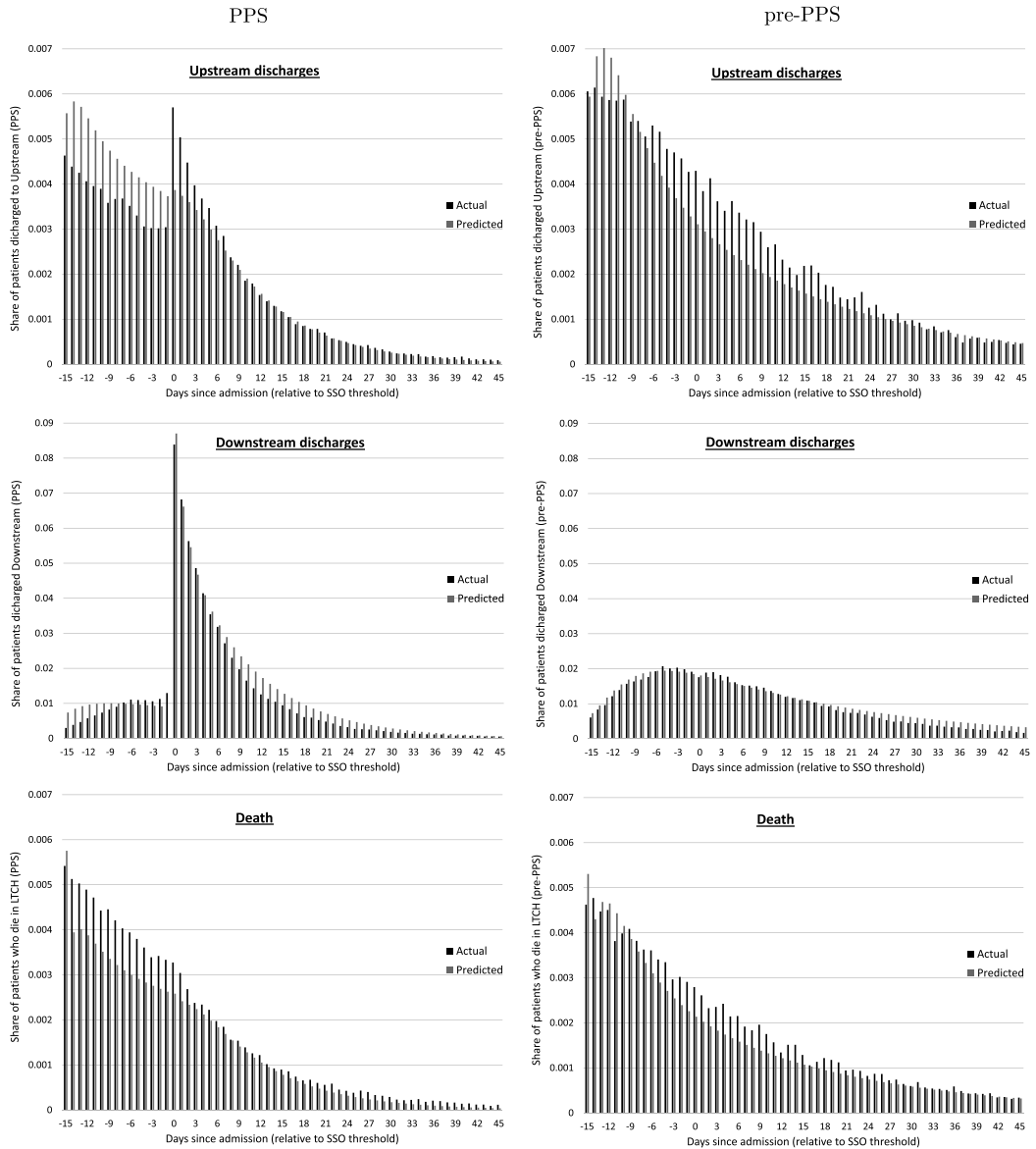
PPS

pre-PPS

FIGURE D.2.—Model fit of the myopic model. Figure is parallel to Figure D.1, and shows how the myopic model described in Appendix D.1 fits the data.

ther upstream or downstream) after a given number of days. As in the baseline model, we use $V^j(h)$ to denote the payoff to discharging the patient to location $j = u, d$ and $p(t) - c$ to denote net revenue. Since the LTCH commits to a discharge decision at admission, we assume that, at admission, the LTCH draws a separate logit error for each length of stay by location of discharge with scale parameter σ_ϵ. Normalizing the patient's utility at the

TABLE D.II

PARAMETER ESTIMATES FROM THE STATIC VERSION OF THE MODEL[a]

	Parameter		Parameter
Health process during pre-PPS:		Preferences:	
μ_0	15.35	c (000s)	1.77
σ_0	4.98	ν_{1u} (000s)	0.71
μ	3.33	ν_{0d} (000s)	−53.77
ρ	0.85	ν_{1d} (000s)	3.72
σ	4.67	σ_ε (00s)	73.38
Health process during PPS:			
μ_0	20.64		
σ_0	8.28		
μ	0.42		
ρ	0.99		
σ	3.07		

[a]Table presents point estimates of the parameters for the static version of the model, which we describe in Appendix D.2.

LTCH to zero yields the discharge probabilities

$$
\Pr(j, t|h_0)
$$

$$
= \frac{\exp\left[\left(\sum_{\tau=1}^{t} \delta^\tau S(\tau|h_0)(p(\tau) - c) + \delta^{t+1} S(t|h_0) \int V^j(h_t)\, dF(h_t|h_0)\right) \Big/ \sigma_\epsilon\right]}{\sum_{k=u,d}\sum_{s=-15,\ldots,+45} \exp\left[\left(\sum_{\tau=1}^{s} \delta^\tau S(\tau|h_0)(p(\tau) - c) + \delta^{s+1} S(s|h_0) \int V^k(h_s)\, dF(h_s|h_0)\right) \Big/ \sigma_\epsilon\right]},
$$

(D.4)

where $S(t|h_0)$ is the survival function, the probability that a patient with initial health h survives to period t. Within the exponents, the first term, $\sum_{\tau=1}^{t} \delta^\tau S(\tau|h_0)(p(\tau) - c)$, is the expected discounted net profits and the second term is the probability the patient survives to a given t multiplied by the payoff $V^l(h)$ from discharging them to location j at that date. The health process and V^j's are parameterized and have the same interpretation as in the baseline model. The model is estimated by GMM to match the standard moments in the pre- and post-PPO period.

Figure D.3 shows the fit of the static model and Table D.II shows the parameter estimates. The static model fits the moments quite well but has two weaknesses relative to the dynamic framework. First, because the LTCH does not condition on heterogeneous future health when it makes its discharge decisions, a much larger share of the heterogeneity loads on the logit error. In particular, the scale term on the logit error is more than 10 times larger in the static model than in the baseline model ($7,337 versus $597). From a modeling perspective, we think it is more desirable for the heterogeneity to load on the health process, which has a clearer economic interpretation. Second, because the discharge locations are largely determined by the logit draws, rather than the evaluation of health, the static model does not capture the relationship between discharge location and observed health. Recall that Figure 4 shows that patients discharged upstream exhibited much larger post-discharge payments than patients discharged downstream. Indeed, these patterns were an important motivation for our "vertical" discharge model.
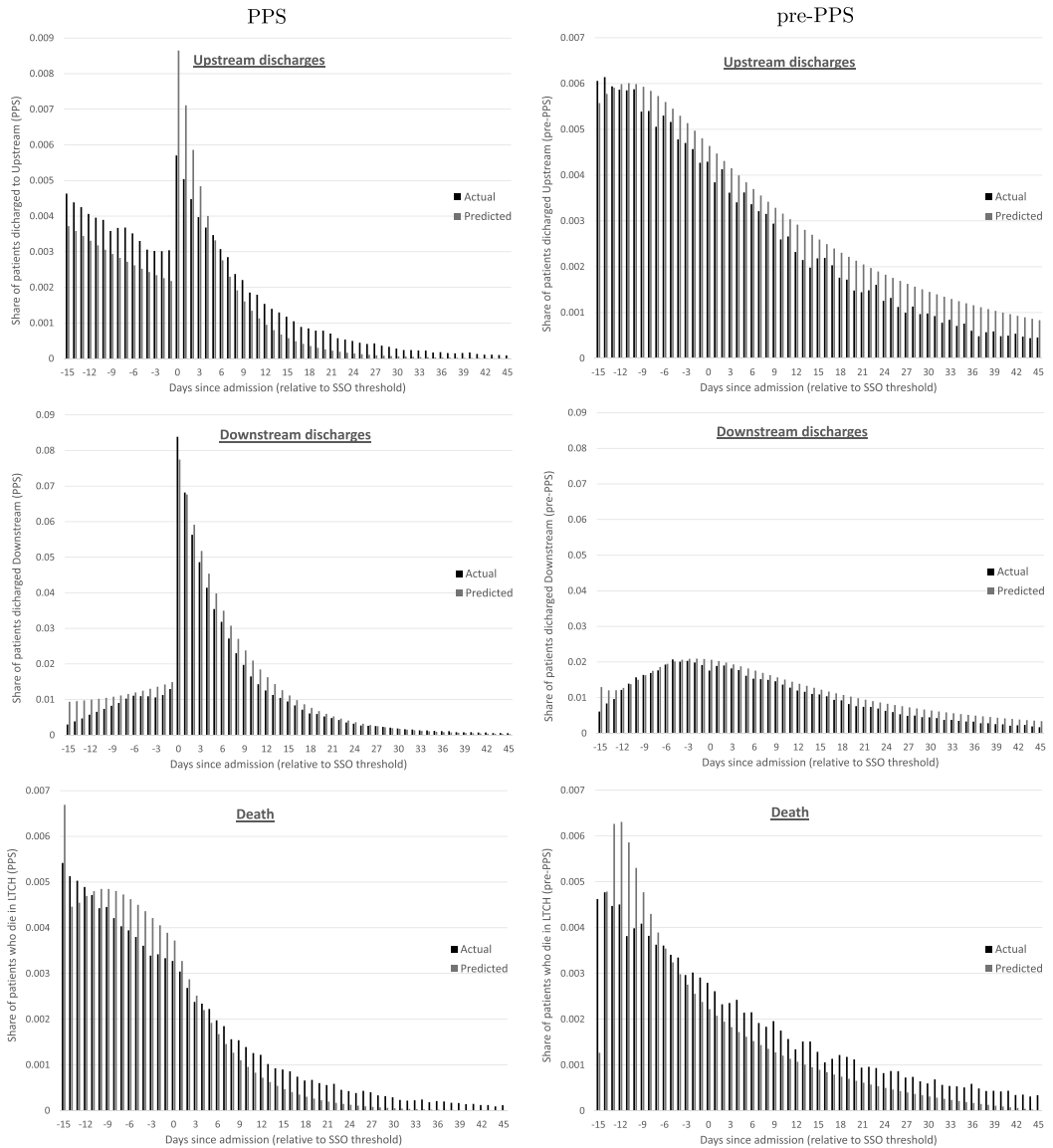
FIGURE D.3.—Model fit of the static model. Figure is parallel to Figure D.1, and shows how the static model described in Appendix D.2 fits the data.

Figure D.4 shows predicted 30-day post-discharge mortality for both the baseline (top panel) and static models (bottom panel) under the assumption that health outside of the LTCH evolves according to the same health process we estimated for patients that remain at the LTCH. Because in the static model LTCHs do not adjust their decisions based on the evolution of health, discharge decisions are largely based on the logit draw, and over longer time horizons, there is very little difference in the projected 30-day post-discharge mortality by location of discharge (upstream versus downstream). While we do not model the health process outside of the LTCH and therefore do not attempt to perfectly fit post-discharge mortality within the baseline model, our baseline model also predicts large
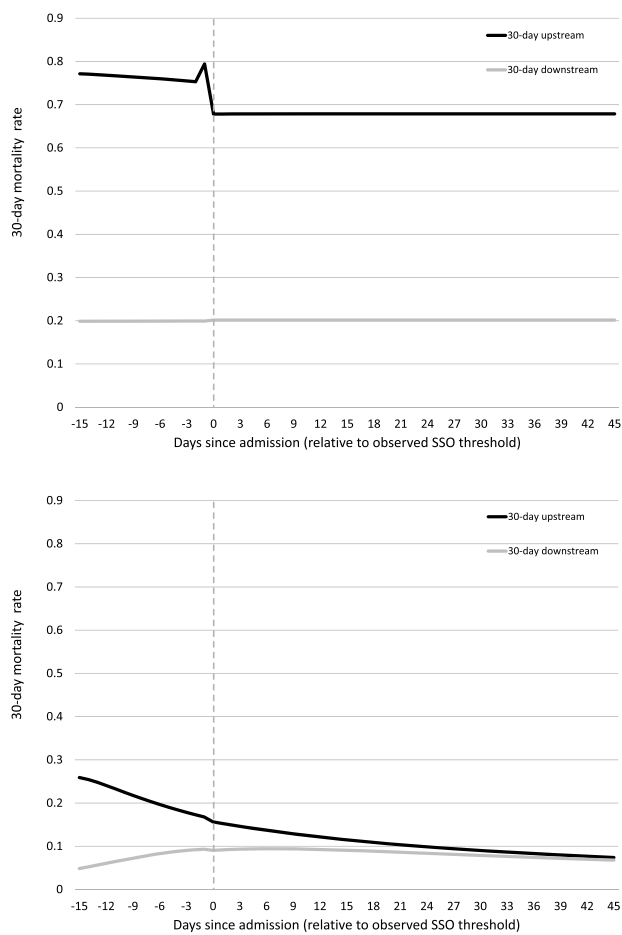
FIGURE D.4.—Discharge patterns implied by the static model. Figure shows the 30-day mortality rate for upstream and downstream dischargees. The top panel presents the results for the baseline model (which trace closely the optimal discharge policy function shown in Figure 7 of the main text), while the bottom panel shows the pattern implied by the parameter estimates of the static model (described in Appendix D.2).

differences in 30-day post-discharge mortality on the upstream and downstream margins, which is much more consistent with this basic feature of our economic environment.

APPENDIX E: IDENTIFICATION

In Section 4.4, we provided intuition for how the variation in our data allows us to separately identify the parameters in our model. In this appendix, we provide some additional details. We first present some quantitative exercises that link perturbations of model parameters to changes in the moments that are used for estimation; the results from this exercise are generally consistent with the overall intuition provided in Section 4.4 of the main text. We then focus on the preference parameters, which are the most critical for our counterfactual exercises, and show how, conditional on patient health, the variation around the jump in payments due to the SSO threshold allows us to separately identify σ_ϵ, c, and the v's.

E.1. *The Impact of Changes in Parameter Values on the Estimation Moments*

To provide more evidence on the mapping between data and parameters, in this section we conduct a perturbation exercise where we adjust each of the parameters on a one-by-one basis and measure the response of the predicted moments we use for estimation.

Figure E.1 provides a few specific examples of the output from this analysis. The left panels show the absolute change in discharges upstream, downstream, and to death by
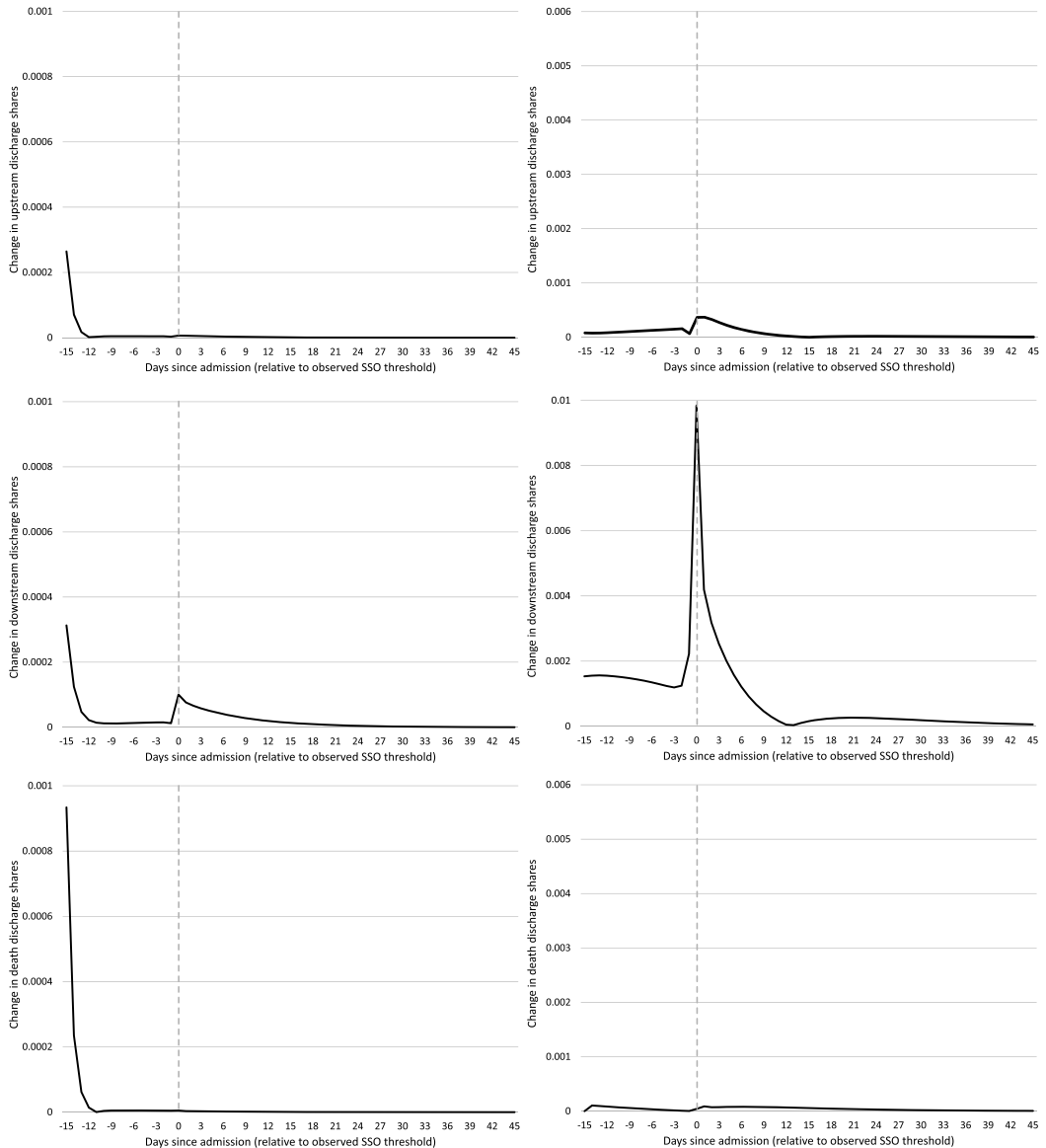

FIGURE E.1.—The impact of perturbing selected parameters on discharge patterns. Figure presents selected perturbation exercises, which are described in Appendix E and summarized in Table E.I. The left panels present the impact of increasing the parameter μ_0 by one standard error on each of the moments used for estimation (discharge patterns by day upstream (top), downstream (middle), and to death (bottom)). The right panel presents a similar exercise for the parameter σ_ε.

TABLE E.I

SUMMARY OF PARAMETER PERTURBATION EXERCISE[a]

	Initial Health	Health Process	Preference
Upstream			
Early	0.443	5.652	0.673
Pre-threshold	0.049	5.166	0.706
Near-threshold	0.043	4.106	1.190
Post-threshold	0.005	0.609	0.172
Downstream			
Early	0.656	4.259	3.484
Pre-threshold	0.137	6.094	3.170
Near-threshold	0.478	13.035	2.466
Post-threshold	0.058	0.625	0.629
Death			
Early	2.368	3.567	0.279
Pre-threshold	0.053	3.870	0.305
Near-threshold	0.029	2.033	0.361
Post-threshold	0.003	0.276	0.060

[a]Table summarizes the perturbation exercise described in Appendix E. We start by perturbing each parameter of the model by a single standard error, and measure its impact on each estimation moment; Figure E.1 presents two selected examples. We then summarize this exercise by averaging the impact across parameters in a given group and aggregating over groups of days. Parameters are grouped into initial health (μ_0 and σ_0), health process (μ, σ, and ρ), and preferences parameters (everything else).

length of stay where we increase the mean of the initial health distribution (μ_0) by its standard error (0.14). The figures show that the effect of mean initial health is concentrated in the first few days of the stay and mainly affects discharges to death. The right panels show the absolute change in discharges upstream, downstream, and to death by length of stay where we increase the scale parameter on the logit errors (σ_ϵ) by five times its standard error (135). Perturbing the scale parameter has the largest effects around the SSO threshold on the upstream and downstream margins.

In Table E.I, we attempt to summarize the output from this analysis across all of the parameters, discharge margins, and lengths of stay. To ease the presentation, we segment the parameters into three groups: initial health parameters (μ_0, σ_0), non-initial health parameters (μ, σ, ρ), and preference and cost parameters (v_{1u}, v_{0d}, v_{1d}, σ_ϵ, c). For each discharge destination (upstream, downstream, death), we also divide the post-PPS time period into early ($t = -15, \ldots, -11$), pre-threshold ($t = -10, \ldots, -3$), near-threshold ($t = -2, \ldots, +10$), and post-threshold ($t = +11, \ldots, +45$). A cell in the table shows the simple average of the absolute change in the discharge share (to the corresponding destination) over a set of the corresponding parameters and days.

Column (1) of Table E.I shows that the initial health parameters are primarily affecting discharges in the early days, and particularly strongly related to discharges to death in the first few days of the stay (consistent with the example shown in Figure E.1). The link between the health parameters and discharge to death is not surprising. There is a one-for-one mapping between a patient's health and their probability of death. The fact that this connection is concentrated in the first few days after admission suggests that the initial health parameters are largely pinned down by the moments over the first few days, and that these parameters are not strongly affecting discharge (and mortality) patterns in later days of the stay.

Column (2) shows that, as may be expected, perturbing the non-initial health parameters generates fairly broad shifts in the discharge moments, both across the different locations (upstream, downstream, death) and different lengths of stay. This is a direct consequence of health being an important state variable that governs behavior across all of the discharge margins and time periods. Finally, column (3) shows that the preference and costs parameters are most strongly connected to the downstream moments, except well after the SSO threshold ($t = +11, \ldots, +45$) where the connection is weaker.

Overall, the table is consistent with our intuitive discussion in Section 4.4 of the main text, and loosely supports a "triangular" identification argument, whereby the initial mortality and discharge patterns identify the initial health parameters, the mortality and upstream patterns during the rest of the LTCH stay identify the health evolution process (non-initial health parameters), and conditional on these, the downstream discharge patterns before and around the SSO threshold identify the preference parameters.

E.2. *Separately Identifying the Scale, Cost, and Preference Parameters*

Our main counterfactual exercises analyze changes in LTCH payments, making it important to pay particular attention to the way we identify the non-health parameters, which are the ones that are most directly affecting the response of the LTCH discharge policy to changes in payments. In this section, we therefore "zoom in" on this group of parameters—σ_ϵ, c, and the v's—and provide more detail on how the variation around the jump allows us to separately identify each of them separately.

In Section 4.4 of the main text, we argue that the jump in payments at the SSO threshold allows us to identify the scale parameter (σ_ϵ) from the cost parameter (c) and preference parameters (the v's) while the costs and preference parameters are separately identified by the differential discharge patterns elsewhere during the LTCH stay. Here, we provide support for these arguments by showing the effect of perturbing these parameters on the choice-specific payoffs. To simplify the exposition, we focus on the downstream margin and hold fixed patient health at a given value h, thus reducing the set of v's to a single parameter. That is, discharges on the downstream margin are characterized by the scale parameter σ_ϵ on the logit errors, the cost parameter c, and the $V^d(h)$ value of discharging the patient downstream. Figure E.2 shows the net payoffs to discharging downstream relative to keeping the patients at the LTCH ($V^d(h) - V^l(h)$) for $h = 10.5$, which is approximately the health index of the marginal patient discharged downstream under the estimated parameters (see Figure 7 in the main text). The solid lines show the net payoffs from discharge at the estimated parameters. The dashed lines show the net payoffs from discharge when we increase the parameter values.

The top panel shows the effect of increasing the scale parameter σ_ϵ by 50% of its estimated value (from 5.97 to 8.96). Intuitively, because we normalize the coefficient on profits ($p(t) - c$) to 1 in the LTCH's objective function, σ_ϵ can be thought of as the inverse "profits sensitivity" of LTCH behavior. Therefore, prior to the SSO threshold, increasing σ_ϵ reduces the option value of retaining the patient until the jump in payments because the LTCH places less weight on the financial value of the jump in payments, thus raising the value of discharging the patient. After the SSO threshold, profits are negative, so increasing σ_ϵ makes retaining the patient not as bad, thus lowering the value of discharging the patient. As a result, σ_ϵ can be thought of as modulating the change in net payoffs at the SSO threshold, with a higher value for σ_ϵ resulting in a smaller change in discharges at the jump in payments conditional on patient health.

The middle panel shows the effects of increasing cost c by 25% of its estimated value (from 1091 to 1363). Increasing the cost parameter c has a negative impact on the LTCH's
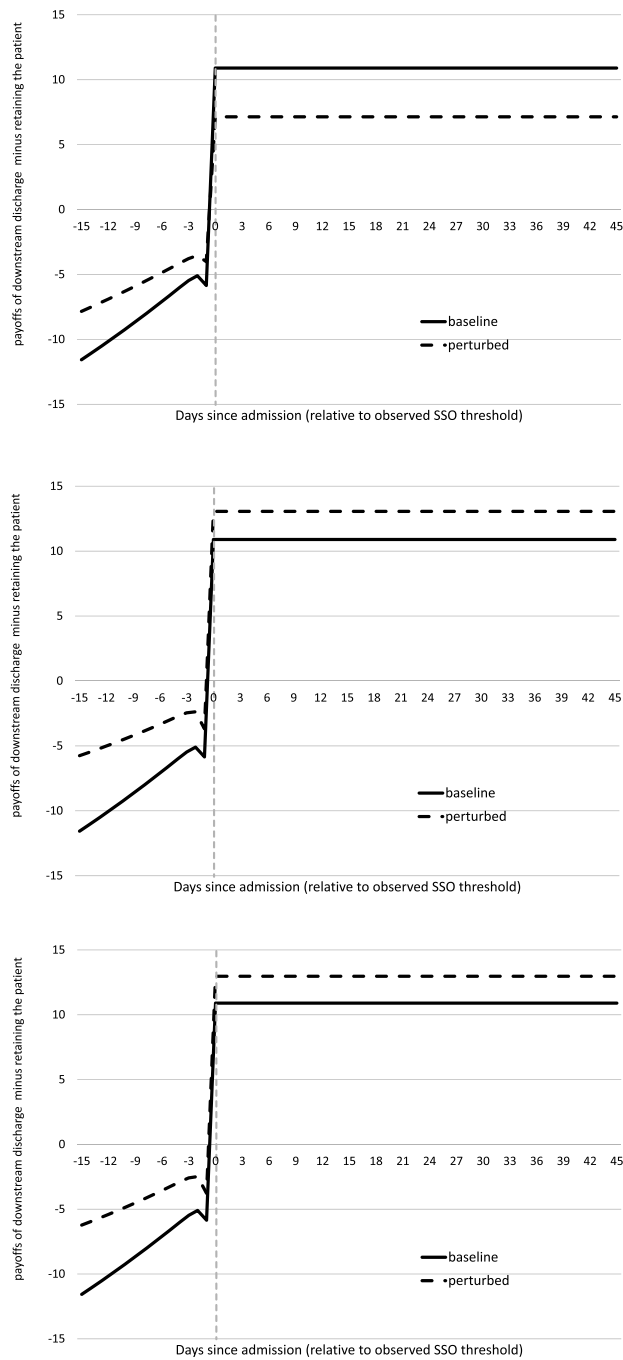
FIGURE E.2.—The impact of perturbing selected parameters on continuation values. Figure shows the impact of increasing each preference parameter on the continuation value of downstream discharge (relative to retaining the patient). The top panel shows this exercise for σ_ε, the middle panel for c, and the bottom panel for $V^d(h)$ (we show values for $h = 10.5$, which is approximately the health index of the marginal patient discharged downstream under the estimated parameters).

value from retaining a patient at the LTCH both before and after the SSO threshold, and the extent of this impact is proportional to the expected length of stay of the patient. The shift in the value of discharge is relatively larger in the first few days of the hospital stay because, at that point, the LTCH expects to keep the patient for a longer number of days (in order to receive the large lump-sum payment at the SSO threshold). The shift is uniform after the SSO threshold because the problem is stationary.

The bottom panel shows the effect of raising the value of downstream discharge $V^d(h)$ for a patient with $h = 10.5$ from -10.8 to 32.4. Raising $V^d(h)$ directly increases the value of discharging the patient downstream. Since most patients who are retained at the LTCH are eventually discharged downstream, raising $V^d(h)$ also increases the continuation value of retaining the patient at the LTCH and discharging them downstream in a later period. The indirect effect is smaller than the direct effect because of discounting and the probability that the patient is discharged upstream or to death instead of downstream at a later date. Thus, there is a net increase in the payoff to discharging the patient, but unlike the effect of costs, which was relatively larger in the first few days of the stay, the effect of increasing v_{0d} is fairly constant over time because discounting is minimal and death probability is low for patients who are close to the downstream margin.

Given the results of these perturbation exercises, it is now easier to see how the model is identified. The scale parameter is separately identified from the costs and $V^d(h)$ because it modulates the size of the shift in net payoffs at the SSO threshold, while the costs and $V^d(h)$ parameters are mostly affected by the level of discharge rates. They are separately identified from each other because of the differential movement in the first few days of the LTCH stay.

APPENDIX F: ROBUSTNESS

In our baseline model, we made a number of parametric assumptions. In order to assess the sensitivity of our main results to these assumptions, Table F.I reports the main results (from Table V) from a subset of the alternative specifications that we examined. The results appear to be qualitatively robust.

In our first alternative specification, we relax the assumption that the health process is stationary by allowing the autocorrelation parameter ρ to vary with length of stay according to $\rho = \rho_0 + \rho_1 \ln(t + 1)$, where the time index is defined such that the patient is admitted on date $t = 0$.[26] While in the pre-PPS period ρ_1 is very close to zero, in the PPS period the estimate is slightly negative, $\rho_1 = -0.005$, which is consistent with health becoming less stable over the course of the stay. However, as shown in Panel B of Table F.I, enriching the specification in this manner has virtually no effect on the counterfactuals. We also specified other models of health processes, such as a random walk and a random walk with a drift, but the ability of these models to fit the data was much worse than our baseline specification.

The second specification reported in Table F.I fits the model using only the PPS moments. The counterfactuals, shown in Panel C of Table F.I, are very similar to the baseline estimates, suggesting that the sharp jump in payments at the SSO threshold, relative to over-time variation from the implementation of LTCH-PPS, is the key driver of the results. The limited importance of the over-time variation presumably stems from the fact

---

[26]We assume that the correlation parameter becomes fixed after 45 days at a value of $\rho = \rho_0 + \rho_1 \ln(46)$ so that the dynamic programming problem becomes stationary for $t > 45$, allowing us to solve the $t > 45$ problem by value function iteration and earlier periods by backwards induction.

TABLE F.I

Robustness to Alternative Specifications[a]

	Observed Schedule	Higher Rate per Day	Lower Cap	Lowest Medicare Payment Within "LTCH Preferred" Schedules
A. Baseline specification				
LTCH Total Payments	27,953	28,316	16,024	26,313
LTCH Total Profits	6,518	8,991	-407	6,876
LTCH Average LOS	19.3	17.4	14.8	17.5
Total Medicare Payments	47,796	47,877	35,399	45,691
B. Alternative specification #1: time-varying health process				
LTCH Total Payments	27,953	28,324	16,043	26,320
LTCH Total Profits	5,153	7,762	-1,460	5,640
LTCH Average LOS	19.3	17.4	14.8	17.5
Total Medicare Payments	47,819	47,916	35,471	45,726
C. Alternative specification #2: post-PPS moments only				
LTCH Total Payments	28,101	28,545	16,506	26,648
LTCH Total Profits	10,446	12,454	2,466	10,421
LTCH Average LOS	19.8	18.0	15.7	18.2
Total Medicare Payments	47,969	48,322	36,211	46,225

[a]Table reports the main results from Table V under two alternative specifications of the model. Panel A reports results from the baseline specification, which corresponds to the numbers that are already reported in Table V. Panel B repeats the analysis, but we allow the AR(1) health process to vary over time by allowing the serial correlation parameter ρ to change linearly with the natural logarithm of days since LTCH admission. Panel C re-estimates the model using only data from the post-PPS period.

that we allow distinct health process parameters in each period, thereby soaking up much of the over-time variation.

REFERENCES

3M Health Information Systems (2015): "Inpatient Medicare Reimbursement Reference Guide GRP-017". [2199]

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2014): "Finite Population Causal Standard Errors," Mimeo. [2204]

American Hospital Association (AHA) (2010): "TrendWatch: Maximizing the Value of Post-Acute Care." Available at http://www.aha.org/research/reports/tw/10nov-tw-postacute.pdf. [2166]

Arcidiacono, P., and P. B. Ellickson (2011): "Practical Methods for Estimation of Dynamic Discrete Choice Models," *Annual Review of Economics*, 3, 363–394.[2186]

Bajari, P., H. Hong, M. Park, and R. Town (2017): "Estimating Price Sensitivity of Economic Agents Using Discontinuity in Nonlinear Contracts," *Quantitative Economics*, 8, 397–433.[2165]

Boards of Trustees for Medicare (2002): "The 2002 Annual Report of the Boards of Trustees of the Federal Hospital Insurance Trust Fund and the Federal Supplementary Medical Insurance Trust Fund". [2165]

——— (2014): "The 2014 Annual Report of the Boards of Trustees of the Federal Hospital Insurance Trust Fund and the Federal Supplementary Medical Insurance Trust Fund". [2165]

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011): "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence From Danish Tax Records," *Quarterly Journal of Economics*, 126, 749–804.[2164,2165]

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014): "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 104, 2593–2632.[2179]

Clemens, J., and J. D. Gottlieb (2014): "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104, 1320–1349.[2164]

Cohen, S. B., and W. Yu (2012): "The Concentration and Persistence in the Level of Health Expenditures Over Time: Estimates for the U.S. Population, 2008–2009," Statistical Brief #354, Agency for Healthcare Research and Quality. Rockville, MD. [2164]

CUTLER, D. (1995): "The Incidence of Adverse Medical Outcomes Under Prospective Payment," *Econometrica*, 63, 29–50.[2164]

CUTLER, D., AND R. ZECKHAUSER (2000): "Chapter 11—the Anatomy of Health Insurance," in *Handbook of Health Economics*, Vol. 1, 563–643. [2164]

DALTON, C. M., G. GOWRISANKARAN, AND R. TOWN (2017): "Salience, Myopia, and Complex Dynamic Incentives: Evidence From Medicare Part D," NBER Working Paper No. 21104. [2165]

EINAV, L., A. FINKELSTEIN, AND N. MAHONEY (2018): "Long-Term Care Hospitals: A Case Study in Waste". NBER Working Paper No. 24946. [2181,2192]

EINAV, L., A. FINKELSTEIN, AND P. SCHRIMPF (2015): "The Response of Drug Expenditure to Nonlinear Contract Design: Evidence From Medicare Part D," *Quarterly Journal of Economics*, 130, 841–899.[2165]

ELIASON, P., P. GRIECO, R. McDEVITT, AND J. ROBERTS (2018): "Strategic Patient Discharge: The Case of Long-Term Care Hospitals," *American Economic Review* (forthcoming). [2164,2174,2178,2186,2194]

HO, K., AND A. PAKES (2014): "Hospital Choices, Hospital Prices and Financial Incentives to Physicians," *American Economic Review*, 104, 3841–3884.[2164]

HU, Y., AND M. SHUM (2012): "Nonparametric Identification of Dynamic Models With Unobserved State Variables," *Journal of Econometrics*, 171, 32–44.[2186]

HU, Y., M. SHUM, W. TAN, AND R. XIAO (2015): "A Simple Estimator for Dynamic Models With Serially Correlated Unobservables," http://people.hss.caltech.edu/~mshum/papers/disestim.pdf. [2186]

KIM, Y., E. KLEERUP, P. GANZ, N. PONCE, K. LORENZ, AND J. NEEDLEMAN (2015): "Medicare Payment Policy Creates Incentives for Long-Term Care Hospitals to Time Discharges for Maximum Reimbursement," *Health Affairs*, 34, 907–915.[2164,2174,2199]

LIU, K., C. BASEGGIO, D. WISSOKER, S. MAXWELL, J. HALEY, AND S. LONG (2001): "Long-Term Care Hospitals Under Medicare: Facility- Level Characteristics," *Health Care Financing Review*, 23, 1–18.[2167, 2168]

MANOLI, D., AND A. WEBER (2016): "Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions," *American Economic Journal: Economic Policy*, 8, 160–182.[2165]

MEDICARE PAYMENT ADVISORY COMMISSION (MEDPAC) (2004): "A Data Book: Health Care Spending and the Medicare Program." [2162,2168]

———— (2014): "March 2014 Report to the Congress: Medicare Payment Policy," Chapter 11. [2199,2202]

———— (2015a): A Data Book: Health Care Spending and the Medicare Program. [2162,2165]

———— (2015b): "Report to the Congress: Medicare Payment Policy, Chapter 7: Medicare's Post-Acute Care: Trends and Ways to Rationalize Payments". [2162,2165]

———— (2016): "Report to the Congress: Medicare Payment Policy. Chapter 10: Long-Term Care Hospital Services." [2164,2174]

NEWHOUSE, J., A. GARBER, R. GRAHAM, M. McCOY, M. MANCHER, AND A. KIBRIA (2013): "Variation in Health Care Spending: Target Decision Making, not Geography," Institute of Medicine, http://www.nationalacademies.org/hmd/~/media/Files/Report%20Files/2013/Geographic-Variation2/geovariation_rb.pdf. [2162,2165]

PAKES, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54, 755–784.[2186]

ROSENBAUM, P. R. (1984): "Conditional Permutation Tests and the Propensity Score in Observational Studies," *Journal of the American Statistical Association*, 79, 565–574.[2204]

———— (2002): "Covariance Adjustment in Randomized Experiments and Observational Studies," *Statistical Science*, 17, 286–327.[2204]

TAUCHEN, G. (1986): "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177–181.[2186]

WEAVER, C., A. W. MATHEWS, AND T. McGINTY (2015): "Hospital Discharges Rise at Lucrative Times," *The Wall Street Journal*, February 17, 2015. [2164,2174]