POMS

Original Article

Production and Operations Management
2025, Vol. 34(1) 45–59

DOI: 10.1177/10591478241276134
journals.sagepub.com/home/pao
Sage

Impact of Capacity Strain on the Health Status of Patients Discharged From an Intensive Care Unit

Song-Hee Kim[1], Edieal Pinker[2] and Joan Rimar[3]

Abstract

Intensive care units (ICUs) play a vital role in hospitals and often experience capacity strain. Using data from a large teaching hospital, we empirically examine the effect of ICU capacity strain on patient discharge. We find that capacity strain has no effect on the health status of discharged patients and that capacity strain leads to a shorter length-of-stay in the ICU, but the magnitude of the effect is less than half a day in the most extreme scenarios. In post-hoc analyses, we find evidence the discharge process is expedited by starting discharges earlier in the day when capacity is strained, demonstrating that more efficient discharges can free up space without compromising patient health. In addition, we find a decrease in patient admissions during periods of capacity strain. To our knowledge, this is the first study to examine the direct relationship between ICU capacity strain and patient health at discharge, and it builds on the existing literature that has examined the effects of ICU capacity strain on proxies for patient health at discharge. We advocate that hospitals use outcome measures based on patient status at discharge to accurately assess the impact of capacity strain on quality of care. In addition, our work highlights the need for tailored research approaches and management practices for different types of ICUs, as generalizing the impact of ICU capacity strain across different settings may lead to inaccuracies.

Keywords

Empirical Operations Management, Healthcare Delivery, Intensive Care Unit, Capacity Strain, Patient Outcome

Date received 9 February 2022; accepted 27 July 2024 after two revisions

Handling Editor: Sergei Savin

1 Introduction

Intensive care units (ICUs) are inpatient units that provide the highest level of care for critically ill patients. ICUs have stochastic patient arrivals and stochastic patient healing processes, making it inevitable that there will be periods when ICU utilization is very high, or even when the immediate demand for ICU beds exceeds capacity. This situation is not unique to ICUs and occurs in many types of service systems, but ICUs have some unique features that make it particularly important to have a good understanding of the effects of their capacity strain. ICUs deal with the most critically ill patients and are expensive, so the stakes are high. The providers who work in the ICU, doctors and nurses, are highly trained and have considerable discretion in how they deliver care and manage the unit, so how they adapt to capacity strain and its impact on them is also important. The hospital must set ICU capacity in such a way that the costs of capacity and the costs of capacity strain are optimally balanced.

There are three main phenomena that ICU capacity strain can generate, leading to poorer quality of care and health outcomes. First, patient admission to the ICU can be rejected or delayed; second, the length-of-stay (LOS) of ICU patients can be shortened to make room for new patients; and third, healthcare providers can be overworked. The first two phenomena can, in theory, have a direct impact on patient health outcomes. The third phenomenon, overwork, can also have direct effects on patient health outcomes, but also longer-term effects on the healthcare workforce. In this paper, we look for empirical

[1]SNU Business School, Seoul National University, Seoul, South Korea
[2]Yale School of Management, Yale University, New Haven, CT, USA
[3]New Haven, USA

Corresponding author:
Song-Hee Kim, SNU Business School, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, South Korea.
Email: songheekim@snu.ac.kr

evidence of the second phenomenon. That is, we want to determine if patients who are in ICUs when the unit's capacity is strained are discharged sooner than they would otherwise have been, and if so, whether these patients suffer poorer outcomes in some way.

A natural mental model of the operation of an ICU that we use in this study and that is implied in the related literature is as follows: Each day a patient stays in an ICU, the patient's health status changes by a random amount. When a patient's health status reaches a certain threshold (note that this threshold may depend on patient characteristics such as age, diagnosis, and pre-admission condition), the patient is ready to be discharged to another less intensive unit in the hospital. We will refer to such patients as "natural discharges." On a daily basis, ICU physicians identify these patients for discharge. Throughout the day, random new demand for ICU beds arises, and if there are not enough beds available to meet this new demand, the ICU physicians may discharge a patient who is not a natural discharge to make room for a newly arriving patient. We will refer to such a patient as a "demand-driven discharge."

Previous research has found evidence of demand-driven discharges. For example, Anderson et al. (2011) examine daily discharge rates across all ICUs in a single tertiary care hospital and find that discharge rates increase on days with high census and more scheduled surgeries. Similarly, in their study of a single 18-bed surgical ICU, KC and Terwiesch (2012) find that patients discharged when the sum of census and planned admissions is high have shorter ICU LOS. Wagner et al. (2013), analyzing data from 155 ICUs, report that capacity strain is correlated with a reduction in LOS. Notably, Long and Mathews (2018), using data from two medical ICUs, decompose ICU LOS into service time—when patients are being treated and stabilized—and time-to-transfer—when patients are waiting to be discharged. They find that while service time is unaffected by census, time-to-transfer is shorter during periods of high census, suggesting that demand-driven discharges may primarily affect the non-value-adding portion of ICU LOS.

The evidence on the impact of these demand-driven discharges on health outcomes is more mixed. Chrusch et al. (2009), studying a 10-bed medical ICU and an eight-bed surgical ICU in a single hospital, find that census is associated with an increased likelihood of early death or ICU readmission, which they conjecture are markers of demand-driven discharge. KC and Terwiesch (2012) find that earlier discharges in turn increase the likelihood that ICU readmission. Wagner et al. (2013) finds only "slightly greater odds of being readmitted to the ICU" and no evidence of negative outcomes in terms of inpatient mortality. An important limitation of all these studies is that they do not measure the *direct* relationship between capacity strain and the health status of patients when they are discharged from such ICUs. That is, are patients more acutely ill when they leave from an ICU with capacity strain?

In this paper, we report on our study of this question using patient-level data from a medical ICU at a large urban U.S. teaching hospital. To do so, we utilize a dynamic measure of patient acuity called the Rothman index (RI; Rothman et al., 2013). The novelty of the RI is that patients' RIs are automatically calculated from the electronic medical record data and updated frequently throughout the patients' hospital stays. In our data, patients' RIs were updated every hour while they were in the ICU, allowing us to track patients' health status very close to the time of their ICU discharge and to examine its direct relationship with capacity strain.

To our knowledge, this study is the first to examine the *direct* relationship between ICU capacity strain and patient health status at discharge. Previous research has largely relied on *proxies* for patient health status at ICU discharge. For example, Chrusch et al. (2009) and KC and Terwiesch (2012) use *ICU readmission rates*. While reducing these rates is important, using these rates as the primary measure of the impact of ICU capacity strain on patient health at ICU discharge may be inaccurate for several reasons, including: (a) extensive interventions in downstream units may prevent or, conversely, necessitate readmissions; that is, the prevention or cause of readmissions is not what happened in the ICU but what happened outside the ICU; (b) some readmissions may not occur due to full ICU capacity; and (c) planned readmissions when beds become available do not necessarily reflect adverse outcomes.

Wagner et al. (2013) use *in-hospital mortality rates*. Similar to ICU readmission rates, in-hospital mortality rates are influenced by the operations of downstream units. In addition, (fortunately) only a small proportion of patients die in the hospital, and ICU capacity strain is likely to affect an even smaller proportion of them. KC and Terwiesch (2012) use *ICU LOS*, operating on the assumption that a longer ICU LOS, indicating more time for treatment and stabilization, correlates with better health outcomes at discharge. However, the relationship between LOS and health status is complicated because patients do not necessarily improve over time (Chalfin, 2005; Moitra et al., 2016). In contrast, we use each patient's health score trajectory in our data to extract the patient's health status *at ICU discharge* and examine how patient's health status at ICU discharge is affected by ICU capacity strain.

Another contribution we make to the existing literature is that we account for census and acuity of patients in the ICU when measuring ICU capacity strain. This differs from the existing literature, which has generally only accounted for census and ignored patient acuity. Previous studies could not account for patient acuity because of the lack of a dynamic patient acuity measure that tracks changes in patient condition. For example, KC and Terwiesch (2012), which is one of the first papers to document demand-driven discharge behavior in ICUs, write on page 61: "Given that the measures of severity (age, gender, and procedure type) are fixed during a hospital stay, and given that we do not observe time-varying health status of patients, our severity measure is fixed for a given patient's stay."

Some studies have attempted to incorporate patient acuity into their capacity strain measures, but have failed to capture changes in patient acuity over the course of an ICU stay. For example, Wagner et al. (2013) include patient acuity in their definition of ICU capacity strain. However, they average the acuity of patients in the ICU based on individual severity of illness scores calculated on the day of ICU admission. Similarly, Wilcox et al. (2020) include severity-weighted bed census as their measure of ICU capacity strain using individual severity of illness scores calculated during the first 24 hours after ICU admission. Because patients' conditions change during an ICU stay, our capacity strain measure based on each patient's current condition is expected to provide a more accurate representation of ICU capacity strain.

We find no evidence that ICU capacity strain affects the health status of discharged patients. In addition, although we find that ICU capacity strain leads to shorter ICU LOS—consistent with the existing literature identifying demand-driven discharge behavior—the magnitude of the effect is less than half a day even in the most extreme scenarios. We investigate the possible mechanisms that allow this to occur despite capacity strain. Consistent with the findings in Long and Mathews (2018) that only time-to-transfer and not service time is affected by congestion, we find evidence that the discharge process is accelerated by starting discharges earlier in the day when there is capacity strain. This shows that more efficient discharges can free up space without compromising patient health. We also find that fewer patients are admitted to the ICU when ICU capacity is strained. Finally, we find support for the assumption that care providers select the healthiest patients for discharge. Taken together, our results suggest that ICU staff in our study ICU use their discretion to manage capacity strain in a way that protects patient safety and mitigates potential risks to patients. This is reassuring, but it does not eliminate the long-term costs of provider burnout when capacity strain persists for long periods of time.

2 Empirical Setting and Data

Our data are from a 36-bed medical ICU at a large urban teaching hospital in the United States. We collected hospitalization data for every patient who received care in the study ICU from January 2013 to March 2014. The data included 3,644 ICU visits from 3,274 hospitalizations during the 15-month period. For each hospitalization, we had detailed patient-level data, including health score trajectory throughout the hospitalization, age, gender, 31 comorbidity indicators—that is, indicators for existing conditions such as congestive heart failure, diabetes, renal failure, and liver disease generated from the International Classification of Diseases codes and the Diagnosis-Related Group classification (Elixhauser et al., 1998)—payer and discharge disposition. The data also included every inpatient unit visited by the patient during the hospitalization, along with unit admission and discharge dates and times.

2.1 Dynamic Health Score

A key feature of our data is the availability of a dynamic health score called the RI (Rothman et al., 2013), which tracks each patient's evolving health status. The RI is a composite score, updated regularly from the electronic medical record, based on changes in 26 clinical measures, including vital signs, cardiac rhythms, laboratory test results, and nursing assessments.[1] We refer the reader to Rothman et al. (2013) for details on the development and validation of the RI as a valid measure of patient health status in the hospital setting.

The RI is designed to be used for any inpatient (i.e., medical or surgical patients, including critical care patients) with any diagnosis. Lower RI scores indicate poorer condition. The RI has a theoretical range of −91 to 100, but the majority of patients fall within the range of 0 to 100. A RI of 65 is the average acuity level of patients discharged to a skilled nursing facility. A RI of 40 corresponds to a modified early warning score of 4, which is considered "a critical score for activating an escalation pathway for intervention or transfer to a higher level of care" (Rothman et al., 2013). Following this interpretation, we define a patient as *acute* if her RI is below 40.

Studies have shown strong associations between RI and longer-term patient outcomes such as 24-hour mortality, 1-year mortality, and discharge disposition (Rothman et al., 2013; Finlay et al., 2014); 30-day hospital readmission rates (Bradley et al., 2013); post-operative complications for colorectal surgery (Tepas et al., 2013); unplanned ICU transfers (Danesh et al., 2012); unplanned surgical ICU readmissions (Piper et al., 2014); and unplanned ICU readmissions within seven days (Stahel et al., 2024). For example, Bradley et al. (2013) find that patients with RI scores lower than 70 (70–79) at hospital discharge had 2.65 (2.40) times higher odds of 30-day unplanned hospital readmission than those with RI scores >90. As another example, using data from a tertiary academic center, Stahel et al. (2024) find that lower RI scores at ICU discharge correlate with a higher risk of unplanned ICU readmission within seven days; the authors conclude that RI is a sensitive predictor of unanticipated ICU readmissions and assert that RI should be considered as a real-time objective measure for predicting a safe downgrade from the ICU to a lower level of care.

To illustrate such associations for our data, in Table B1 in the Appendix we examine the relationship between the health scores at ICU admission and ICU discharge measured by RI and three patient outcome measures: hospital LOS after ICU discharge, ICU readmission, and in-hospital mortality. In columns (1), (3), and (5), in addition to other patient characteristics, we include patient health score at ICU admission as an explanatory variable. All three models show that health score at ICU admission is a statistically significant explanatory variable. In columns (2), (4), and (6), we add health score at ICU discharge in addition to health score at ICU admission. All three models again show that the health score at ICU discharge

Table 1. Summary statistics of variables of interest and control variables.

	Mean	Std. Dev.	Min	Median	Max	Corr(*, 1)	Corr(*, 2)	Corr(*, 3)	Corr(*, 4)
Variables of interest									
(1) $Census_i$	31.0	3.4	19.0	31.0	38.0	1.00			
(2) $PctAcute_i$	70.0	8.6	43.8	70.0	96.9	0.04	1.00		
(3) $ICULOS_i$ (days)	3.6	4.4	0.5	2.1	48.0	−0.02	0.01	1.00	
(4) $HealthScoreAtDischarge_i$	47.0	22.0	−24.9	46.6	96.8	−0.06**	−0.07***	−0.24***	1.00
Control variables									
(5) $HealthScoreAtAdmission_i$	41.3	21.9	−21.9	39.9	97.1	−0.02	−0.07***	−0.26***	0.66***
(6) Age_i	61.5	17.5	18.0	62.0	102.0	0.02	0.01	−0.03	−0.31***
(7) $Female_i$	0.45					−0.02	−0.02	−0.04	0.02
Payer									
(8) $Medicare_i$	0.46					−0.01	−0.01	−0.04	−0.16***
(10) $Medicaid_i$	0.24					0.00	−0.01	0.01	0.12***
(11) $ManagedCare_i$	0.19					0.02	0.00	0.05*	0.09***
(12) *Other*	0.01					−0.02	0.00	−0.03	0.07***
Unit before ICU									
(13) $OtherICU_i$	0.01					−0.01	0.00	0.12***	−0.05*
(14) $StepDownUnit_i$	0.08					−0.04	−0.02	0.11***	−0.03
(15) $GeneralWard_i$	0.29					0.02	−0.03	0.01	−0.15***
(16) $None/Procedure/Test_i$	0.10					−0.01	−0.03	0.05*	0.01
(17) $EmergencyRoom_i$	0.52					0.01	0.05*	−0.12***	0.15***

Note: ICU = intensive care unit; LOS = length-of-stay. The unit of observation is each ICU visit, indexed by *i*. $N = 2,589$ visits. For binary variables, only the mean is reported. Summary statistics of 31 Elixhauser comorbidity indicators (e.g., congestive heart failure, diabetes, renal failure, and liver disease indicators), ICU admission day of week, and ICU admission month indicators are not provided due to space limitations. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

is a statistically significant explanatory variable. Note that in two of the three models, the health score at ICU admission is no longer statistically significant, suggesting that the health score at ICU admission has no additional explanatory power when controlling for health score at ICU discharge.

2.2 Sample Selection

In examining the effect of capacity strain on LOS and patient health status, we attempt to eliminate possible confounding events by restricting our sample. First, to avoid censored estimation of variables, we restrict our study to the 365 days in the middle of the data period—from 1 March 2013 to 28 February 2014. There were 3,112 ICU visits during this one-year study period. We removed 36 ICU visits with missing health score trajectories and eight ICU visits with missing patient-level information, leaving 3,068 ICU visits. We then excluded 343 ICU visits where the patient died and 136 ICU visits where the ICU stay was <12 hours. Our final sample consists of 2,589 ICU visits.[2]

2.3 Variables

Table 1 provides the summary statistics of the variables of interest.

2.3.1 Measures of ICU Capacity Strain. To measure ICU capacity strain, we account for census and how acutely ill the patients in the ICU are. We let $Census_i$ denote the number of patients occupying ICU beds at the beginning of the day that patient i is discharged from the ICU. We let $PctAcute_i$ denote the percentage of acute patients among all patients in the unit not including the focal patient i at the beginning of the day that patient i is discharged from the ICU. We deliberately do not include patient i in the calculation of the percentage of acute patients because we want to examine the effect of patient i's environment on her patient outcome. Consequently, patients i and j discharged on the same day will have $Census_i = Census_j$, but $PctAcute_i$ and $PctAcute_j$ may be different.

2.3.2 Measures of Patient Outcomes. We let $ICULOS_i$ denote the ICU LOS of visit i. On average, patients stayed in the ICU for 3.6 days, and the standard deviation was 4.4 days. We let $HealthScoreAtDischarge_i$ denote the health score of patient i at ICU discharge. The average health score at ICU discharge was 47.0 with a standard deviation of 22.0, indicating a large variation in health scores at ICU discharge.

The last four columns in Table 1 show the correlation values between $Census_i$, $PctAcute_i$, $ICULOS_i$, and $HealthScoreAtDischarge_i$ as well as the correlation values between the variables of interest and the control variables. $ICULOS_i$ is negatively associated with $Census_i$ and positively associated with $PctAcute_i$, although the magnitudes of the correlations are small and not statistically significant (−0.02

and 0.01, respectively). $HealthScoreAtDischarge_i$ is negatively associated with both $Census_i$ and $PctAcute_i$ (−0.06 and −0.07, respectively, and both are statistically significant).

2.3.3 Control Variables. When estimating the effect of a treatment on a patient outcome, it is critical to control for factors that may affect the patient outcome (KC, 2018). Our data allow us to control for patient characteristics and seasonality factors that may affect ICU LOS and health score at ICU discharge. Specifically, we control for the health score at ICU admission of each stay i ($HealthScoreAtAdmission_i$), age, gender, payer indicators, indicators for the unit before ICU, indicators for 31 Elixhauser comorbidity, indicators for ICU admission day of week, and indicators for ICU admission month.

3 Effects of ICU Capacity Strain

3.1 Estimation Models

The unit of observation is an ICU visit, indexed by i. To examine the effect of capacity strain on ICU LOS, we use the following model to examine the effect of census and percentage of acute patients on $log(ICULOS_i)$, which is the logarithm of the LOS of visit i:

$$log(ICULOS_i) = \alpha_0 + \alpha_1 Census_i + \alpha_2 PctAcute_i + X_i'\theta + \epsilon_i. \tag{1}$$

Note that we take the logarithm of the LOS to account for its right-skewed nature. $Census_i$ is the number of patients occupying beds in the ICU at the beginning of the day patient i is discharged from the ICU. $PctAcute_i$ is the percentage of acute patients among all patients in the unit, not including the focal patient i, at the beginning of the day patient i is discharged from the ICU. Importantly, our data allow us to control for patient characteristics and seasonality factors that can potentially affect $log(ICULOS_i)$. We let X_i denote a vector of these control variables—a detailed description of the control variables is provided in Section 2.3.3.

Similarly, to examine the effect of capacity strain on health status at ICU discharge, we use the following model to examine the effect of census and percentage of acute patients on $HealthScoreAtDischarge_i$, which is the patient outcome of visit i measured by the patient's health score at ICU discharge:

$$HealthScoreAtDischarge_i = \beta_0 + \beta_1 Census_i + \beta_2 PctAcute_i + X_i'\lambda + \varepsilon_i. \tag{2}$$

The main coefficients of interest are α_1 and β_1, which capture the effects of the census, and α_2 and β_2, which capture the effects of the percentage of acute patients.

3.2 Main Results

Table 2 presents the estimation results. Columns (1) and (2) show the results of estimating equation (1) without and with the control variables, respectively. Columns (3) and (4) show the results of estimating equation (2) without and with the control variables, respectively.

Table 2. Effects of census and percentage of acute patients on ICU LOS and on health status at discharge.

	log(ICULOS)		*HealthScoreAtDischarge*	
	(1)	(2)	(3)	(4)
Census	−0.01*	−0.01*	−0.40**	−0.10
	(0.01)	(0.01)	(0.12)	(0.10)
PctAcute	−0.00	0.00	−0.17***	0.02
	(0.00)	(0.00)	(0.05)	(0.04)
HealthScoreAtAdmission		−0.01***		0.53***
		(0.00)		(0.02)
Age		−0.00*		−0.24***
		(0.00)		(0.02)
Female		0.02		0.37
		(0.03)		(0.65)
Payer				
ManagedMedicare		0.11*		0.88
		(0.05)		(1.02)
Medicaid		0.06		−0.87
		(0.05)		(0.98)
ManagedCare		0.11*		2.20*
		(0.05)		(0.94)
Other		−0.23		6.49*
		(0.12)		(2.69)
Unit before ICU				
StepDownUnit		−0.31		3.14
		(0.18)		(2.64)
GeneralWard		−0.61***		1.67
		(0.17)		(2.44)
None/Procedure/Test		−0.49**		1.76
		(0.18)		(2.57)
EmergencyRoom		−0.70***		5.75*
		(0.17)		(2.43)
Constant	1.22***	2.29***	71.55***	42.89***
	(0.20)	(0.30)	(5.17)	(5.31)
N	2589	2589	2589	2589
R^2	0.00	0.21	0.01	0.54

Note: ICU = intensive care unit; LOS = length-of-stay. Columns (1)–(4) are linear regression models estimated at ICU visit level with *log*(*ICULOS*) and *HealthScoreAtDischarge* as the dependent variables. Control variables not shown for columns (2) and (4) include 31 Elixhauser comorbidity indexes, ICU admission day of week indicator variables, and ICU admission month indicator variables. Robust standard errors are shown in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

3.2.1 Effects of Capacity Strain on ICU LOS. We begin by examining the effect of *Census* on *log*(*ICULOS*). Column (2) of Table 2 shows that the coefficient for *Census* is statistically significant ($p < 0.05$). The coefficient value of −0.01 indicates that one additional patient in the ICU can reduce the ICU LOS by approximately 1%. Put another way, a one standard

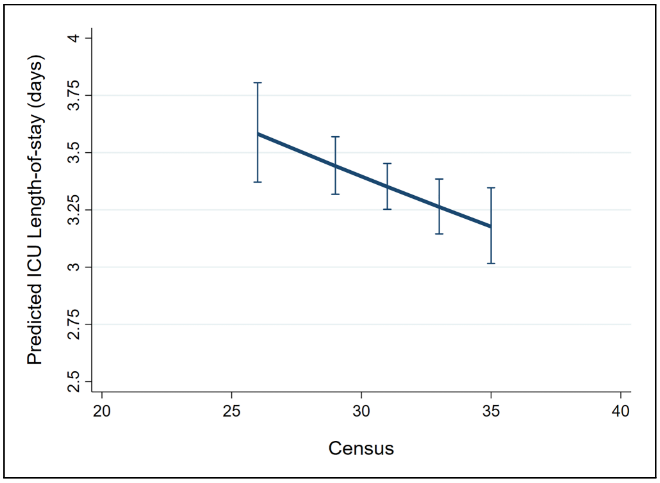

Figure 1. Predicted ICU LOS as a function of census. Note. ICU = intensive care unit; LOS = length-of-stay. Predicted *ICULOS* as a function of *Census* (Column (2) of Table 2). Error bars show the 95% confidence interval of predicted ICU LOS at the 10th, 25th, 50th, 75th, and 90th percentiles of ICU census.

deviation increase in the census (i.e., 3.4 more patients) results in a 3.4% decrease in ICU LOS.

In addition, for ease of interpretation, we illustrate the magnitude of this effect by plotting the predicted ICU LOS at the 10th, 25th, 50th, 75th, and 90th percentiles of *Census* in Figure 1. For example, if there are 35 patients (90th percentile value of *Census*) in the unit at the beginning of the day, as opposed to 26 patients (10th percentile of *Census*), we predict a 0.4 day or 9.6 hour decrease in ICU LOS. This means that even in the most extreme scenarios, the impact of ICU capacity strain on ICU LOS is likely to be less than half a day.

Next, we examine the effect of *PctAcute* on *log*(*ICULOS*). Column (2) of Table 2 shows that the coefficient for *PctAcute* is not statistically significant at the 0.05 significance level. To ensure that this null result is not due to insufficient data, we use equivalence tests (Wellek, 2002). An equivalence test reverses the question that is asked in a null hypothesis significance test. That is, instead of testing whether we can reject the null hypothesis of no effect of *PctAcute* (i.e., $H_0 : \alpha_2 = 0$), an equivalence test asks whether we can reject that $|\alpha_2|$ is equal to or greater than the smallest effect size of interest, Δ (i.e., $H_0 : |\alpha_2| \geq \Delta$ and $H_a : -\Delta < \alpha_2 < \Delta$). We find that we can reject the equivalence null hypothesis for any Δ that is $\geq$0.0036 at the 0.05 significance level. Since an effect size of 0.0036 is so small that it can be considered negligible—for example, $\alpha_2 = -0.0036$ translates to one percentage point increase in the percentage of acute patients reducing the ICU LOS by approximately 0.36%—we conclude that we have sufficient data to support the lack of an effect of *PctAcute* on *log*(*ICULOS*).

Lastly, we examine the effect of including the control variables in our estimation model for *log*(*ICULOS*). Column (1) shows the results of estimating equation (1) without the control variables. Compared to the results in column (2) in which the control variables are included, we find that there is no change in the coefficient values for *Census* and *PctAcute*.

In summary, we find that a patient's LOS is slightly shorter if the patient is discharged from an ICU with capacity strain than from an ICU without capacity strain. This effect is driven by census and not by the percentage of acute patients. However, the magnitude of the effect is small and is expected to be less than half a day even in the most extreme scenarios in our study ICU.

3.2.2 Effects of Capacity Strain on Health Status at ICU Discharge. We begin by examining the effect of *Census* on *HealthScoreAtDischarge*. Column (4) Table 2 shows that the coefficient for *Census* is not statistically significant at the 0.05 significance level. Using equivalence tests, we find that we can reject the equivalence null hypothesis for any Δ that is $\geq$0.27 at the 0.05 significance level. This also means that if the effect size of interest is smaller than 0.27, there is insufficient data to support either an effect or the lack of an effect between *Census* and *HealthScoreAtDischarge*. Note that $\beta_1 = -0.27$ translates to one additional patient in the ICU reducing patient health score at ICU discharge by 0.27. Given that *HealthScoreAtDischarge* has a large variation (i.e., it has a standard deviation of 22.0, a minimum of -24.9, and a maximum of 96.8), we consider an effect size <0.27 to be small and conclude that we have sufficient data to support the lack of an effect of *Census* on *HealthScoreAtDischarge*.

Next, we examine the effect of *PctAcute* on *HealthScoreAtDischarge*.[3] Column (4) Table 2 shows that the coefficient for *PctAcute* is also not statistically significant at the 0.05 significance level. Using equivalence tests, we find that we can reject the equivalence null hypothesis for any Δ that is $\geq$0.10 at the 0.05 significance level. Note that $\beta_2 = -0.10$ translates to a one percentage point increase in the percentage of acute patients reducing patient health score at ICU discharge by 0.10. As before, given the large variation in *HealthScoreAtDischarge*, we consider an effect size <0.10 to be small and conclude that we have sufficient data to support the lack of an effect of *PctAcute* on *HealthScoreAtDischarge*.

Lastly, we examine the effect of including the control variables in our estimation model for *HealthScoreAtDischarge*. Column (3) shows the results of estimating equation (2) without the control variables. Compared to the results in column (4) in which the control variables are included, we find that the coefficients for *Census* and *PctAcute* become statistically significant at the 0.05 significance level. This change highlights the importance of controlling for patient characteristics and seasonality factors in our model.

In summary, our findings indicate that a patient's health status at ICU discharge is not adversely affected by capacity strain. In other words, we find that capacity strain, as measured by census and percentage of acute patients does not affect patient health status at ICU discharge in our study ICU.

3.3 Robustness Checks

3.3.1 Sensitivity Analysis for the Null Effect of Capacity Strain on Health Status at ICU Discharge. We have estimated the causal effect of capacity strain on health status at ICU discharge using observational data. Our analysis relies on the assumption of "no-unobserved-confounding," a condition that, while common in observational studies, cannot be directly verified from the data (Cinelli and Hazlett, 2020). To address this, we use the methodology from Cinelli and Hazlett (2020) to perform sensitivity analyses to assess how our estimates might be influenced by potential unobserved confounding.

Figure 2 presents sensitivity contour plots that illustrate the potential impact of confounders on the point estimates of the coefficients for *Census* and *PctAcute* in column (4) of Table 2. The x-axis is the hypothetical fraction of the residual variation in *Census* or *PctAcute* explained by a confounder, while the y-axis shows the same for the dependent variable, *HealthScoreAtDischarge*. The triangles show the original unadjusted estimates, assuming no confounders. The contours show the adjusted estimates for hypothetical values of the sensitivity parameters, assuming that unobserved confounders make the estimates more negative, although the main insights remain consistent when positive adjustments are assumed.

As reference points, three diamond markers are included in each plot to illustrate the hypothetical influence of confounders as strong as patient health score at ICU admission, patient age, and month of ICU admission, respectively. The markers are colored red if the corresponding hypothetical estimates are statistically significant at the 5% level. The remaining observed control variables in our model, including 31 Elixhauser comorbidity conditions and admission day of the week, generate reference points that cluster around the origin. To avoid clutter, we have omitted these reference points from the plots.[4] The proximity of most observed control variables to the origin suggests that any unobserved confounders, if present, are likely to have a similar magnitude of influence, supporting the robustness of our estimates.

The white-filled diamonds show that even unobserved confounders of patient health status as strong as patient health score at ICU admission or patient age would not significantly alter our estimates. We believe that it is highly unlikely that there are other unobserved measures of patient health status that are as strong or stronger than patient health score at ICU admission or patient age, and conclude that our results are robust to unobserved measures of patient health status.

On the other hand, the red-filled diamonds show that unobserved confounders as strong as month of ICU admission could significantly alter the point estimates of the coefficients of *Census* and *PctAcute* to −0.81 and −0.27, respectively, both achieving statistical significance at the 5% level. The strength of the month of ICU admission variable in explaining capacity strain is consistent with the well-documented seasonal variation in ICU utilization (Garfield et al., 2001; Cheng et al., 2014). In our data, the average midnight census in February, the peak month, is 33.7 compared to 27.3 in August, the lowest month, a difference of 6.4 ($p < 0.001$). Similarly, the average percentage of acute patients at midnight varies significantly, from 78.2% in February to 68.6% in August, a difference of 9.6 percentage points ($p < 0.001$). Given these significant seasonal influences of month of ICU admission on *Census* and *PctAcute*, it is difficult to imagine that any other seasonal or operational unobserved confounding would be as strong or stronger than month of ICU admission.

In summary, although our results are subject to the usual caveats associated with observational studies, the results of

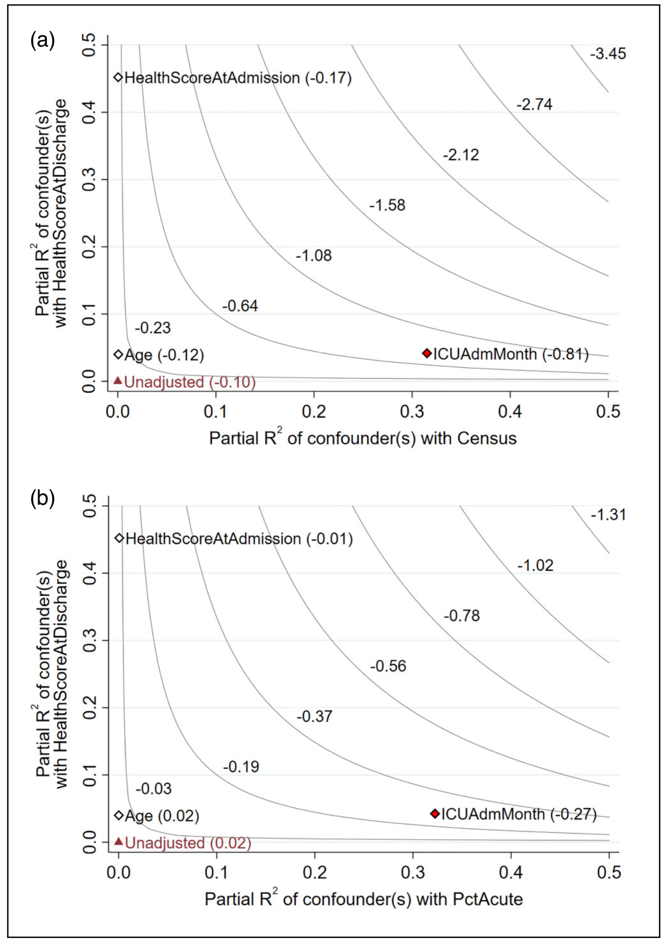

Figure 2. Sensitivity contour plots of the point estimates of the coefficients of *Census* and *PctAcute* in column (4) of Table 2.
(a) Sensitivity contour plot of the point estimate of the coefficient of *Census*; (b) sensitivity contour plot of the point estimate of the coefficient of *PctAcute*. Note. ICU = intensive care unit. The original unadjusted estimate, assuming no confounders, is indicated by a triangle. The contours show adjusted estimates for hypothetical values of the sensitivity parameters, assuming that unobserved confounders skew the estimates negatively. Three reference points, indicated by diamonds, show the influence of confounders as strong as patient health score at ICU admission, patient age, and month of ICU admission, respectively. They are colored red when the hypothetical estimate is significant at the 5% level.

Table 3. Effects of census and percentage of acute patients on hazard rate of ICU discharge .

	(1)	(2)	(3)	(4)	(5)
Census (time-varying)	0.00	0.00	−0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
PctAcute (time-varying)	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
HealthScoreAtAdmission			0.01***		−0.00
			(0.00)		(0.00)
HealthScore (time-varying)				0.02***	0.02***
				(0.00)	(0.00)
Age		0.00	0.00	0.01***	0.01***
		(0.00)	(0.00)	(0.00)	(0.00)
Female		0.01	0.02	0.02	0.02
		(0.04)	(0.04)	(0.04)	(0.04)
Payer					
ManagedMedicare		−0.08	−0.12	−0.09	−0.09
		(0.07)	(0.07)	(0.07)	(0.07)
Medicaid		−0.10	−0.10	−0.09	−0.09
		(0.06)	(0.06)	(0.06)	(0.06)
ManagedCare		−0.15*	−0.10	−0.13*	−0.13*
		(0.06)	(0.06)	(0.06)	(0.06)
Other		0.21	0.28	0.18	0.17
		(0.19)	(0.18)	(0.18)	(0.18)
Unit before ICU					
StepDownUnit		0.34	0.24	0.25	0.25
		(0.21)	(0.21)	(0.21)	(0.21)
GeneralWard		0.63**	0.50*	0.50*	0.51*
		(0.20)	(0.20)	(0.20)	(0.20)
None/Procedure/Test		0.50*	0.38	0.36	0.37
		(0.21)	(0.21)	(0.21)	(0.21)
EmergencyRoom		0.69***	0.59**	0.54**	0.54**
		(0.20)	(0.20)	(0.20)	(0.20)
N	9125	9125	9125	9125	9125
Pseudo R^2	0.00	0.01	0.01	0.02	0.02
Likelihood Ratio Test	0.783	<0.001	<0.001	<0.001	<0.001

Note: ICU = intensive care unit. Columns (1)–(5) are Cox proportional hazard models. Control variables not shown for columns (2)–(5) include 31 Elixhauser comorbidity indexes, ICU admission day of week indicator variables, ICU admission month indicator variables, and day of week indicator variables. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

the sensitivity analyses suggest that it is unlikely that there is unobserved confounding that would substantially alter our conclusions.

3.3.2 Cox Proportional Hazard Model for ICU LOS. Following KC and Terwiesch (2012), as a robustness check, we use the Cox proportional hazard model to examine the effect of ICU capacity strain on the hazard rate of a patient's discharge from the ICU. To be consistent with other analyses in the paper, we let the unit of time be each day. Note that using the Cox proportional hazard model allows us to take advantage of the fact that we have each patient's health score trajectory. That is, we can control for *PctAcute*, the percentage of acute patients, and let it change for each day the focal patient stays in the ICU. Similarly, we can control for the patient's evolving health score by controlling for the time-varying *HealthScore*.

Table 3 shows the estimation results. First, we find that the coefficients for the time-varying *Census* and *PctAcute* are not statistically significant at the 0.05 significance level. Recall that although we found evidence that a patient's LOS was shorter if the patient was discharged from an ICU with capacity strain than from an ICU without capacity strain in Section 3.2.1, the magnitude of the effect was small and less than half a day. Because our analysis models the daily hazard rate of a patient's discharge from the ICU, our estimation results are consistent with what we found in Section 3.2.1.

Importantly, we estimate the coefficient for the time-varying *HealthScore* to be 0.02 ($p < 0.001$), which corresponds to a hazard ratio of 1.02. In other words, we find that the hazard rate of a patient's discharge from the ICU increases by 2% for each unit increase in the patient's current health score. This result supports the validity of using the RI as a measure

Table 4. Effects of census and percentage of acute patients—alternate measures of the dependent variable.

	(1)	(2)	(3)	(4)
Census	−0.02	−0.14	−0.17	−0.04
	(0.09)	(0.10)	(0.10)	(0.10)
PctAcute	−0.04	0.04	0.04	0.01
	(0.04)	(0.04)	(0.04)	(0.04)
X (control variables)	Yes	Yes	Yes	Yes
N	2125	2409	2542	2584
R^2	0.08	0.55	0.56	0.54

Note: ICU = intensive care unit. Columns (1)–(4) are linear regression models estimated at the ICU visit level. For column (1), the dependent variable is *HealthScoreAt24hrAfterDischarge–HealthScoreAtDischarge*. Among the 2,589 ICU visits, we focus on 2,439 visits whose next unit is a step-down unit (264 visits) or a general ward (2,175 visits). Among the 2,439 visits, 314 visits were missing the health score measured 24 hours after ICU discharge (we used a time window of $\pm$ 3 hours; the results are similar when using different time windows). For column (2), the dependent variable is the health score measured close to the ICU discharge but before the discharge day. For column (3), the dependent variable is the health score measured close to the ICU discharge but before 6 am of the discharge day. For column (4), the dependent variable is the health score measured close to the ICU discharge but before 12 pm of the discharge day. For columns (2)–(4), the reductions in sample size are mostly due to patients who were discharged on the same day they were admitted and had their first health score measured in the ICU after the time cutoffs. For all columns, controls not shown include patient characteristics summarized in Table 1 as well as 31 Elixhauser comorbidity indexes, ICU admission day of week indicator variables, and ICU admission month indicator variables. Robust standard errors are shown in parentheses. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

of health status, as we do in this paper. In addition, by comparing the coefficients of *HealthScoreAtAdmission* in columns (3) and (5), we note that while the coefficient is statistically significant in column (3), it is no longer significant in column (5) at the 0.05 level. This change highlights the added value of including current health scores, which may provide more relevant information than the health score at ICU admission alone.

3.3.3 Alternate Measures of Patient Outcome at ICU Discharge. Given our findings in Section 3.2.2 that capacity strain, as measured by census and percentage of acute patients, has no effect on patient outcomes after controlling for patient characteristics and seasonality factors, one might wonder whether the effect of capacity strain on patient outcomes is reflected in the evolution of patient health status *after* ICU discharge. To explore this possibility, we examine the effect of capacity strain on the difference between the patient health score measured 24 hours after ICU discharge and the patient health score measured at ICU discharge. The regression results are shown in Table 4. We find no evidence that capacity strain affects the evolution of patient health status after ICU discharge.

Furthermore, we define $HealthScoreAtDischarge_i$ as the final health score recorded during visit I. Since ICU discharges occur at different times of the day, the timing of this measurement also varies. If there is a systematic variation in patients' health scores based on the time of day, this could affect our results. For example, if some patients typically have better health scores in the morning, those discharged in the morning and whose care may have been sped up may appear healthier at the time of discharge compared to their condition if they were discharged later in the day. To address this concern, we use alternative dependent variables that capture health scores at specific times relative to ICU discharge. Specifically, the results in columns (2) through (4) use health scores recorded on the day of discharge but before various time cutoffs: before the day of discharge (column (2)), before 6 am on the day of discharge (column (3)), and before 12 pm on the day of discharge (column (4)). Our results show robustness in all three models.

Table 5. Effects of census and percentage of acute patients with spline specifications.

	log(ICULOS)		*HealthScoreAtDischarge*	
	(1)	(2)	(3)	(4)
Census $\leq$ 34	−0.02*		−0.15	
	(0.01)		(0.12)	
Census > 34	0.01		0.27	
	(0.03)		(0.53)	
Census $\leq$ 35		−0.01*		−0.13
		(0.01)		(0.11)
Census > 35		0.00		0.53
		(0.04)		(0.85)
PctAcute $\leq$ 77.1	0.00		0.05	
	(0.00)		(0.05)	
PctAcute > 77.1	0.00		−0.09	
	(0.01)		(0.13)	
PctAcute $\leq$ 80.8		−0.00		0.03
		(0.00)		(0.05)
PctAcute > 80.8		0.00		−0.08
		(0.01)		(0.19)
X (control variables)	Yes	Yes	Yes	Yes
N	2589	2589	2589	2589
R^2	0.21	0.21	0.54	0.54

Note: ICU = intensive care unit; LOS = length-of-stay. Columns (1)–(4) are linear regression models estimated at the ICU visit level with *log(ICULOS)* and *HealthScoreAtDischarge* as the dependent variables. The knots used for *Census*, 34 and 35, are the 80th and 90th percentiles of *Census*, respectively. The knots used for *PctAcute*, 77.1 and 80.8, are the 80th and 90th percentiles of *PctAcute*, respectively. Controls not shown include patient characteristics summarized in Table 1 as well as 31 Elixhauser comorbidity indexes, ICU admission day of week indicator variables, and ICU admission month indicator variables. Robust standard errors are shown in parentheses. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

3.3.4 Spline Specifications for ICU Capacity Strain Measures. We used continuous specifications for *Census* and *PctAcute* in our main models. However, the effects of *Census* and *PctAcute* may change for different intervals of *Census* and *PctAcute*. To test this, we use spline specifications for *Census* and *PctAcute*.

The estimation results of the models that include *Census* and *PctAcute* as spline variables are presented in Table 5. Note that the knots used for *Census* were 34 and 35, which are the 80th and 90th percentiles of *Census*, respectively. The knots used for *PctAcute* were 77.1 and 80.8, which are the 80th and 90th percentiles of *PctAcute*, respectively. The results suggest that the effect of the census on ICU LOS is driven by when the census is <34 or 35, which are the 80th and 90th percentiles of the census. Overall, we find that our main results are robust to the use of these alternative specifications for *Census* and *PctAcute*.

4 Post-hoc Analysis to Understand Underlying Mechanisms for Our Findings

The results in Section 3.2 show that ICU capacity strain leads to shorter ICU LOS, but the magnitude of the effect is less than half a day. Most importantly, we find no evidence that ICU capacity strain negatively affects patient health at discharge. Overall, we find no evidence for aggressive demand-driven ICU discharge behavior associated with deterioration in patient care in our study ICU. As evidenced by the effect of ICU capacity strain on ICU LOS, there may be a modest level of demand-driven ICU discharges, but they occur in a manner that does not compromise the quality of care.

What might be the mechanism by which one could reduce the ICU LOS without compromising patient outcomes at discharge? The findings of Long and Mathews (2018) provide one possible explanation. Long and Mathews (2018) decompose ICU LOS into service time (when patients are treated and stabilized) and time-to-transfer (when patients are waiting to leave the ICU). They find that service time is not affected by the census, but that the less important time-to-transfer is accelerated when the census is high. In their setting, the average service time was 3.3 days and the average time-to-transfer was 15.1 hours. Our results in Section 3.2.1 show that the estimated effect of the high census on ICU LOS is not likely to exceed 10 hours in our study ICU (note that the average ICU LOS in our setting is 3.6 days). This suggests that it is plausible that the census reduces ICU LOS when the census is high, but in a way that does not compromise patient care.

Our data do not distinguish between service time and time-to-transfer, so we cannot directly test the mechanism identified by Long and Mathews (2018) in our setting. Instead, we examine the effect of capacity strain on the time of the first discharge of the day. Our premise is that providers have identified some patients as ready for discharge and can respond to capacity strain by starting their transfer process earlier than "usual." If we find that discharges start earlier when the ICU is capacity strained, this will provide more evidence that the acceleration occurs in the transfer process and not necessarily in the service process, because there is no reason to believe that speeding up the transfer of a patient to another unit by a few hours will affect the care that the patient receives.

Table 6. Effects of census and percentage of acute patients on time of first discharge of the day and number of daily admissions.

	FirstDischargeTime (1)	*Admissions* (2)
Census	−0.43***	−0.20***
	(0.07)	(0.04)
PctAcute	0.02	−0.00
	(0.03)	(0.02)
Day of week	Yes	Yes
Month	Yes	Yes
N	365	365
R^2	0.15	0.11

Note: Columns (1) and (2) are linear regression models estimated at day level. *FirstDischargeTime* is the dependent variable for column (1) and *Admissions* is the dependent variable for column (2). Robust standard errors are shown in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

We define $FirstDischargeTime_t$ as the time of the first discharge of day t. The average value of $FirstDischargeTime_t$ is 8.0 or 8 am with a standard deviation of 4.8. To examine the effect of capacity strain on $FirstDischargeTime_t$, we estimate the following model:

$$FirstDischargeTime_t = \eta_0 + \eta_1 Census_t + \eta_2 PctAcute_t + \rho_{DayOfWeek_t} + \omega_{Month_t} + u_t. \quad (3)$$

Here, $Census_t$ is the number of patients occupying beds in the ICU at the beginning of day t. $PctAcute_t$ is the percentage of acute patients among all patients in the unit at the beginning of day t. We control for day of week and month. We let $\rho_{DayOfWeek_t}$ denote the coefficients for day of week indicator variables, where $DayOfWeek_t$ is the day of week of day t. We let ω_{Month_t} denote the coefficients for month indicator variables, where $Month_t$ is the month of day t.

Column (1) in Table 6 shows the result of the estimation. It shows that for one more patient in the ICU, the first discharge of the day occurs 0.43 hours or 26 minutes earlier ($p < 0.001$). This result suggests that when the ICU is capacity strained, the acceleration of patient care occurs in the transfer process.

Another possible mechanism by which one could reduce the ICU LOS without compromising patient outcomes at discharge is for physicians to accurately identify the healthiest patients in the ICU for discharge, thereby minimizing the impact of shorter LOS on patient outcomes. We look for evidence of such a mechanism in our data by examining the relationship between patient health scores and ICU discharge. As there are many factors other than a patient's health status score that influence discharge decisions, health scores alone cannot fully explain discharge decisions. However, health scores should be able to show whether healthier patients are more likely to be selected for discharge than less healthy patients. For instance, how often is the patient with the highest health score—the most medically stable patient according to the RI—discharged from the ICU? Among the 365 days from

1 March 2013 to 28 February 2014 in our data, we find that the patient with the highest health score at the beginning of the day is discharged during the day on 204 days (56% of the days). The patient with the lowest health score—the least medically stable patient according to the RI—at the beginning of the day stays in the ICU for 266 days (73% of the days). Additionally, suppose we define a patient to have a rank of 0 if she has the lowest health score at the beginning of each day and a rank of 1 if she has the highest health score. For example, if there are six patients, then the patient with the second lowest health score has a rank of 0.2. We find that the mean median rank of patients who stay is 0.46, while the mean median rank of patients who are discharged is 0.64. The difference in mean median rank is statistically significant ($p < 0.001$). These results provide some evidence that physicians deliberately select healthier patients for discharge to minimize the impact of shorter lengths of stay on patient outcomes.

Next, if demand-driven discharges are made aggressively to make room for newly arriving patients, then capacity strain at the beginning of the day should have no effect on how many new patients are admitted that day (KC and Terwiesch, 2012). Since we find evidence that demand-driven discharges are not being made aggressively, the ICU must be using some other lever to adjust demand on the unit. In fact, Kim et al. (2015) show that a high level of ICU census lowers the likelihood of ICU admission. Thus, we hypothesize that the ICU will admit fewer patients when there is capacity strain.

We define $Admissions_t$ as the number of admissions to the ICU on day t. The average value of $Admissions_t$ is 8.6 with a standard deviation of 2.7. To examine the effect of capacity strain on $Admissions_t$, we estimate the model in (3) with $Admissions_t$ as the dependent variable. Column (2) in Table 6 shows the estimation results. It shows that for each additional patient in the ICU at the beginning of the day, the number of ICU admissions decreases by 0.2 patients ($p < 0.001$). In sum, our results suggest that reducing the number of admissions is one lever the ICU uses to adjust demand on the unit.

5 Discussion and Conclusions

5.1 Limitations

Our results should be interpreted with limitations in mind. First, we use the RI to measure patient health status at ICU discharge. While this health score is validated and widely trusted (Rothman et al., 2013), such as any health score, it is not without limitations and may have measurement errors. However, as long as the measurement errors do not introduce systematic bias that our control variables cannot control for, we believe our results remain reliable. Future research could complement our findings by using a different, dynamic health score system, which we anticipate will become increasingly available. Second, we were unable to control for nurse staffing levels due to data unavailability; although we believe this limitation should not alter our results because the study ICU strictly adhered to the 1-to-1 or 1-to-2 staffing ratio depending on patient condition, future work could complement our findings by confirming this.

5.2 Summary of Findings

In this study, we examine the direct relationship between ICU capacity strain and patient health status at the time of ICU discharge, extending the existing literature that has only been able to examine the indirect relationship using proxies for patient health status at discharge. We find that there is no effect of ICU capacity strain on the health status of discharged patients in the ICU we study. We also find that ICU capacity strain leads to shorter LOS in the ICU, but the magnitude of the effect is less than half a day. In other words, we do not find evidence of aggressive demand-driven discharge practices that compromise patient care. Instead, we find evidence that our study ICU manages the quality of care under capacity strain by implementing earlier discharges in the day, prioritizing the healthiest patients for discharge, and reducing admissions, thereby minimizing the need for and impact of demand-driven discharges.

5.3 Reconciling Our Findings With Previous Research and Managerial Implications

Although evidence on the impact of ICU capacity strain on patient health outcomes is mixed in the existing literature, there is an intuitive appeal to the view that such strain compromises the quality of care through aggressive, demand-driven discharge practices. Most notably, KC and Terwiesch (2012), an influential study of an 18-bed surgical ICU, show that capacity strain shortens ICU LOS and subsequently increases ICU readmissions.

We believe that the main reason for the difference between our result and that of KC and Terwiesch (2012) is the different characteristics of the two study ICUs. In fact, in Table B3 of the Appendix, we replicate their instrumental variable estimation approach and find that while ICU census does reduce LOS, this reduction in LOS does not increase readmissions in our setting. As suggested by Long and Mathews (2018), ICUs differ in many ways. Patient population, size, decision-making processes, and the role the ICU plays in the hospital all vary across institutions. KC and Terwiesch (2012) examine an 18-bed ICU, while we examine a larger, 36-bed ICU. Based on the queuing theory, we know that larger units can operate at higher utilization rates without having to discharge patients early. Another difference is that KC and Terwiesch (2012) examine a surgical (cardiac) ICU while we examine a medical ICU. Notably, although we find that capacity strain leads to fewer admissions, KC and Terwiesch (2012) find no correlation between ICU census and admissions. They note that the only lever that surgical ICUs can pull in the face of ICU capacity strain is early discharge of patients, since surgical case cancellations due to ICU congestion are rare. Similarly, Kim et al. (2015) find demand-driven discharge behavior in elective, surgical ICUs, but not in medical ICUs where patients are admitted through the emergency department.

These distinctions between ICUs are important because setting and managing capacity in a hospital are key areas where operations management can contribute to improving healthcare. Too much capacity is wasteful and too little capacity leads to frequent episodes of capacity strain with potentially negative impact on patient care. Given the stochastic nature of these systems, it is inevitable that they will sometimes be under strain. How well can they handle this strain?

Our work illustrates three major lessons. First, analyses of capacity in ICUs should take into account the levers the providers have to manage workload. Second, if a hospital wants to assess whether or not capacity strain is hurting care it should use outcome measures based on patient status at discharge. Third, it is possible to use modest shortening of ICU stays to reduce strain without hurting patients.

This third point is consistent with Wagner et al. (2013) and Long and Mathews (2018) but stands in contrast to the conclusions in KC and Terwiesch (2012) who found that for some categories of patients, the readmission to the ICU used more capacity than was saved by shortening their stays originally. Finding more evidence supporting the benefits of modest LOS shortening helps keep an important lever in the ICU manager's toolbox. However, there is still uncertainty about how generalizable these findings are given the heterogeneity across ICUs.

This highlights the risks of generalizing about the impact of ICU capacity strain across different ICU settings. It also highlights the need for tailored research approaches and management practices for different types of ICUs. Future studies should continue to examine how the impact of ICU capacity strain varies by ICU characteristics. In doing so, we believe that researchers and hospitals should consider using the increasingly available dynamic health status measures (Landro, 2015; Leenen et al., 2020; Starr et al., 2022), which will help pinpoint problems more accurately and benefit both the research community and hospital operations.

Appendix A. Inputs to the RI

A key feature of our data is the availability of a dynamic health score called the RI, which tracks each patient's evolving health status. The RI is a composite score, updated regularly from the electronic medical record, based on changes in 26 clinical measures, including vital signs, cardiac rhythms, laboratory test results, and nursing assessments. Example definitions of standards for nursing assessments, provided in Table 1 of Rothman et al. (2013), are as follows:

- Cardiac standard: Pulse regular, rate 60–100 BPM, skin warm and dry. Blood pressure less than 140/90 and no symptoms of hypotension.
- Food/nutrition standard: No difficulty with chewing, swallowing, or manual dexterity. Patient consuming >50% of daily diet ordered as observed or stated.
- Gastrointestinal standard: Abdomen soft and nontender. Bowel sounds present. No nausea or vomiting. Continent. Bowel pattern normal as observed or stated.
- Genitourinary standard: Voids without difficulty. Continent. Urine clear, yellow to amber as observed or stated. Urinary catheter patent if present.
- Musculoskeletal standard: Independently able to move all extremities and perform functional activities as observed or stated (includes assistive devices).
- Pain standard: Without pain or visual analog pain scale <4 or experiencing chronic pain that is managed effectively.
- Neurological standard: Alert, oriented to person, place, time, and situation. Speech is coherent.
- Peripheral/vascular standard: Extremities are normal or pink and warm. Peripheral pulses palpable. Capillary refill <3 s. No edema, numbness, or tingling.
- Psychosocial standard: Behavior appropriate to situation. Expressed concerns and fears being addressed. Adequate support system.
- Respiratory standard: Resp. 12–24/min at rest, quiet and regular. Bilateral breath sounds clear. Nail beds and mucous membranes pink. Sputum clear, if present.
- Safety/fall risk standard: Safety/fall risk factors not present. The patient is not a risk to self or others.
- Skin/tissue standard: Skin clean, dry, and intact with no reddened areas. The patient is alert, cooperative, and able to reposition self-independently. Braden scale >15.

For details on the development and validation of the RI, see Rothman et al. (2013).

Appendix B. Additional Tables

Table B1. Relationship between health scores and other patient outcome measures.

	Hospital LOS after discharge		ICU readmission		In-hospital mortality	
	(1)	(2)	(3)	(4)	(5)	(6)
HealthScoreAtAdmission	-0.11^{***}	-0.05	-0.01^{***}	-0.01	-0.04^{***}	-0.02^{***}
	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)
HealthScoreAtDischarge		-0.12^{***}		-0.02^{***}		-0.05^{***}
		(0.03)		(0.00)		(0.01)
N	2439	2439	2439	2439	2589	2589
R^2	0.16	0.17				
Pseudo R^2			0.18	0.19	0.25	0.32

Note: ICU = intensive care unit; LOS = length-of-stay. Columns (1) and (2) are linear regression models and columns (3)–(6) are logit regression models estimated at the ICU visit level. In columns (1)–(4), we focus on patients who are discharged from the ICU to step-down units or general wards. Controls not shown include patient characteristics summarized in Table 1 as well as 31 Elixhauser comorbidity indexes and ICU admission month indicator variables. Robust standard errorsare in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table B2. Effects of census and percentage of acute patients on ICU LOS and on health status at discharge. *PctAcute − Adm* is used instead of *PctAcute*.

	log(ICULOS)		*HealthScoreAtDischarge*	
	(1)	(2)	(3)	(4)
Census	-0.01^{*}	-0.01^{*}	-0.39^{**}	-0.09
	(0.01)	(0.01)	(0.13)	(0.10)
PctAcute − Adm	0.00	-0.00	-0.09	0.09^{*}
	(0.00)	(0.00)	(0.05)	(0.04)
X (control variables)	No	Yes	No	Yes
N	2589	2589	2589	2589
R^2	0.00	0.21	0.01	0.54

Note: ICU = intensive care unit; LOS = length-of-stay. Columns (1)–(4) are linear regression models estimated at the ICU visit level with *log*(*ICULOS*) and *HealthScoreAtDischarge* as the dependent variables. In defining *PctAcute − Adm*, a patient is labeled as acute if her health score at ICU admission is <40 and this label does not change throughout her ICU stay. Controls not shown include patient characteristics summarized in Table 1 as well as 31 Elixhauser comorbidity indexes, ICU admission day of week indicator variables, and ICU admission month indicator variables. Robust standard errors are shown in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table B3. Effects of early discharge on ICU readmission and in-hospital mortality rates.

	ICU readmission	In-hospital mortality
	(1)	(2)
Main		
log(*ICULOS*)	-0.31	-0.46
	(1.07)	(0.81)
HealthScoreAtAdmission	-0.01	-0.02^{***}
	(0.01)	(0.00)
Controls	Yes	Yes
First stage (*log*(*ICULOS*))		
Census	-0.01^{+}	-0.01^{*}
	(0.01)	(0.01)
HealthScoreAtAdmission	-0.01^{***}	-0.01^{***}
	(0.00)	(0.00)
Controls	Yes	Yes
N	2439	2589

Note: ICU = intensive care unit; LOS = length-of-stay. Following KC and Terwiesch (2012), we used *Census* as an instrument for endogenous regressor *log*(*ICULOS*) and used instrumental variable probit maximum likelihood to estimate the coefficients for *log*(*ICULOS*). In column (1), we focus on patients who are discharged from the ICU to step-down units or general wards. Controls not shown include patient characteristics summarized in Table 1 of the main manuscript as well as 31 Elixhauser comorbidity indexes, ICU admission day of week indicator variables, and ICU admission month indicator variables. Robust standard errors are shown in parentheses. $^{+}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Acknowledgments

This study was supported by the Institute of Management Research at Seoul National University. Edieal Pinker thanks Avraham Ohana for wise comments. The authors would also like to thank the editors and reviewers for their constructive and insightful feedback, which greatly improved this paper.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Song-Hee Kim https://orcid.org/0000-0002-3106-5726

Notes

1. Example definitions of standards for nursing assessments are provided in Appendix A.
2. We used all of our data from 3,644 ICU visits to construct the ICU capacity strain measures.
3. In Table B2 in the Appendix, we present the estimation results of models where $PctAcute - Adm$ is used instead of *PctAcute*. $PctAcute - Adm$ differs from *PctAcute* in that a patient is labeled as acute if her health score at ICU admission is <40 and this label does not change throughout her ICU stay. We find that the coefficient values for the percentage of acute patients change when $PctAcute - Adm$ is used instead of *PctAcute*, which highlights the value of our dynamic measure of patient health status.
4. The reference points generated by the remaining observed control variables in our model would all lie within a rectangle that extends horizontally from 0 to 0.0043 and vertically from 0 to 0.0153 in Figure 2(a) and a rectangle that extends horizontally from 0 to 0.0017 and vertically from 0 to 0.0153 in Figure 2(b).

References

Anderson D, Price C, Golden B, et al. (2011) Examining the discharge practices of surgeons at a large medical center. *Health Care Management Science* 14(4): 338–347.

Bradley E, Yakusheva O, Horwitz LI, et al. (2013) Identifying patients at increased risk for unplanned readmission. *Medical Care* 51(9): 761–766.

Chalfin DB (2005) Length of intensive care unit stay and patient outcome: The long and short of it all. *Critical Care Medicine* 33(9): 2119–2120.

Cheng KC, Lu CL, Chung YC, et al. (2014) ICU service in Taiwan. *Journal of Intensive Care* 2: 1–8.

Chrusch CA, Olafson KP, McMillan PM, et al. (2009) High occupancy increases the risk of early death or readmission after transfer from intensive care*. *Critical Care Medicine* 37(10): 2753–2758.

Cinelli C and Hazlett C (2020) Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82(1): 39–67.

Danesh V, Guerrier L and Jimenez E (2012) Proactive vs. reactive rapid response systems: Decreasing unplanned ICU transfers. *Critical Care Medicine* 40(12): 1–328.

Elixhauser A, Steiner C, Harris DR, et al. (1998) Comorbidity measures for use with administrative data. *Medical Care* 36(1): 8–27.

Finlay GD, Rothman MJ and Smith RA (2014) Measuring the modified early warning score and the Rothman index: Advantages of utilizing the electronic medical record in an early warning system. *Journal of Hospital Medicine* 9(2): 116–119.

Garfield M, Ridley S, Kong A, et al. (2001) Seasonal variation in admission rates to intensive care units. *Anaesthesia* 56(12): 1136–1140.

KC DS (2018) *Econometric methods*. In: Dai T and Tayur S (eds) *Handbook of Healthcare Analytics: Theoretical Minimum for Conducting 21st Century Research on Healthcare Operations*. Hoboken, NJ: John Wiley & Sons, 381–402.

KC DS and Terwiesch C (2012) An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing & Service Operations Management* 14(1): 50–65.

Kim SH, Chan CW, Olivares M, et al. (2015) ICU admission control: An empirical study of capacity allocation and its implication for patient outcomes. *Management Science* 61(1): 19–38.

Landro L (2015) Hospitals find new ways to monitor patients 24/7. *The Wall Street Journal*. Available at: https://www.wsj.com/articles/hospitals-find-new-ways-to-monitor-patients-24-7-1432560825 (accessed 29 July 2024).

Leenen JP, Leerentveld C, van Dijk JD, et al. (2020) Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: Systematic review. *Journal of Medical Internet Research* 22(6): e18636.

Long EF and Mathews KS (2018) The boarding patient: Effects of ICU and hospital occupancy surges on patient flow. *Production and Operations Management* 27(12): 2122–2143.

Moitra VK, Guerra C, Linde-Zwirble WT, et al. (2016) Relationship between ICU length of stay and long-term mortality for elderly ICU survivors. *Critical Care Medicine* 44(4): 655.

Piper GL, Kaplan LJ, Maung AA, et al. (2014) Using the Rothman index to predict early unplanned surgical intensive care unit readmissions. *Journal of Trauma and Acute Care Surgery* 77(1): 78–82.

Rothman MJ, Rothman SI and Beals J (2013) Development and validation of a continuous measure of patient condition using the electronic medical record. *Journal of Biomedical Informatics* 46(5): 837–848.

Stahel PF, Belk KW, McInnis SJ, et al. (2024) The Rothman index predicts unplanned readmissions to intensive care associated with increased mortality and hospital length of stay: A propensity-matched cohort study. *Patient Safety in Surgery* 18(1): 10.

Starr AJ, Julka M, Nethi A, et al. (2022) Parkland trauma index of mortality: Real-time predictive model for trauma patients. *Journal of Orthopaedic Trauma* 36(6): 280–286.

Tepas JJ, Rimar JM, Hsiao AL, et al. (2013) Automated analysis of electronic medical record data reflects the pathophysiology of operative complications. *Surgery* 154(4): 918–926.

Wagner J, Gabler NB, Ratcliffe SJ, et al. (2013) Outcomes among patients discharged from busy intensive care units. *Annals of Internal Medicine* 159(7): 447–455.

Wellek S (2002) *Testing Statistical Hypotheses of Equivalence*. 1st ed. New York: Chapman and Hall/CRC.

Wilcox ME, Harrison DA, Patel A, et al. (2020) Higher ICU capacity strain is associated with increased acute mortality in closed ICUs. *Critical Care Medicine* 48(5): 709–716.

How to cite this article

Kim S-H, Pinker E and Rimar J (2025) Impact of Capacity Strain on the Health Status of Patients Discharged from an Intensive Care Unit. *Production and Operations Management* 34(1): 45–59.