Should Hospitals Keep Their Patients Longer? The Role of Inpatient Care in Reducing Post-Discharge Mortality

Ann P. Bartel
Economics, Columbia Business School; NBER, apb2@columbia.edu

Carri W. Chan
Decision, Risk, and Operations, Columbia Business School, cwchan@columbia.edu

Song-Hee Kim
Data Sciences and Operations, Marshall School of Business, University of Southern California, songheek@marshall.usc.edu

This version: February 13, 2019

The Centers for Medicare & Medicaid Services (CMS) and the National Quality Forum have endorsed the 30-day mortality rate as an important indicator of hospital quality. Concerns have been raised, however, as to whether *post-discharge* mortality rates are reasonable measures of hospital quality as they consider the frequency of an event that occurs after a patient is discharged and no longer under the watch and care of hospital staff. Estimating the *causal effect* of length-of-stay (LOS) on post-discharge mortality from retrospective data introduces a number of econometric challenges. We describe three potential sources of (endogeneity and censoring) biases and propose an approach that provides conservative estimates of the true treatment effect. Using a large dataset comprised of all hospital encounters of every Medicare Fee-for-Service patient with acute myocardial infarction from 2000 to 2011, we find evidence that an increase in LOS is associated with a decrease in 30-day mortality rates. An additional day in the hospital could decrease 30-day mortality rates by over 6%. Moreover, we find that, from a social planner's perspective, the gains achieved in reducing mortality rates likely exceed the cost of keeping the patients in the hospital for an additional day.

Key words: Healthcare, Medicare, Econometric Analysis, Length-of-stay, Mortality, Hospital quality

1. Introduction

The National Quality Forum (NQF), a non-profit organization that conducts research on hospital quality measures, and the Centers for Medicare and Medicaid Services (CMS) provide various publicly accessible measures of hospital quality. Both of these organizations publish their measures of hospital quality for two main purposes: (1) to provide information to potential patients regarding the quality of care at different hospitals and (2) to encourage hospitals to improve their quality of care. When CMS first launched the Hospital Compare website (www.medicare.gov/hospitalcompare/) in 2005, only process-of-care measures, such as the percentage of pneumonia patients given the most appropriate initial antibiotic(s), were reported. By 2008, CMS began reporting outcome measures such as 30-day mortality for acute myocardial infarction (AMI), heart failure (HF), and pneumonia (PNE) on its website (Centers for Medicare & Medicaid Services

2015). CMS's perspective that 30-day mortality rates are an important measure of quality is made clear under the Hospital Value Based Purchasing plan, which began in fiscal year 2013. Under this plan, hospital reimbursements are adjusted based on a Total Performance Score, which includes 30-day mortality rates for AMI, HF, and PNE (Centers for Medicare & Medicaid Services 2016). NQF also endorses 30-day mortality as a hospital quality measure because "it allows for a broad view of quality of care that encompasses more than what can be captured by individual process-of-care measures" (National Quality Forum 2009).

Although both CMS and NQF report 30-day mortality rates as a measure of hospital quality, there are conflicting views as to whether the reported 30-day mortality measure is a fair measure of quality. Indeed, there is an "ongoing debate about Hospital Compare, whose measures, critics say, do not necessarily reflect quality of care provided at hospitals" (Fleming 2012). One argument is that this 30-day mortality measure, which captures deaths which occur following hospital discharge, is likely to be strongly influenced by what happens *after hospital discharge*. Some people have raised questions as to whether it is fair to attribute post-discharge outcomes to hospitals whose focus is on *inpatient* care delivery. With this in mind, in-hospital mortality (rather than post-discharge mortality) is often considered as a measure of quality (e.g. Clark and Huckman (2012), Kc et al. (2013), Kuntz et al. (2015)). On the other hand, hospitals play a large role in what happens to patients after hospital discharge in that they typically arrange for follow-up appointments as well as communicate with follow-up care providers such as the primary care and specialist physicians. In this respect, "[i]t's reasonable to say that hospitals have some responsibility for what happens when the patient leaves" (Clark 2012).

Our study contributes to this debate by exploring whether there are factors related to *inpatient care* which can impact the *post-discharge outcome* of 30-day mortality, which, in this paper, is defined as whether a patient dies within 30 days of hospital discharge. Studies of process-of-care measures and their impact on mortality have led to mixed conclusions. For example, Jha et al. (2007) found that better process-of-care measures reported on Hospital Compare are related to lower risk-adjusted inpatient mortality, while Ryan et al. (2009) found that when controlling for hospital fixed effects, the process-of-care performance measures are not associated with 30-day mortality. The authors conclude that "this suggests that the relationship between hospital-level process-of-care performance and [30-day] mortality is not causal."

In contrast, our paper considers the impact of increasing a patient's hospital length-of-stay (LOS), with the idea that an extra day in the hospital may provide benefits such as allowing a patient to reach a higher level of stability as well as providing more time for patients to be educated about expectations with respect to their post-discharge behavior, thereby resulting in a reduction in the risk of mortality. Prior studies have shown that early discharges are associated with worse outcomes such as increased 30-day hospital readmissions for AMI patients (Carey 2015) and increased intensive care unit (ICU) readmissions (Kc and Terwiesch 2012). We hypothesize that a similar effect exists between LOS and mortality. A number of articles

have studied the relationship between LOS and mortality. Focusing on Medicare fee-for-service patients hospitalized for HF during the time period 1993-2006, Bueno et al. (2010) documented a decrease in hospital LOS and found that mortality rates had either stayed the same or increased, but they were unable to show a causal relationship between these trends. Chan et al. (2012) found no association between shorter ICU LOS and in-hospital mortality and Almond and Doyle (2011) found no effect of shorter postpartum hospital LOS on mortality or readmissions of mother and new-born. Other papers found that an increase in LOS was associated with an increase in the risk of post-discharge mortality (e.g. Williams et al. (2010), Nichols et al. (2014), Reynolds et al. (2015)). Looking at patients with similar ailments to those we study, Kaboli et al. (2012) examined Veterans Hospitals and found that reductions in LOS from 1997 to 2010 did not come at the expense of higher mortality. Clark (2012) argued that some hospitals may be 'cherry-picking' healthier patients who have shorter LOS and lower mortality rates or they may even be discharging/transferring patients with poor prognoses, so that "they look better when their death rates are compared with hospitals that keep patients longer." While these papers considered the relationship between LOS and mortality, none of them – other than Almond and Doyle (2011) –conducted a rigorous study of the *causal* effect.Note that the setting considered in Almond and Doyle (2011) is very different; moreover, their identification relies on a regression discontinuity design that cannot be used in our setting.

Our objective is to measure the *causal* effect of an increase in hospital LOS on post-discharge mortality, using a dataset from CMS that consists of all Medicare Fee-for-Service inpatient hospital visits between 2000 and 2011. Estimating the impact of LOS on the probability of post-discharge mortality is complicated for a number of reasons. First, it is not possible to perfectly measure a patient's severity level and unobservable severity factors might be positively correlated with both LOS and mortality risk. To address the possible endogeneity of LOS, we use an instrumental variable (IV) approach that is based on a patient's admission day-of-week. To circumvent the concerns that patient severity might differ by day-of-week, our analysis focuses on non-elective patients whose admission diagnosis is AMI because this diagnosis can be considered to be "non-deferrable" (Card et al. 2009), i.e. admissions are equally likely on the weekdays and weekend. In our data, we find that the residuals from a LOS equation for AMI patients admitted on Monday or Tuesday are negative, suggesting that they are 'prematurely' discharged. The average LOS for AMI patients is 5.3 days, which implies that patients admitted early in the week would be ready for discharge on the weekend. However, because hospitals prefer to discharge patients before the weekend (see Varnava et al. (2002) and Wong et al. (2009)), these patients end up with a shorter than normal LOS. This variation in LOS based on admission day-of-week helps us capture the impact of shorter LOS on increased mortality risk. To mitigate concerns about variation in hospital resource availability by day-of-week, we exclude weekend admits.

Our identification strategy relies on the non-deferrable nature of AMI along with our finding that observed patient-level covariates for those admitted on IV days (Monday or Tuesday) are not significantly different from the covariates of patients admitted on non-IV days (Wednesday, Thursday, or Friday). To address any concerns that admission day-of-week might be correlated with the unobserved severity of the patient's condition, we also conduct a sensitivity analysis that examines how an unobservable covariate could potentially affect our results. We find that in order for there to be an unobserved confounder that would explain away our results, the effect size of this unobserved confounder (or aggregation of multiple unobserved confounders) would have to be much larger than that of any observed confounder. Although the sensitivity analysis provides some support for the robustness of our results to violations in the exclusion criteria, we also use two alternative instrumental variables that may be less susceptible to the limitations of our Monday/Tuesday instrument. The alternative instrumental variables are (1) instruments for each admission day of the week, and (2) an indicator for whether a patient is predicted to be discharged on a Saturday. For both of these alternative instruments, we find that longer LOS is associated with a reduction in 30-day post discharge mortality risk, providing additional evidence to support the results found with the Monday/Tuesday instrument.

The second empirical challenge that results from using retrospective data is the potential for censoring biases in our outcome variable, 30-day post-discharge mortality. There are two types of censoring bias. First, since patients who die in the hospital cannot also die post-discharge, the post-discharge mortality outcome is censored for these patients. That is, in-hospital mortality and post-discharge mortality are competing risks, which can lead to biased estimates of the causal effect of LOS on post-discharge mortality. We address this concern by coding patients with in-hospital death as survivors, i.e. they do not experience 30-day post-discharge death.

The second source of censoring bias results from the fact that the risk of death is decreasing over time for patients with AMI. Thus, patients discharged later will *necessarily* have lower post-discharge mortality risk due solely to the fact that the window of time for the event to take place is later. To address this concern, we utilize an adjusted time-window for mortality, so that we consider whether patients die between 2 to 31 days after discharge if they may be discharged early, compared to a window of 1 to 30 days for patients who are not prematurely discharged. The two adjustments for censoring biases should result in a conservative estimate of the true treatment effect because we are using a reduced mortality rate by considering all patient with in-hospital death as surviving and we are using a shifted time window with lower mortality risk for all patients who are 'encouraged' to be discharged early, even though some will not be. We use simulation to demonstrate that, indeed, the adjustments for the two censoring biases result in conservative estimates of

the true treatment effect. That is, while the estimates are still biased, they are biased in the direction that makes it more difficult to detect an effect of shorter LOS increasing post-discharge mortality risk.

Our results show that keeping patients in the hospital for one more day could decrease 30-day mortality rates by over 6%. Using hospital cost estimates and value of life estimates, we calculate that keeping an AMI patient admitted on a Monday or Tuesday one more day in the hospital is likely to be cost-effective from a social welfare perspective. Moreover, we find that our results are very robust to a range of estimates of hospital costs and the value of life. One interpretation of our findings is that hospitals should consider moving to a 7-day discharge cycle, thereby reducing the likelihood of premature discharge prior to the onset of the weekend.

The remainder of the paper is structured as follows: Section 2 describes the dataset and the sample we use for our analyses. Section 3 describes the econometric challenges of our setting and our econometric approach to address them. In Section 4 we present our main results, including a number of robustness checks. In Section 5, we conduct a cost-benefit analysis and discuss the implications of our findings from a social planner's as well as a hospital administrator's perspective. We conclude in Section 6.

2. Setting

2.1. Data

We utilize data on all inpatient hospitalizations from 2000 to 2011 for Medicare Fee-For-Service (FFS) beneficiaries. Medicare FFS is the typical version of Medicare under which 70-80% of beneficiaries are covered. These data are drawn from the 100% sample in the Medicare Provider Analysis and Review (MedPAR) inpatient file[1]. Note that this does not include patients treated at the same hospitals, but not covered by Medicare. As a result, we do not have information about the congestion at each hospital. Although our dataset does not include operational metrics, it has huge value for a study of the determinants of post-discharge mortality. By virtue of it being a national dataset that covers more than 3500 hospitals in all states in the U.S. and includes all Medicare FFS patients hospitalized for AMI, it enables us to provide an extremely comprehensive analysis for the U.S.

Our observations are at the patient-visit level. For each hospitalization, we have the patient's demographic information including age, gender, race, coverage choice, and hospitalization characteristics including admission and discharge dates (which enable us to compute the patient's LOS and account for potential seasonal variations), the primary condition or other coexisting conditions identified by up to 10 International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes, the Medicare Severity adjusted Diagnosis Related Group (MS-DRG) classification (indicating the DRG to which the claims that

[1] See http://www.resdac.org/cms-data/files/medpar-rif for a description of this dataset.

comprise the stay belong for payment purposes), hospital, and admission type (e.g., elective or emergency basis). We also generate a severity of illness measure, the Elixhauser index (Elixhauser et al. 1998), using the ICD-9-CM codes and the MS-DRG classification. The patient outcome variable, 30-day post-discharge mortality, is defined as death within 30 days of discharge from a hospitalization.

2.2. Selection of Patient Sample

Our analysis is restricted to patients with AMI. AMI was one of the first three diagnoses for which mortality rates were reported on Hospital Compare and incorporated into the Value Based Purchasing Program. Moreover, as we will describe later, patients with AMI embody a number of desirable properties from an econometric standpoint. We use the primary ICD-9 codes to identify patients with AMI.

Appendix Table A.3 describes our sample selection process. We only consider hospital stays with admission and discharge that occur between January 1, 2000 and November 30, 2011. Because we study 30-day mortality, an event which occurs within 30 days of discharge, we exclude admissions and discharges that occur during December 2011 to avoid potential censoring of our outcome variable. Due to data fidelity concerns, we exclude visits with overlapping admissions (i.e., admissions that occur prior to discharge of the previous hospital stay). Following CMS (Grady et al. 2013), we focus on acute care stays. Stays that involve hospital transfers are excluded as it is difficult to control for what happens in two different hospitals and during the transfer time. Furthermore, we do not know why patients were transferred; e.g., they needed specialized treatment or wanted to be closer to home.

We exclude stays that are not paid under the current DRG code based prospective payment system (PPS) which Medicare switched to in 1983; after this significant payment change, patient care also began to change since payments were no longer based on the amount of time patients spent in the hospital, but rather based on the average cost to treat the particular DRG. Because we use DRG codes to control for patient severity, it is important these codes are used in a similar manner across hospitals.

We then keep the patients with AMI. Following CMS (Grady et al. 2013), we exclude admissions within 30 days of a prior hospitalization's discharge as it is unclear whether the first or second hospital stay impacts patient outcomes most. Since the hospitals with fewer than 25 visits for each corresponding condition do not have their performance publicly reported because of concerns regarding statistical power to assess performance with so few observations, we exclude hospitals that have less than 25 visits for AMI. Patients who are discharged to destinations that provide inpatient related services are excluded as we do not know the reason for transferring the patient between services, e.g., requirements of specialty services available at specific institutions versus requests due to personal preferences. In Section 4.3 we run robustness checks including these patients. We only include patients 65 years and older, which is the primary indication for

Medicare eligibility. Patients who left against medical advice, who do not have their race reported, or who do not reside in the U.S. are also excluded.

We focus on emergency and urgent (i.e., non-elective) patients to leverage the *random* variation in admission day-of-week to construct an instrumental variable (see Section 3.2 for details). Such an identification strategy is not possible for elective patients whose admissions are mostly scheduled. Following CMS, we exclude patients who are discharged on the same day they are admitted as they are unlikely to have been admitted for a true AMI. Next we exclude LOS outliers (greater than the 99th percentile value) and cost outliers (as identified in the MedPAR inpatient file). Note that we include patients with in-hospital death in our sample. As discussed in detail in Section 3.3, we include them as survivors (of 30-day post-discharge mortality) to address the in-hospital death bias.

2.3. Summary Statistics

Table 1 presents means and standard deviations for our patient sample as well as for the subset of patients who survive to hospital discharge. The 30-day post-discharge mortality rate is 6.4%. The average LOS for these patients is 5.3 days. Moreover, there is considerable variation in length of stay depending on the day of admission with a more than 7% gap between the shortest LOS for Monday admits and the longest LOS for Friday admits; in Section 3, we discuss how this variation enables us to construct an instrument to deal with the bias attributable to unobservable patient severity characteristics.

Note that the sample size given in Table 1 may not be exactly equivalent to the sample sizes in our regressions in Section 4. This is because some samples are dropped because they are perfect predictors of 30-day post-discharge mortality[2].

3. Econometric Model and Approach

Our goal is to estimate the impact of hospital LOS on 30-day post-discharge mortality. We start with the following reduced form equation:

$$y_i^* = \beta X_i + \theta \log(LOS_i) + \xi M_i + \psi Y R_i + \eta H_i + \epsilon_i \quad (1)$$

$$y_i = 1_{\{y_i^* > 0\}} \quad (2)$$

where y_i is the binary outcome with 1 indicating the patient died within 30-days following discharge from the hospital. Thus, y_i^* can be interpreted as the latent risk of death occurring in the 30 days post-discharge.

In equation (1), X_i is a vector of patient characteristics: age, gender, race, Elixhauser co-morbidities[3], DRG code, a dummy variable for having one or more surgical procedures (any minor/major diagnostic or

[2] For example, if all patients in hospital i die within 30 days of discharge, then a hospital fixed effect for hospital i would be a perfect predictor of mortality and all patients treated in hospital i would be dropped from the mortality regression.

[3] Elixhauser et al. (1998) defines 30 comorbid conditions using the ICD-9-CM and MS-DRG codes. Equation (1) includes 30 dummy variables, one for each of the 30 conditions.

Table 1 Summary statistics

	Including weekend admits		Excluding weekend admits	
	All	Excluding patients with in-hospital death	All	Excluding patients with in-hospital death
N	1,808,889	1,609,866	1,317,944	1,174,353
Age	78.8 (8.3)	78.4 (8.3)	78.7 (8.3)	78.4 (8.3)
Elixhauser Score	2.9 (1.5)	2.9 (1.5)	2.9 (1.5)	2.9 (1.5)
Female	0.50	0.50	0.50	0.50
Race				
White	0.89	0.89	0.89	0.89
Black	0.08	0.08	0.08	0.08
Hispanic	0.02	0.02	0.02	0.02
Other	0.02	0.02	0.02	0.02
Had surgical procedure(s)	0.71	0.71	0.71	0.71
Intensive care use				
No	0.49	0.50	0.49	0.50
General	0.26	0.24	0.25	0.24
Surgical	0.02	0.02	0.02	0.02
Medical	0.03	0.03	0.03	0.03
Intermediate	0.19	0.20	0.19	0.20
Other	0.02	0.02	0.02	0.02
Average LOS (days)	5.3 (3.4)	5.4 (3.4)	5.3 (3.4)	5.4 (3.4)
Sunday	5.2 (3.4)	5.4 (3.4)		
Monday	5.1 (3.5)	5.2 (3.4)	5.1 (3.5)	5.2 (3.4)
Tuesday	5.1 (3.4)	5.3 (3.4)	5.1 (3.4)	5.3 (3.4)
Wednesday	5.2 (3.4)	5.4 (3.4)	5.2 (3.4)	5.4 (3.4)
Thursday	5.4 (3.4)	5.5 (3.4)	5.4 (3.4)	5.5 (3.4)
Friday	5.5 (3.4)	5.6 (3.3)	5.5 (3.4)	5.6 (3.3)
Saturday	5.4 (3.4)	5.5 (3.3)		
Post-Discharge Death in 30 days	0.064	0.072	0.064	0.071

Note. Mean and standard deviation (in parentheses) shown.

therapeutic procedures)[4], and intensive care use indicators. M_i, YR_i and H_i are all vectors: M_i is the month of hospital admission; YR_i is the year of hospital admission; and, H_i is the hospital in which patient i is treated. Hence, we include month and year dummies as well as hospital fixed effects; the inclusion of the hospital fixed effects controls for the potential impact of unobservable attributes of the more than 3500 hospitals in our study. As is standard practice, we take the logarithm of the patient's LOS in order to account for the heavy tail in the distribution. We assume that the error term ϵ_i is a standard normal random variable to fit the Probit model.

[4] While we do not have any data on the socio-economic status of the patients, we believe this is of minimal concern. Although we expect patients with lower socio-economic status to be more likely to die, we also expect them to be less likely to be prematurely discharged as hospitals are hesitant to send patients home without a solid support system to help manage their recovery. If there were a positive correlation between premature discharge and socio-economic status, this could result in us erroneously concluding that premature discharge increases the likelihood of mortality when the true effect may be due to socio-economic status.

3.1. Empirical Challenges

While our retrospective patient dataset is quite rich, we are faced with a number of estimation challenges which we describe next. In Sections 3.2 and 3.3, we describe in detail how we address these potential sources of bias.

3.1.1. Endogeneity Although the Elixhauser co-morbidity conditions have been widely used in previous research, these measures are not a perfect control for patient severity. Unobservable severity factors might be positively correlated with both LOS and the dependent variable in equation (1). Since sicker patients tend to stay longer in the hospital and are also more likely to die, we might draw an erroneous conclusion that longer LOS leads to higher mortality risks. To address this concern, we utilize an instrumental variable approach.

3.1.2. Censoring due to in-hospital deaths We focus on post-discharge mortality for two primary reasons. First, this is a measure of quality that is commonly used to compare performance across different hospitals. Second, since we want to estimate the effect of LOS on mortality, post-discharge mortality is more appropriate than alternative mortality measures such as in-hospital mortality or post-admission mortality. If one examines *in-hospital mortality*, it is not clear how to interpret the coefficient of LOS as it does not make sense to keep a patient in the hospital an extra day if s/he has already died. In the case of post-admission mortality, other factors, such as care in the first 24 hours of hospital stay, may be more likely to have an impact than a patient's LOS.

However, by focusing on post-discharge mortality, we must recognize a potential bias introduced by patients with in-hospital death. In particular, in-hospital death is a competing risk to post-discharge death; a patient with in-hospital death cannot also have post-discharge death. As such, excluding patients who died in hospital from the sample could bias our results. To address this problem, we include patients with in-hospital death in our cohort, but code them as survivors of 30-day post-discharge mortality.

3.1.3. Censoring due to decreasing mortality hazard Studies have shown that for patients with AMI, the likelihood of death decreases with time following the AMI event (Dharmarajan et al. 2015). This is true in our data, as shown in Figure 1. Because of the decreasing hazard rate for mortality, when considering any single patient, if one looks at a later time window (e.g. for a patient discharged 1 day later), the mortality risk will necessarily be lower even if there is no effect of LOS on mortality. Ignoring this could lead one to conclude that shorter LOS increases mortality risk, *even if no effect exists*. To account for this potential bias, we propose a shifted time-window over which we define post-discharge mortality.

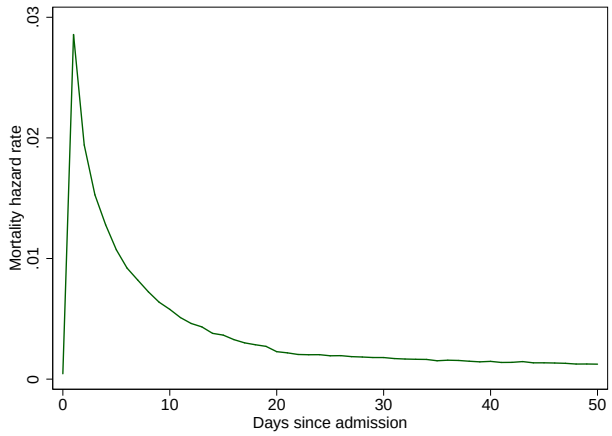


Figure 1 Mortality hazard rate of AMI patients in our data.

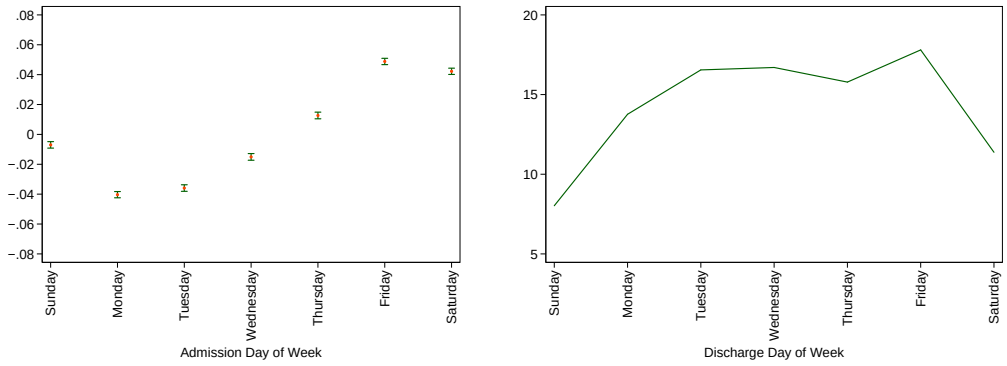
3.2. Addressing Endogeneity Bias

A valid instrumental variable (IV) is correlated with the endogenous variable ($\log(LOS)$) and uncorrelated with the unobservable noise (Wooldridge 2010). We propose an IV that is based on a patient's admission day-of-week and evaluate whether it satisfies these two properties. A number of studies in other healthcare settings have used time-of-day or day-of-week as an IV as the timing of admission has been shown to have an impact on the type of care patients receive (e.g. Ho et al. (2000), Hamilton et al. (2000), Ryan et al. (2005), Bhattacharya et al. (2008), Goyal et al. (2013), Baiocchi et al. (2014a)).

Relationship Between Admission Day-of-Week and LOS: We start by examining whether admission day-of-week is correlated with our endogenous variable, LOS. Table 1 shows that the average LOS for patients differs based on admission day-of-week. We estimated a separate regression of the logarithm of LOS on patient observables (age, gender, race, Elixhauser, DRG, had surgical procedure(s) or not), time dummies (month and year of hospital admission) and hospital fixed effects. Figure 2(a) shows the average residual from this regression plotted against the admission day-of-week for our patient cohort. We can see that patients admitted on Monday or Tuesday have negative average residuals, suggesting that they are 'prematurely' discharged.

The average LOS for AMI patients in our sample is 5.3 days. As such, patients admitted on Sunday, Monday or Tuesday are likely to be ready for discharge on the weekend. There is substantial evidence (e.g. Varnava et al. (2002), Wong et al. (2009)) that hospitals prefer to discharge patients just prior to the weekend rather than keeping the patients over the weekend when many services are not available; this is also consistent with patients' preferences to be discharged, if possible, prior to the weekend[5]. Indeed, there

[5] In discussions with administrators at a major medical center, we were informed that on the weekend social workers are generally not available and it is difficult to arrange for home health aides.



(a) Average of residuals and their 95% lower and upper bounds from regressions of $\log(LOS)$ on observables plotted against admission day-of-week

(b) Percentage of discharges by day-of-week

Figure 2 Day-of-week effect

seems to be evidence of this preference in our data where discharge rates peak on Friday and fall sharply on Saturday and Sunday (see Figure 2(b)). This suggests that we may be able to leverage the variation in LOS due to this "discharge before the weekend effect" as an identification strategy and isolate a valid instrumental variable based on admission day of week. Note that what we are considering a weekend effect is different than that seen in Rinne et al. (2014), which examines the impact of a weekend discharge on hospital readmissions and finds no effect. In contrast, we consider the effect of being discharged 'early' due to the hospitals' practice to discharge before the weekend. We start by using Monday/Tuesday admissions as our instrument, which allows us to conduct a sensitivity analysis in Section 4.1. Later, we use multiple instruments defined by admission on each day of the week.

Unobserved Severity by Admission Day-of-Week: For an IV to be valid, it must be uncorrelated with the unobservable noise; in our case, the admission day-of-week must be uncorrelated with the unobserved severity of the patient condition. We follow the approach used by Card et al. (2009) who define non-deferrable diagnoses as those for which admissions are equally likely on the weekend and weekdays. The rationale behind this approach is that patients will only go to the hospital on the weekend "for a relatively severe set of conditions that require immediate hospitalization", i.e. their condition is 'non-deferrable' (Card et al. 2009). Card et al. (2009) relies on non-deferrability to argue that hospitalized patients with these types of conditions who are just under age 65 are no different than those admitted who are just over age 65 (and, hence, Medicare eligible), while we use it to argue that unobserved patient severity measures are unlikely to be correlated with admission day of week. Specifically, for each ICD9 code, we calculate t-statistics to see if the proportion of patients admitted on the weekend is significantly different from 2/7 of the total

weekly admissions[6]. Recall that our analysis is restricted to patients admitted on an emergency or urgent basis and for these patients we find that 77% of ICD9s for AMI satisfy this criterion for non-deferrability. These results are consistent with Card et al. (2009), which finds that AMI satisfies the criterion for non-deferrability, implying that patient severity for these conditions does not differ by admission day-of-week.

A number of studies (e.g. Hamilton et al. (2000), Bhattacharya et al. (2008), Kc and Terwiesch (2012)) examine the exclusion criteria by comparing their IV to observable measures of severity. We do this as well and compare observed patient-level covariates for those admitted on Monday/Tuesday versus Wednesday/Thursday/Friday: age, gender, Elixhauser score, race, number of procedures a patient experiences during the hospital stay, intensive care use indicators, and the five most prevalent DRG codes (Table 2). Due to our large sample size, we focus on standardized differences, which is the mean difference divided by the average standard deviation (Flury and Riedwyl 1986). We find that the standardized differences of these covariates are well below the 0.2 rule-of-thumb (Rosenbaum and Rubin 1985). By this measure, there is no evidence that patients admitted on the IV days are different than those admitted on other days. Note, we also compared standardized differences between all pairs of weekdays and found no evidence of differences between any of the pairs.

Table 2 Standardized Differences

	MTu admissions	WThF Admissions	Std. Diff.
Age	78.69	78.78	-0.01
Female	0.50	0.50	-0.01
Race			
White	0.89	0.89	0.00
Black	0.08	0.08	0.00
Hispanic	0.02	0.02	0.00
Other	0.02	0.02	0.00
Elixhauser	2.93	2.93	0.00
# Procedures	3.13	3.10	0.01
Intensive care use			
No	0.49	0.49	0.00
General	0.25	0.25	0.00
Surgical	0.02	0.02	0.00
Medical	0.03	0.03	0.00
Intermediate	0.19	0.19	0.00
Other	0.02	0.02	0.00
Top 5 DRGs			
121 (before 2007)	0.27	0.27	0.00
122 (before 2007)	0.09	0.09	0.00
280 (after 2007)	0.08	0.08	0.00
123 (before 2007)	0.07	0.07	0.00
516 (before 2007)	0.06	0.06	0.00

Note. N=1,317,944. For patients admitted on weekdays only.

In sum, our identification strategy relies on the non-deferrable nature of AMI, the fact that observed

[6] To account for potential differences across states, we run each t-test by state.

covariates of AMI patients admitted on IV and non-IV days are similar, and that admission on an IV day does not directly affect a patient's mortality risk. However, as it is impossible to completely rule out the existence of an unobservable variable violating the exclusion criteria, we conduct a sensitivity analysis in Section 4.1 to examine the impact of potential violations of the assumption that our instrument is independent of unmeasured confounders.

Resource Availability on IV days: The motivation for using day of admission as an instrument is that patients have shorter LOS when admitted earlier in the week because of the desire to discharge patients before the weekend. However, there might be concerns that the type of care provided to patients on weekends is different than other days of the week. For instance, with less staffing and resources available on weekends, one may question if patients admitted on weekends have worse outcomes because of lack of access to care and not because of their LOS. Ryan et al. (2005) found that while cardiac patients admitted on the weekend have longer delays to catheterization, there does not seem to be any difference in outcomes. On the other hand, Dobkin (2003) found that weekend admission is associated with higher risk of mortality. In our main analysis, we exclude weekend admits to address any concerns regarding differences in care provided on the weekend and provide a robustness check in Section 4.3 that includes weekend admits.

It is common for many surgeries, especially the more complex ones, to be scheduled on Mondays. Thus, while we focus on emergency and urgent patients, the availability of surgical staff may be reduced for patients admitted to a surgical service on Monday. As 71% of AMI patients have some sort of surgical procedure, this may have an impact on their care. However, surgical schedules are unlikely to have a significant impact on care for patients admitted to a medical service. In Section 4.3, we will provide a robustness check which excludes patients who have a surgical procedure during their hospital stay. In addition, one could argue that patients admitted on Fridays may be more likely to die compared to patients who are admitted on Monday to Thursday because Friday admits will not get the same level of service during the critical first few days that overlap with the weekend. In Section 4.3, we will provide a robustness check which excludes patients admitted on Fridays.

3.3. Addressing Censoring Biases

In Sections 3.1.2 and 3.1.3 we described two potential censoring biases that can impact our study of post-discharge mortality: in-hospital deaths and decreasing mortality risks. To address these biases we propose two 'adjustments' to our outcome variable: 1) patients with in-hospital death will be included and coded as survivors and 2) patients who are discharged early will have a shifted time-window to define post-discharge mortality. In this section, we describe these adjustments.

In-hospital Death: To address the potential bias introduced by patients with in-hospital death, we incorporate the competing risk approach within the Probit model by setting 30-day post-discharge mortality to 0

for those who died in hospital. That is, setting 30-day mortality to 0 for those with in-hospital death assumes that everyone who died in the hospital would have otherwise survived 30 days, which would give us a conservative estimate of the effect of LOS on 30-day mortality. This idea is similar to the method used in Kuntz et al. (2015) to address the competing risk of hospital discharge when studying the impact of hospital occupancy on hospital mortality.

Decreasing Failure Rate: To address the potential bias introduced by the fact that the risk of death for patients with AMI has a decreasing failure rate, we propose to shift the window for which mortality is considered based on whether a patient may have been discharged early.

Consider a patient who is discharged (alive) on day t. Whether the patient dies on any day in the interval $[t+1, t+30]$ determines his 30-day post-discharge mortality. If, instead, the patient is discharged (alive) on day $t-1$, 30-day post-discharge mortality would be death on any day in the interval $[t, t+29]$. However, we need to consider the same time interval for death $[t+1, t+30]$ in order to avoid potential biases introduced by the decreasing failure rate. Thus, for any patient discharged early (on day $t-1$), we should consider the interval $[t+1, t+30]$ to define 30-day post-discharge mortality. If the patient dies on day t, we will count that as an in-hospital death.

These two adjustments together will bias our estimates for the treatment effect in the positive direction. This is because we are 1) using a lower mortality rate by including those with in-hospital deaths as surviving and 2) using a shifted time window with lower mortality risk for all patients who are 'encouraged' (by the IV) to be discharged early, even though some may not be. Simulation experiments in the Appendix demonstrate that this approach will result in conservative estimates of the true treatment effect.

3.4. Estimation Approach

As explained above, we use an instrumental variable approach as well as adjustments to the outcome variable of interest. In the first stage, we fit a linear model for $\log(LOS)$:

$$\log(LOS_i) = \hat{\beta} X_i + \hat{\xi} M_i + \hat{\psi} YR_i + \hat{\eta} H_i + \hat{\lambda} Z_i + \nu_i \tag{3}$$

In the second stage, a Probit model is estimated:

$$y_i^* = \beta X_i + \theta \log(LOS_i) + \xi M_i + \psi YR_i + \eta H_i + \epsilon_i \tag{4}$$

$$y_i = 1_{\{y_i^* > 0\}} \tag{5}$$

Thus, the first stage uses Z_i as an instrument for $\log(LOS)$ in the second stage. We let Z_i be an indicator that equals 1 if the patient is admitted on Monday or Tuesday, and 0 otherwise. These equations are estimated jointly via Maximum Likelihood Estimation (Wooldridge 2010). We estimate robust standard errors, clustered by admission day-of-week, month, and year. As part of our robustness checks in Section 4.3, we

use alternative clusters: DRG code; hospital; DRG code and year; hospital and year; DRG code, hospital, and year.

Considering the admission day-of-week effect in Equation (3), we expect the coefficient for our Monday/Tuesday admission day instrument (Z_i), $\hat{\lambda}$, to be negative. Finally, we hypothesize that these 'premature' discharges due to hospitals' desire to discharge patients before the weekend will increase the risk of mortality, so that θ is negative.

We use the data with adjustments made to the outcome variable y_i as described in Section 3.3. Specifically, any patient with in-hospital death is included in the sample with post-discharge mortality set to 0, i.e. $y_i = 0$[7]. A challenge that arises in considering the shifted time-window is that, unlike in a simulation, we cannot know for sure which patients are discharged early. We can only know that patients admitted on Monday or Tuesday are more likely to have shortened LOS. It is unreasonable to expect that *all* patients admitted on Monday/Tuesday were discharged early. We calculated the LOS residual among patients admitted on IV days and discharged from the hospital alive; patients with residual near -1 are likely to have been discharged 1 day early. We used the modified time-window for the mortality outcome (days $[2, 31]$ post-discharge, instead of days $[1, 30]$) for the 5% of patients with residual around -1. In our robustness checks, we consider larger groups of patients who may have been discharged early and require the adjusted outcome variable.

4. Results

Columns (1) and (2) of Table 3 show that when we do not instrument $\log(LOS)$, the coefficient of $\log(LOS)$ on the probability of mortality is positive and statistically significant at the .1% level. This bias is likely due to less sick patients (by unobservable measures of severity) being more likely to have shorter LOS and lower mortality risk.

In columns (3) and (4) we use Monday/Tuesday admission to instrument $\log(LOS)$, and find that the coefficient on log(LOS) is negative and significant at the .1% level. The F statistic for the significance of the instrument in the first-stage regression is 913.66 (adjusted for the day-month-year clusters), which indicates our instrument is quite strong. The average marginal effect of a one-day increase in LOS among patients who survived hospital discharge is a reduction in the 30-day mortality risk from .0612 to .0572, which is about a 6.5% decrease. We will use these estimates of the marginal effects when considering different patient care strategies in Section 5.

The results of the Wald χ^2 test suggest that our instrument is able to control for a substantial portion of the endogeneity bias in our sample. This, along with the non-deferrability results (Card et al. 2009) and

[7] Note that in doing so, some DRGs of patients with in-hospital death (e.g., DRG code 123: "Circulatory disorders with AMI, expired") become perfect predictors of $y_i = 0$. We replace such DRGs by their counterparts for survivors of in-hospital death, e.g., DRG code 123 is replaced by DRG code 121 "Circulatory disorders with AMI & major complication, discharged alive".

Table 3 30-Day Mortality Model Results

Model	(1) Probit (excluding patients w/ in-hospital death)	(2) Probit (including patients w/ in-hospital death as survivors and w/ adjustments to mortality measure)	(3) IV Probit (excluding patients w/ in-hospital death)	(4) IV Probit (including patients w/ in-hospital death as survivors and w/ adjustments to mortality measure)
Second Stage (Mortality)				
log(LOS)	0.14***	0.28***	-0.19**	-0.17*
	(0.00)	(0.00)	(0.07)	(0.07)
	$p=$0.000	$p=$0.000	$p=$0.006	$p=$0.015
Age, Gender, Race	Yes	Yes	Yes	Yes
Elixhauser Vars	Yes	Yes	Yes	Yes
Had surgical procedure(s)	Yes	Yes	Yes	Yes
Month, Year Dummies	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes
First Stage (log(LOS))				
IV			-0.06***	-0.05***
			(0.00)	(0.00)
			$p=$0.000	$p=$0.000
Age, Gender, Race			Yes	Yes
Elixhauser Vars			Yes	Yes
Had surgical procedure(s)			Yes	Yes
Month, Year Dummies			Yes	Yes
Hospital FE			Yes	Yes
Num. of obs.	1171148	1314136	1171148	1314136
Wald χ^2 test			22.71	39.31
Wald p-value			0.00	0.00

Note. Robust standard errors clustered by admission day-of-week, month, and year in parentheses. $^{+}$ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Table A.4 in the appendix provides the coefficients, robust standard errors, and p-values for all control variables for the model in column 4.

the results of our standardized differences comparing the observed patient-level covariates of AMI patients admitted on Monday/Tuesday versus other days, supports the reliability of the IV estimates. In results not reported here we found that the instrument is statistically significant and of the same magnitude even when stratifying by different levels of patient severity, as measured by the Elixhauser score, and by different patient conditions as measured by ICD9 codes, suggesting that the admission day of week effect is present across a large group of patients in our data. As such, it is reasonable to conclude that the population which complies with our instrument is broadly spread across heterogeneous patients. Thus, the IV results in Table 3 provide estimates for the local average treatment effect of LOS across heterogeneous patients. Finally, comparing columns (3) and (4), we see that, as expected by our analysis in the Appendix, the adjustments to the mortality measure bias the coefficient on LOS more positive.

4.1. Sensitivity Analysis

In order to mitigate the possibility that unobservable patient characteristics may be correlated with admission day of week, we study AMI patients who are admitted on an emergency or urgent basis and therefore

do not choose which day to be admitted to the hospital. One potential concern is that we observe a slightly higher admission rate on Monday compared to other days of the week: see Figure 3. This is in line with the literature that has found a Monday preference for AMI and stroke onset[8] (Spielberg et al. 1996, Manfredini et al. 2001, Arntz et al. 2000, Kinjo et al. 2003). The proposed factors for such a preference include stress from returning to work after the weekend and difference in family life and leisure patterns during the weekend versus weekdays. That said, one could argue that the higher admission rates of AMI patients on Mondays may be due to patients waiting until Monday to go to the hospital. This would only be possible for patients with 'deferrable' AMI, which we believe is likely to be a very small portion of the AMI population, if it exists at all. Patients who wait to go to the hospital may be sicker due to delays in getting treatment, which would bias our estimates more negative. On the other hand, patients who can afford to wait (instead of going immediately to the hospital over the weekend) may be less severe, which would bias our estimates more positive. Thus, if patients admitted on Monday are unobservably different than those admitted on other days of the week, it is not clear in which direction this would bias the results.

While we do not find evidence that AMI patients presenting to the hospital on Monday versus other days of the week are different (e.g., Table 2 shows that observed covariates for Monday/Tuesday admits versus Wednesday/Thursday/Friday admits are similar), we cannot completely rule out the possibility of an unobservable variable being correlated with admission day-of-week. For instance, we do not have data on congestion in the hospitals which, if correlated with day-of-week, could potentially bias our results[9]. We therefore use the methodology from Baiocchi et al. (2014b) to conduct a sensitivity analysis to see if our results could be explained away by an unobserved confounder that is correlated with our instrument.

Following Baiocchi et al. (2014b), suppose there exists an unobserved confounder, u_i, with mean 0 and variance 1 that is independent of the measured confounders, X_i, M_i, YR_i, H_i, but is correlated with the outcome (30-day mortality) and whether a patient is admitted on an IV day or not. We assume the following model:

$$y_i = \beta X_i + \theta \log(LOS_i) + \xi M_i + \psi YR_i + \eta H_i + \delta u_i + \epsilon_i$$
$$u_i = \phi Z_i + v_i$$

[8] In the literature, the day of AMI and stroke onset is defined as the time of onset of the first symptoms, usually reported by the patient, and not the day of hospital admission.

[9] There is some evidence that hospital occupancy is increases from Monday and Tuesday (e.g.Kanter and Moran (2007), Fieldston et al. (2011)). These findings suggest that *if* high occupancy rates increase mortality risks, this would disproportionately impact patients admitted on Wednesday to Friday (recall that we focus on weekday admissions in our analysis). Based on our hypothesis that shorter LOS will increase mortality risk and that Monday and Tuesday admits (our instrument) have shorter LOS, having high occupancy rates on Wednesday to Friday could result in conservative estimates of the true effect of LOS on mortality. That is, while the estimates may be biased, they are likely to be biased in the direction that makes it more difficult to detect an effect of shorter LOS increasing post-discharge mortality risk.

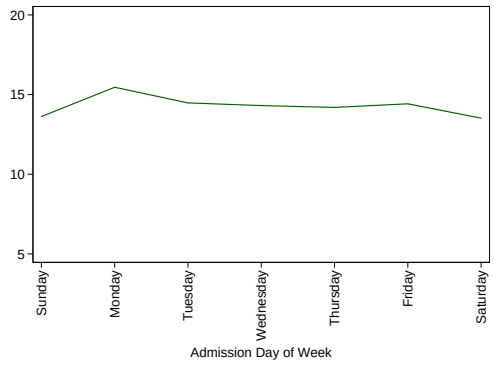


Figure 3 Percentage of AMI admissions by day-of-week.

$$E(v_i|X_i, M_i, YR_i, H_i, Z_i) = 0$$
$$E(\epsilon_i|X_i, M_i, YR_i, H_i, Z_i) = 0$$

We are interested in establishing a consistent estimate for θ. δ and ϕ are sensitivity parameters, where δ measures the effect of a one standard deviation increase in the unmeasured confounder on 30-day mortality under no treatment and ϕ measures how much higher the mean of the unobservable, u_i, is in standard deviation units of being admitted on the instrument days (Monday/Tuesday), Z_i. Under this model, Z_i would be a valid instrument if we could control for u_i. Thus, we can provide a consistent estimate for θ using 2SLS with $y_i - \delta\phi Z_i$ as the outcome variable and Z_i as the IV:

$$y_i - \delta\phi Z_i = \beta X_i + \theta \log(LOS_i) + \xi M_i + \psi YR_i + \eta H_i + \delta v_i + \epsilon_i$$

Note that the noise term, $\delta v_i + \epsilon_i$, is zero-mean conditional on X_i, M_i, YR_i, H_i and Z_i. We can run regressions for different values of δ and ϕ to determine parameter regimes where the estimate of θ is i) negative and statistically significant at the 5% level, ii) positive and statistically significant at the 5% level, or iii) statistically not different than 0 at the 5% level. If the regime where i) holds is large, we can conclude that our estimates for θ are reasonably robust to potential violations in the exclusion criteria of the IV being independent of unobserved confounders. This methodology requires that we use the 2SLS estimator, and hence we use a linear probability model for the 30-day mortality instead of a Probit model.

Figure 4 summarizes our sensitivity analysis for the 30-day mortality model. In particular, we see that the white region, where the estimates for θ are negative and statistically significant at the 5% level, is quite large. The gray (black) regime depicts where the estimates for θ are statistically not different than zero (are positive and statistically significant at the 5% level.) We note that the magnitude of the coefficients does not change (though the standard errors do) in the white region. Moreover, the (δ, ϕ)-values for all observed covariates (as indicated by x's) are well within the white region. That is, in order for there to be an unobserved

confounder that would explain away our LOS result, the effect size of this unobserved confounder would have to be much larger than that of any observed confounder. Specifically, unobserved patient severity factors must have a larger effect size than observable measures of health such as the MS-DRG codes or Elixhauser comorbidity conditions. Though it may be unlikely that any single unobserved confounder would have larger predictive power, it is possible that the cumulative effect of many unobservable factors together could impact the significance of our results.

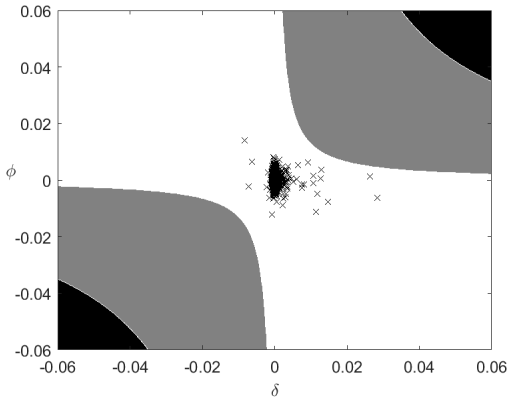


Figure 4 **Sensitivity Analysis for the 30-day mortality model. (1) White area represents regime for δ and ϕ where the estimates for θ are negative and statistically significant at the 5% level. (2) Gray area represents regime for δ and ϕ where the estimates for θ are statistically not different than 0 at the 5% level. (3) Black area represents regime for δ and ϕ where the estimates for θ are positive and statistically significant at the 5% level. x's represent δ and ϕ values for observed covariates.**

4.2. Alternative Instruments

In order to provide more evidence to support our conclusion that shorter LOS is associated with an increase in 30-day post-discharge mortality, we consider alternative instrumental variables that may be less susceptible to the potential limitations of the Monday/Tuesday instrument.

Indicators for all Admission Days-of-Week The rationale for using a single indicator for admissions on Monday or Tuesday as the instrument was that patients admitted earlier in the week are more likely to be discharged early. Using a single indicator also allowed us to conduct the sensitivity analysis of Section 4.1. However, Figure 2(a) shows that LOS varies across all admission days of the week. We therefore consider a specification with multiple instruments: four indicator variables, one for each admission day of week, with one day (Friday) being the base case.

Column (1) of Table 4 shows the results of this alternative specification. In the first stage, we see that patients admitted earlier in the week have shorter LOS than those admitted later in the week. We saw this previously when we first introduced the idea of using admission day of week as an IV. In addition, we

see that the coefficient on log(LOS) is negative and statistically significant in the second stage, providing additional support for the hypothesis that longer LOS is associated with a reduction in 30-day post discharge mortality risk.

Table 4 Alternative Instruments

	Excluding patients with in-hospital death			Including patients with in-hospital death as survivors and with adjustments to mortality measure		
	(1) All	(2) Cohort 1	(3) Cohort 2	(4) All	(5) Cohort 1	(6) Cohort 2
Second Stage (Mortality)						
log(LOS)	-0.15**	-0.78*	-1.13***	-0.13*	-0.96**	-1.10***
	(0.06)	(0.33)	(0.34)	(0.06)	(0.29)	(0.30)
First Stage (log(LOS))						
Instrument(s)						
Mon admit	-0.10***			-0.09***		
	(0.00)			(0.00)		
Tues admit	-0.09***			-0.08***		
	(0.00)			(0.00)		
Wed admit	-0.07***			-0.06***		
	(0.00)			(0.00)		
Thurs admit	-0.04***			-0.04***		
	(0.00)			(0.00)		
Fri admit	(base)			(base)		
Predicted Sat Discharge		0.02***	0.02***		0.02***	0.02***
		(0.00)	(0.00)		(0.00)	(0.00)
Num. of obs.	1171148	399713	309617	1314136	420377	322419
Wald χ^2 test	26.40	5.99	7.82	47.43	9.74	9.81
Wald p-value	0.00	0.01	0.01	0.00	0.00	0.00

Note. Robust standard errors clustered by admission day-of-week, month, and year in parentheses.$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Predicated Saturday Discharge We consider another alternative instrument which is also based on the idea that patients are less likely to be discharged on the weekend, thereby impacting LOS for some patients. Rather than examining this phenomenon by admission day-of-week, we consider the predicted day of discharge. Specifically, we define our instrumental variable as an indicator for whether a patient is predicted to be discharged on a Saturday.

In order to predict a patient's day of discharge, we start by predicting a patient's LOS. First, for all patients who were admitted on a weekday and survived to hospital discharge, we fit a length of stay regression including the following covariates: DRGs, Elixhauser dummies, age, gender, race, whether or not had

surgery, admission month, admission year, and hospital. Next, we use the results of this regression to predict a patient's LOS. A patient's predicted day of discharge is then determined by adding the predicted LOS to the patient's admission day. If it falls on a Saturday, then the patient is classified as having a predicted Saturday discharge; otherwise, he is classified as not having a predicted Saturday discharge. We use the indicator variable for whether a patient is predicted to be discharged on Saturday as our alternative IV. Because the predicted Saturday discharge IV is constructed using observed covariates, we believe it is reasonable to assume that there does not exist an unobservable variable that is correlated with both the predicted Saturday discharge and mortality, and hence satisfying the exclusion restriction.

We restrict our analysis to patients who are predicted to be discharged before the first Sunday following admission. In other words, we exclude patients admitted on Monday, Tuesday, Wednesday, Thursday, or Friday if their predicted LOS is greater than 5, 4, 3, 2, 1 days, respectively. We do this because these patients are more likely to stay in the hospital over the weekend, when fewer services are available and the type of care provided may be different than that available on weekdays. We define this group of patients as Cohort 1. Next, in order to better balance patient covariates, we conduct a four-to-one matching with replacement based on predicted LOS to generate the second cohort. Specifically, using the initial cohort of patients predicted to stay in the hospital only on weekdays, we match four patients with the same predicted LOS who are not predicted to have a Saturday discharge to one with the same predicted LOS but who is predicted to have a Saturday discharge. Table 5 shows the standardized differences between those with and without predicted Saturday discharge for these two cohorts. We can see that prior to matching (Cohort 1), the predicted Saturday discharge patients are different than those who are not. However, after matching, the standardized differences are well below the 0.2 rule-of-thumb.

Columns (2), (3), (5), and (6) of Table 4 show the results with this alternative IV based on predicted Saturday discharge for Cohort 1 (without matching) and Cohort 2 (with matching). We can see in the first stage that patients with predicted Saturday discharge have *longer* LOS than those without predicted Saturday discharge. Because both cohorts consider patients with relatively short predicted LOS (≤ 5), it seems that the reluctance to discharge patients on the weekend manifests itself by having patients stay longer, being discharged on Monday instead of Friday. Moreover, in the second stage, we see that the coefficient on log(LOS) is negative and statistically significant for both cohorts.

Because of the non-linearities introduced by the Probit model and the differences in the cohorts for the predicted Saturday discharge IV and our main specification in Table 3, it is difficult to directly compare the magnitude of the coefficients. For Cohort 1, the -0.96 coefficient in column (5) signifies an average treatment effect which reduces the mortality rate from 8.16% to 4.77% with an additional day in the hospital. For Cohort 2, the -1.10 coefficient in column (6) signifies a reduction in the mortality rate from 9.96% to

Table 5 Standardized Differences: Cohort 1 and Cohort 2

	Cohort 1			Cohort 2		
Predicted Sat discharge?	No	Yes	Std. Diff.	No	Yes	Std. Diff.
N	250677	180642		202707	133979	
Age	76.50	78.22	-0.21	77.30	77.32	0.00
Female	0.43	0.49	-0.11	0.46	0.46	0.00
Race						
White	0.91	0.90	0.05	0.91	0.91	0.00
Black	0.05	0.06	-0.05	0.06	0.06	0.00
Hispanic	0.01	0.01	-0.02	0.01	0.01	0.00
Other	0.02	0.02	0.00	0.02	0.02	-0.01
Elixhauser	2.28	2.64	-0.26	2.41	2.41	0.00
# Procedures	3.69	3.08	0.23	3.42	3.40	0.01
Intensive care use						
No	0.61	0.56	0.09	0.59	0.59	0.00
General	0.16	0.20	-0.08	0.18	0.18	0.00
Surgical	0.01	0.01	-0.01	0.01	0.01	0.01
Medical	0.02	0.02	-0.04	0.02	0.02	-0.01
Intermediate	0.19	0.19	0.00	0.19	0.19	0.00
Other	0.01	0.01	-0.02	0.01	0.01	0.00
Top 5 DRGs						
121 (before 2007)	0.10	0.22	-0.33	0.15	0.15	0.00
122 (before 2007)	0.15	0.13	0.04	0.15	0.15	-0.01
247 (after 2007)	0.14	0.07	0.23	0.10	0.09	0.00
516 (before 2007)	0.08	0.08	0.01	0.09	0.09	-0.01
281 (after 2007)	0.08	0.06	0.05	0.08	0.08	0.01

Note. For Cohort 2, means and variances are adjusted by observation weights resulting from the four-to-one matching with replacement.

5.45%. Cohort 2 is better balanced on observed covariates, suggesting it is less likely to be unbalanced by unobservable factors than Cohort 1. This suggests the results for Cohort 2 are likely to be more reliable. However, it is a slightly different subset of the patients, so the estimation results fundamentally only apply to the matched cohort rather than the full population.

Note that the compliers with respect to the predicted Saturday discharge IV are different from the compliers with respect to our original IV, Monday/Tuesday admission, and that each IV identifies the average treatment effect for its group of compliers. The compliers with respect to the predicted Saturday discharge IV are patients who would stay longer in the hospital if predicted to be discharged on the first Saturday after admission compared to other days while the compliers with respect to the Monday/Tuesday admission IV are patients who would stay shorter in the hospital if admitted on Monday or Tuesday compared to Wednesday, Thursday, or Friday. Thus, it is not surprising that the magnitude of the effect of LOS on mortality is different when different IVs, which capture different types of variation, are used. We emphasize, however, that getting similar directional results from two IVs that rely on two different groups of compliers increases our confidence in our finding that an increase in LOS is associated with a decrease in 30-day mortality rates.

As we also cannot rule out the possibility of confounding factors in these results with the new IV, we also conduct a sensitivity analysis similar to that done in Section 4.1 for the predicted Saturday discharge IV. This is possible since it is a single instrument. Figure A.2 in the Appendix summarizes the sensitivity analysis for

this IV for Cohort 2. The white region, where the estimates for θ are negative and statistically significant at the 5% level, is quite large. The gray regime depicts where the estimates for θ are statistically not different than zero. A number of observed covariates (x's) are close to the gray region suggesting that these results may be more sensitive to violations in the exclusion criteria than the Monday/Tuesday IV. However, because Cohort 2 was created through matching, it is less likely that there are substantial differences in unobservable factors.

4.3. Robustness Checks

This section presents a number of robustness checks for the primary specification that uses Monday/Tuesday admission as the instrument. We start by considering different ways of computing standard errors. In Table 3 we clustered standard errors by admission day-of-week, month, and year. Since there may be correlations between patient characteristics and care within a hospital or a DRG group, we also tried a number of alternative clustering approaches: by DRG code; hospital; DRG code and year; hospital and year; DRG code, hospital, and year. For our main model presented in column 4 in Table 3, IV Probit (w/ adjustments to mortality measure), the estimated robust standard errors from these alternative clusters are 0.08, 0.07, 0.08, 0.07, and 0.07, respectively, which gives p-values 0.030, 0.011, 0.023, 0.009, and 0.010, respectively. In addition, the robust standard errors reported in Table 3 are computed through closed-form equations. If we compute standard errors through bootstrapping, the bootstrap standard error is 0.06 with p-value 0.004 based on 100 replications.

Next, we consider different groups of patients for which we use a shifted time-window to define the post-discharge mortality outcome. Recall that since we cannot know for sure which patients were discharged early, we adjusted the time-window for the 5% of patients admitted on Monday or Tuesday whose LOS residuals were around -1. In Table 6, we use the adjusted time-window for a larger group of patients (10%, 15%, and 20%). We see that the coefficient on LOS is negative for all models, though it is not statistically significant (even at the $p < .10$ level) when using 20% of patients admitted on Monday/Tuesday. Recall from our simulation results that our coefficients are conservative estimates for the true effect, so the negative coefficients are suggestive that shorter LOS is associated with increased mortality risk.

Table 7 presents a number of additional robustness checks. The first issue that we address is the potential that the ICD9 and DRG codes (used to indicate the patients' conditions and severity) are inaccurate due to potential upcoding hospitals may utilize to increase Medicare payments which are provided on a fee-for-service basis. To address this, we restrict our analysis to non-profit hospitals because the prior literature has found that these hospitals have lower rates of upcoding than for-profits (see Silverman and Skinner (2004), Dafny (2005), and Powell et al. (2012)). The results in Column (1) show that while the coefficient on LOS is negative, it is no longer significant. If the DRG and ICD9 codes at non-profit hospitals are more accurate

Table 6 Robustness Checks - Varying % Monday and Tuesday admits for whom the 30-day mortality window is shifted by 1 day

	10%	15%	20%
log(LOS)	-0.16*	-0.14^{+}	-0.12
	(0.07)	(0.07)	(0.07)
IV	-0.05***	-0.05***	-0.05***
	(0.00)	(0.00)	(0.00)
Num. of obs.	1314136	1314136	1314155
Wald χ^2 test	36.25	33.16	29.81
Wald p-value	0.00	0.00	0.00

Note. Robust standard errors clustered by admission day-of-week, month, and year in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

due to lower rates of upcoding, then these results may be more reliable than those which include for-profit hospitals. Since our results are conservative estimates for the true treatment effect, it is possible that there is still a significant effect. Alternatively, it may be the case that for-profit hospitals are indeed doing more early discharges. Bueno et al. (2010) found that when hospitals switched to a Fee-for-Service payment model, LOS steadily decreased, likely with the intent to generate more revenue.

In Column (2), we restrict the analysis to hospitals that are in the top quartile for number of patients as one might anticipate that the effects of LOS on mortality could be different for these hospitals that are more likely to treat complex cases. We see that the coefficient is still negative and statistically significant as well as being very slightly larger in magnitude.

In Column (3), we exclude patients with a surgical procedure during their hospital stay as an increase in complex scheduled surgeries on Mondays may reduce the availability of staff to treat the emergency and urgent patients who are admitted to a surgical service. The coefficient on LOS is still negative and statistically significant as well as being larger in magnitude. Additionally, the coefficient for the IV is slightly larger in magnitude. As patients with surgeries may be subject to strict protocols which dictate their LOS (e.g. Gustafsson et al. (2011), Lassen et al. (2013), Miller et al. (2014), Thiele et al. (2015) among many other), there may be more discretion to discharge patients before the weekend for patients without surgical procedures.

In Column (4), we exclude patients admitted on Fridays because there may be fewer resources available for these patients during the critical first few days of hospitalization which coincides with the weekend. The coefficient on LOS is negative and statistically significant. Moreover, the effect size is quite a bit larger.

In Column (5), we include patients who are admitted on the weekends. While the coefficient is still negative, it is no longer statistically significant. Since our adjustments result in conservative estimates of the true treatment effect, it is possible that despite the lack of statistical significance in Column (5), there is still an effect of shorter LOS on increasing post-discharge mortality in these cohorts.

In Column (6), we randomly select one hospital encounter per patient so that we have a single hospitalization per patient. This is because an individual patient may have multiple admissions, which are each counted as a separate observation. In our sample, about 30% of our observations are from patients with multiple visits during our study period. Again, the results are consistent with that of Table 3.

In Column (7), we include patients who are discharged to destinations that provide inpatient related services. Recall that we initially excluded these patients from our analysis as we cannot control for the reason for such transfers. However, this might introduce a sample selection bias if the transferred patients are more likely to die. Indeed, the 30-day post-discharge mortality of patients sent to inpatient services is 13.5% and that of patients not sent to inpatient services is 6.7%. Still, when we include these patients in our analysis, the treatment effect estimate is still negative, statistically significant, with similar magnitude as our main specification.

Table 7 Robustness Checks - Different Subsets (In-hospital mortality included as survivors; 30 day mortality time window is shifted for 5% of Mon and Tues admits)

	(1) Non-profit	(2) Big	(3) w/o Surgical Procedures	(4) w/o Friday Admits	(5) w/ Weekend Admits	(6) Random Episode	(7) Include dis. to inpat. service
log(LOS)	-0.09	-0.20*	-0.22*	-0.27*	-0.06	-0.19*	-0.19**
	(0.09)	(0.08)	(0.10)	(0.11)	(0.07)	(0.08)	(0.07)
IV	-0.05***	-0.05***	-0.06***	-0.04***	-0.05***	-0.05***	-0.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Num. of obs.	926980	1028584	380333	1053176	1805879	1198312	1356838
Wald χ^2 test	19.05	30.56	17.96	22.70	27.66	38.03	44.29
Wald p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note. Robust standard errors clustered by admission day-of-week, month, and year in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In all seven columns of Tables 7, the coefficient on $\log(LOS)$ is negative and the results of the Wald χ^2 test show that our instrument is able to control for some of the endogeneity. In five out of seven cases, the coefficient estimate is statistically significant, suggesting our results are reasonably robust.

5. Cost-Benefit Analysis and Implications

In this section, we utilize our results from Section 4 to estimate the impact of a specific intervention that hospitals can implement or CMS can impose to reduce post-discharge deaths. In our analysis, we take the perspective of the social planner who aims to reduce adverse outcomes and overall costs. Having observed that keeping a patient in the hospital for one more day is an effective intervention to reduce 30-day mortality, we compare the following two policies: (1) **Keep the status quo** or (2) **Increase LOS by one day**. This allows us to compare the effect of inpatient care on post-discharge mortality.

To compare the cost-effectiveness of these two policies, we first discuss the cost estimates we will use. Taheri et al. (2000) estimate the *cost of an additional day in the hospital* to be $420 in 1998, which is $635 in 2017 when adjusted for inflation. Importantly, Taheri et al. (2000) show that the direct cost of the last day represents only 2.4% of the total hospitalization cost. The Henry J. Kaiser Family Foundation (2014) provides an alternative measure and reports that the average hospital expenses for a day of inpatient care in the U.S. was $1,960 in 2011, or $2,156 in 2017 dollars. However, this measure includes an adjustment for outpatient care and is therefore likely to be an overestimate of the actual costs of inpatient care. Based on these two references, we assume that the cost of keeping a patient in the hospital one more day is $635 or $2,156, depending on whether one uses the *marginal* or *average* cost estimate. We note that the cost of a hospital day may also change over time according to other factors besides inflation; an example is the introduction of more costly procedures and tests. This change could be more or less than overall inflation, so we will also consider the robustness of our results to various rates of increase between 1998 and 2017 for the Taheri et al. (2000) cost estimates and between 2011 and 2017 for the The Henry J. Kaiser Family Foundation (2014) cost estimates.

Next, we use the estimates provided in Murphy and Topel (2006) for the *benefits of reduced mortality*. They calculated the value of a life-year for an average 80 year-old (the approximate mean age of the patients in our sample) to be $150,000 per person in 1999, which translates to $221,926 per year or $18,494 per month in 2017 dollars. Recognizing this may be an overestimate for individuals with serious medical conditions, we also consider how robust our calculations are to alternative value of life estimates.

Using the results from our IV Probit models summarized in Tables 3 and 4, we can compute the average estimated mortality rate under the aforementioned two policies. First consider our main model in Column 4 of Table 3 where Monday or Tuesday admission is the IV. We note that our identification strategy relies on the increased likelihood that patients are discharged early if admitted on Monday or Tuesday due to the 'discharge before the weekend effect.' Thus, while we find evidence that longer LOS is associated with a reduction in mortality risk for patients admitted on Monday or Tuesday, we cannot make projections for the effect on other patients. As such, we focus our intervention on keeping those patients admitted on a Monday or Tuesday one more day. To be conservative, we restrict our sample to patients who are discharged on the first Friday after admission and had LOS shorter than the predicted LOS (i.e., it is likely that these patients experienced the 'discharge before the weekend effect'). In our dataset, there are 46,504 such patients admitted over 11.91 years, so we estimate there are approximately 3,905 such patients per year.

When there is no change, the average estimated mortality percentage for such patients predicted by the model is 9.60%. If we increase their LOS by 1 day, the average estimated mortality rate decreases by 0.67 percentage points (a 7% decrease).

Table 8 Estimated annual mortality rates for keeping patients in the hospital an extra day.

IV	# of affected patients/year	Mortality Rate: Baseline LOS	Mortality Rate: Intervention LOS + 1	# Lives saved
Mon/Tue admit (Table 3 Col 4)	3,905	9.60	8.93 (-7%)	26
All days of week (Table 4 Col 4)	7,241	8.21	7.45 (-9%)	55
Predicted Sat discharge (Table 4 Col 5)	5,575	14.93	7.61 (-49%)	408

Note. Only patients who are 'encouraged' by the IV to have shorter LOS and whose LOS is shorter than their predicted LOS are kept an additional day in the hospital.

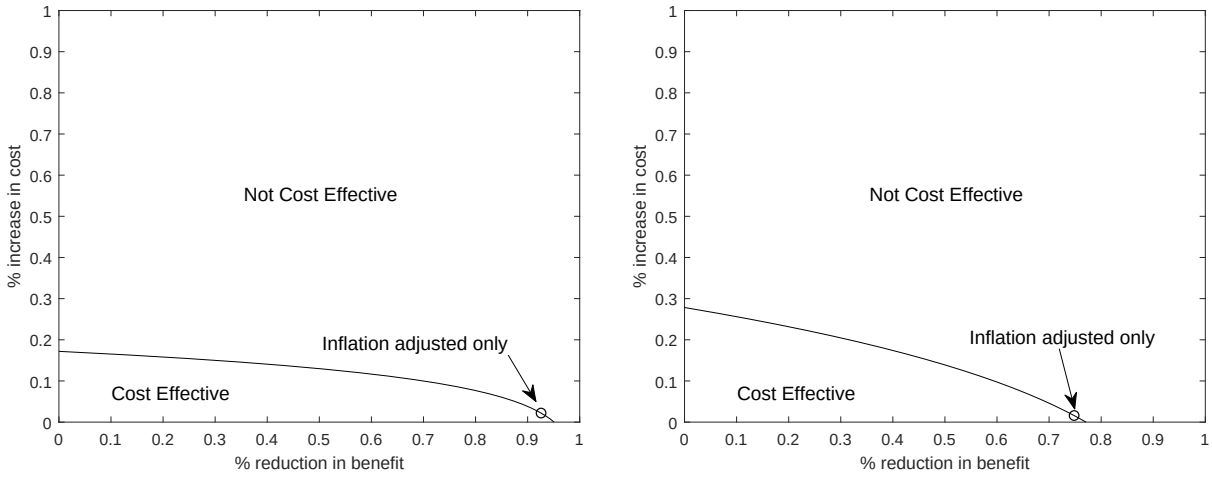
We explore whether keeping these patients in the hospital an extra day is cost-effective over the baseline of doing nothing. As summarized in Table 8, this intervention would result in savings of $3,905 \times 0.0067 = 26$ lives per year. The total value of these saved lives is $\$18,494 \times 26 = \$480,839$ for each month these patients live when we use Murphy and Topel (2006) to estimate the value of an additional month. If hospitals kept all 3,905 patients for one more day, the extra costs would range from $\$635 \times 3,905 = \$2.48m$ (using Taheri's estimate of the marginal cost of an extra day) to $\$2,156,905 = \$8.42m$ (using Kaiser's estimate of the average cost of a hospital day). This means that the patients would need to live $\$2,480,00/\$480,839 = 5.16$ months (using the marginal cost estimate) or 17.51 months (using the average cost estimate) in order for the inpatient intervention to be cost-effective over the baseline of doing nothing.

Note our results refer to the likelihood that a patient will live for 30 days post-discharge, but they do not provide insight into how long patients will survive after 30 days. Our cost-benefit analysis suggests that patients will need to survive for 5.16 or 17.51 months under the marginal or average cost estimates, respectively. If all patients died on day 31, then it would certainly not be cost-effective to keep all patients an extra day. On the other hand, if we believe that the patients who survive to 30 days due to the extra day in the hospital are similar to the patients who survive up to 30 days, then keeping patients an extra day is likely to be cost-effective. This is because on average, among the patients we consider, those who survive 30 days post-discharge live another 5.8 years[10]. In fact, when considering the distribution of the number of months survived after 30 days post-discharge for AMI patients, about 85% (72%) of patients survive beyond the cost-benefit break-even life-spans of 5.16 (17.51) months.

We next consider the robustness of the cost-effectiveness of our proposed intervention. In particular, since the average survival of AMI patients is 5.8 years, we consider the range of rates of increase in cost and the range in reductions in the value of living an additional month for which it is still cost-effective to keep patients an additional day. That is, if the percentage increase in cost is x, then the marginal cost of an additional day in 2017 would be $\$420 \times (1+x)^{19}$. The area under the curve in Figure 5(a) demonstrates

[10] Note that our estimates for average survival are conservative as our data are truncated with the last recorded date of death being December 26, 2012; for any patient missing a date of death (i.e., they did not die before 12/26/2012) we assigned a death date of December 26, 2012.

the regime where it is cost-effective to keep patients an additional day in the hospital when costs are given by the marginal cost from Taheri et al. (2000). The circle indicates the case where the cost of a hospital day is adjusted by inflation only (to $635 in 2017 dollars). We can see in this case the cost-effectiveness of keeping a patient an additional day is robust to reductions in the value of living an additional month of up to 92.58% of the estimates from Murphy and Topel (2006). It is also robust to increases in annual cost of up to 17.19% . Further reductions in the benefits of living an additional month and/or larger annual increases in cost would render the baseline as more cost-effective. Figure 5(b) shows the same for when the average cost estimate of an additional hospital day is used. We see that the cost-effectiveness of keeping patients in the hospital an additional day is very robust.



(a) Taheri et al. (2000) costs (b) The Henry J. Kaiser Family Foundation (2014) costs

Figure 5 Robustness of cost-effectiveness for keeping patients in the hospital an additional day, when treatment effect is based on the Monday/Tuesday IV (Table 3 Col 4). Baseline benefits are given by the value of living an additional month as estimated in Murphy and Topel (2006).

We conducted a similar analysis using the results based on the alternative instruments described in Section 4.2. For the specification with multiple instruments, one for each weekday admission, we consider keeping one more day those patients who were discharged on their first Friday in the hospital and had shorter LOS than predicted LOS. As summarized in Table 8, we estimate there to be 7,241 such patients per year and 55 lives saved per year. Table 9 summarizes the cost-effectiveness of such an intervention when using the Taheri et al. (2000) and The Henry J. Kaiser Family Foundation (2014) estimates for the cost of a hospital day. The results are very similar to the results based on the Monday/Tuesday admission IV.

Next, we used the results for the predicted Saturday discharge. Because we found that patients who were predicted to be discharged on Saturday had longer LOS, we considered the intervention of keeping all patients with predicted non-Saturday discharge who were discharged before the weekend and had LOS

shorter than predicted LOS one more day. We note that the 14.93% predicted baseline mortality rate for these patients is substantially higher than the observed mortality rate of 3.88%. The estimated treatment effect is much larger than our other IV analysis suggests; as such, the cost-effectiveness (summarized in Table 9) is very robust to reductions in the value of life and to increases in the cost of a hospital day. Figures A.3 and A.4 in the Appendix summarize the cost effectiveness of the intervention based on the treatment effect estimated by our alternative instruments.

Table 9 Cost-effectiveness of keeping patients in the hospital an extra day.

IV	Taheri costs				Kaiser Costs			
	Cost of int.	Months to live	Cost robustness	Val. of life robustness	Cost of int.	Months to live	Cost robustness	Val. of life robustness
Mon/Tue admit	$2.48m	5.16	17.19%	92.58%	$8.42m	17.51	27.85%	74.81%
All days of week	$4.60m	4.52	18.01%	93.50%	$15.61m	15.35	30.38%	77.92%
Pred. Sat dis.	$3.54m	0.47	32.95%	99.32%	$12.02m	1.59	90.63%	97.71%

5.1. Managerial and Operational Implications

Our cost-benefit analysis suggests that keeping AMI patients who are discharged early due to the preference to avoid discharges over the weekend in the hospital one more day is cost-effective from a social planner's perspective. Over the past few years, the U.S. government and CMS have taken steps through legislation, e.g. the Affordable Care Act, to provide incentives to healthcare providers to improve quality of care. Therefore, it is useful to consider the operational changes that hospitals would need to introduce if CMS were to require them to keep patients in the hospital a day longer.

Since we have documented that premature discharges occur right before the start of the weekend, our analysis suggests that one way to reduce the number of patients who are discharged too early is for hospitals to discharge patients 7-days a week rather than preferentially discharging Monday through Friday. In fact, this is an approach that the United Kingdom government is intending to implement by 2020 for the National Health Services[11]. In order to move to a 7-day-a-week discharge cycle, hospital managers would need to provide sufficient staff on the weekends, notably social workers and others who facilitate the discharge process. This would enable patients to avoid premature discharges due to hospitals' desire to discharge before the weekend. Doing so would introduce a number of operational decisions such as: how many staff are needed; what types of staff (e.g. physician assistants, social workers, nurses, etc.) are needed; and when staff should be scheduled. If the additional costs required to provide these resources are substantial and exceed our cost estimates for an extra hospital day, the robustness of our cost-benefit analysis will be reduced.

[11] https://www.gov.uk/government/collections/nhs-7-day-services, Accessed 12/15/2017.

Premature discharges may arise for other reasons (e.g. congestion in inpatient beds (Kc and Terwiesch 2012, 2009), family/patient preferences, etc.) and our results suggest that hospitals should also take steps to avoid these types of discharges whenever possible. For instance, hospital administrators should consider the potential benefits of increasing bed capacity if congestion in inpatient beds appears to be a frequent initiator for early discharges. Separately, when faced with pressure to discharge a patient early, physicians and social workers should educate patients and family members and emphasize the potential benefits of remaining in the hospital an extra day.

We note that though our analysis does not enable identification of the mechanism by which the extra hospital day helps reduce mortality risk, being able to do so is an important next step as it could provide insights into the operational implications of these findings. For instance, if the extra day serves to educate patients about post-discharge behavior, then it might be possible to achieve reductions in mortality risk by hiring more social workers and discharge nurses to begin the education process earlier. Alternatively, if the extra day helps patients reach a higher level of stability, this might imply that some hospitals will need to increase their bed capacity to accommodate the patients who will need to stay in the hospital longer. As such, our work suggests there are important resource allocation questions that must be studied to understand how to tactically reduce 30-day post-discharge mortality rates.

6. Conclusions

This paper examines the potential reductions in post-discharge mortality due to keeping heart attack (AMI) patients in the hospital longer. We identify and address three econometric challenges that researchers have faced in studying this question: (1) Endogeneity bias relating to the possibility of unobservable severity factors being correlated with length of stay and post-discharge mortality; (2) Censoring bias if patients who died in-hospital are excluded from the analysis; and (3) Censoring bias resulting from the decreasing mortality hazard that characterizes AMI. We use a very comprehensive data set from CMS that consists of all Medicare FFS in-hospital AMI patient visits between 2000 and 2011. AMI patients admitted on Monday or Tuesday are at risk of being discharged prematurely and we find that keeping these patients in the hospital for one more day saves an additional 26 lives per year and the value of these saved lives exceeds the cost of the extra hospital day under reasonable assumptions. These results are supported by additional evidence based on analysis with alternative instruments.

The fact that we find compelling evidence that keeping certain AMI patients in the hospital for one more day significantly decreases their mortality rates shows that there are factors within a hospital's control, i.e. LOS, that impact post-discharge mortality. This finding has significant implications in that it indicates that the 30-day mortality measures reported by Hospital Compare are indeed reasonable indicators of hospital

quality and the government should explore how to more widely disseminate the information available on the Hospital Compare website. Additionally, hospitals may need to revisit staffing and bed allocation decisions in order to be able to keep some patients in the hospital for an extra day.

Although we use a very comprehensive database, we excluded elective patients from our analysis because our instrument for LOS, day-of-week on which the patient was admitted, is most valid for patients admitted on an emergency or urgent basis. Hence, one limitation of our study is that the results may not apply to elective patients. A second limitation is that although we provide a number of important tests that support the validity of our instrument, we cannot completely rule out the possibility of an unobservable variable being correlated with admission day of week. It is important to note that such unobservable variables could have biased our results. While our results are consistent when using alternative instruments that utilize the impact of the weekend on early/later discharges, it would be interesting to study a setting where an instrument based on operational metrics or a randomized policy change might be available. A third limitation is that because we do not have information in our data about non-Medicare patients, we are not able to control for congestion, which has been shown to impact patient outcomes (e.g. Kc and Terwiesch (2012), Kuntz et al. (2015)). While it may be possible to get such detailed data at the level of individual hospital systems, to the best of our knowledge, unlike our MedPAR dataset, such a comprehensive data set does not exist at the national level. A fourth limitation of our work is that, while we provide evidence that an extra day in the hospital significantly reduces mortality risk for heart attack patients admitted on Monday or Tuesday we do not know exactly why the extra day is beneficial. An extra day may provide more time for patients to be educated about their post-discharge behavior and/or it may enable the patient to reach a higher level of stability. In addition, studies have shown that longer hospital stays are associated with high risks of adverse events such as adverse drug reactions, infections, and ulcers (Hauck and Zhao 2011). To assess the true benefit of an extra day, the findings of this paper need to be evaluated against such risks. Future research should explore the underlying causes of the relationship between hospital LOS and post-discharge mortality and also evaluate the benefit of increased hospital LOS against potential risks, which can help hospitals to improve their quality of care.

Acknowledgments

We thank the Strategy Cross-Disciplinary Area at Columbia Business School for its financial support of this project and Jean Roth of the NBER for data support. We are especially thankful to Jose Zubizarreta for pointing us to the sensitivity analysis with IVs and to Fanyin Zheng for many helpful conversations. We also thank participants in the Microeconomics Faculty Lunch at Columbia Business School and the INFORMS and POMS Meetings as well as the anonymous referees and Management Science Editorial team for very helpful comments. Outstanding research assistance was provided by Raymond Lim, Abraham Bae and Jimmy Charite.

References

Almond, Douglas, Joseph J Doyle. 2011. After midnight: A regression discontinuity design in length of postpartum hospital stays. *American Economic Journal: Economic Policy* **3**(3) 1–34.

Arntz, H.-R., S. N. Willich, C. Schreiber, T. Brüggemann, R. Stern, H.-P. Schultheiß. 2000. Diurnal, weekly and seasonal variation of sudden death. population-based analysis of 24061 consecutive cases. *European Heart Journal* **21**(4) 315–320.

Baiocchi, M., J. Cheng, D. S. Small. 2014a. Instrumental variable methods for causal inference. *Statistics in Medicine* **33**(13) 2297–2340.

Baiocchi, M., J. Cheng, D. S. Small. 2014b. Tutorial in biostatistics: Instrumental variable methods for causal inference. *Statistics in Medicine* **33** 2297–2340.

Bhattacharya, J., A. M. Shaikh, E. Vytlacil. 2008. Treatment effect bounds under monotonicity assumptions: An application to swan-ganz catheterization. *The American Economic Review* **98**(2) 351–356.

Bueno, H., J. S. Ross, Y. Wang, J. Chen, M. T. Vidán, S. T. Normand, J. P. Curtis, E. E. Drye, J. H. Lichtman, P. Keenan, M. Kosiborod, H. M. Krumholz. 2010. Trends in length of stay and short-term outcomes among medicare patients hospitalized for heart failure, 1993-2006. *JAMA: The Journal of the American Medical Association* **303**(21) 2141–2147.

Card, D., C. Dobkin, N. Maestas. 2009. Does Medicare Save Lives? *The Quarterly Journal of Economics* **124**(2) 597–636.

Carey, K. 2015. Measuring the hospital length of stay/readmission cost trade-off under a bundled payment mechanism. *Health Economics* **24**(7) 790–802.

Centers for Medicare & Medicaid Services. 2015. Hospital quality initiative: Hospital compare (accessed 8/14/15). URL https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Hos

Centers for Medicare & Medicaid Services. 2016. Hospital value-based purchasing. URL https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospit Accessed March 10, 2016.

Chan, C. W., V. F. Farias, N. Bambos, G. Escobar. 2012. Optimizing icu discharge decisions with patient readmissions. *Operations Research* **60** 1323–1342.

Clark, C. 2012. Stop using in-hospital mortality rates to judge quality. HealthLeaders Media (accessed 08/14/15). URL http://healthleadersmedia.com/content.cfm?content_id=275021&page=1&topic=qua.

Clark, J. R., R. S. Huckman. 2012. Broadening focus: Spillovers, complementarities, and specialization in the hospital industry. *Management Science* **58**(4) 708–722.

Dafny, Leemore S. 2005. How do hospitals respond to price changes? *American Economic Review* **95**(5) 1525–1547.

Dharmarajan, K., A. F. Hsieh, V. T. Kulkarni, Z. Lin, J. S. Ross, L. I. Horwitz, N. Kim, L. G. Suter, H. Lin, S.-L. T. Normand, H. M. Krumholz. 2015. Trajectories of risk after hospitalization for heart failure, acute myocardial infarction, or pneumonia: retrospective cohort study. *The BMJ* **350** h411.

Dobkin, C. 2003. Hospital staffing and inpatient mortality. *working paper* .

Elixhauser, A., C. Steiner, D. R. Harris Robert, R. M. Coffey. 1998. Comorbidity measures for use with administrative data. *Medical care* **36**(1) 8–27.

Fieldston, Evan S, Meera Ragavan, Bhuvaneswari Jayaraman, Kelly Allebach, Susmita Pati, Joshua P Metlay. 2011. Scheduled admissions and high occupancy at a children's hospital. *Journal of hospital medicine* **6**(2) 81–87.

Fleming, C. 2012. Medicare hospital quality reporting brings little or no mortality improvement. Health Affairs Blog (accessed 08/14/15). URL http://healthaffairs.org/blog/2012/03/06/medicare-hospital-quality-reporting-brings-li

Flury, B. K., H. Riedwyl. 1986. Standard distance in univariate and multivariate analysis. *The American Statistician* **40**(3) 249–251.

Goyal, N., J. R. Zubizarreta, D. S. Small, S. A. Lorch. 2013. Length of stay and readmission among late preterm infants: an instrumental variable approach. *Hospital Pediatrics* **3** 7–15.

Grady, J. N., Z. Lin, Y. Wang, C. Nwosu, M. Keenan, K. Bhat, H. Krumholz, S. Bernheim. 2013. 2013 measures updates and specifications: Acute myocardial infarction, heart failure, and pneumonia 30-day risk-standardized mortality measure (version 7.0). Yale University/Yale-New Haven Hospital-Center for Outcomes Research & Evaluation (Yale-CORE): Technical Report. Accessed August 8, 2015.

Gustafsson, U. O., J. Hausel, A. Thorell, O. Ljungqvist, M. Soop, J. Nygren, Enhanced Recovery After Surgery Study Group. 2011. Adherence to the enhanced recovery after surgery protocol and outcomes after colorectal cancer surgery. *Archives of Surgery* **146**(5) 571–577.

Hamilton, Barton H, Vivian Ho, Dana P Goldman. 2000. Queuing for surgery: is the us or canada worse off? *Review of Economics and Statistics* **82**(2) 297–308.

Hauck, Katharina, Xueyan Zhao. 2011. How dangerous is a day in hospital? a model of adverse events and length of stay for medical inpatients. *Medical care* 1068–1075.

Ho, V., B. H. Hamilton, L. L. Roos. 2000. Multiple approaches to assessing the effects of delays for hip fracture patients in the United States and Canada. *Health Services Research* **34** 1499–518.

Jha, Ashish K., E. John Orav, Zhonghe Li, Arnold M. Epstein. 2007. The inverse relationship between mortality rates and performance in the hospital quality alliance measures. *Health Affairs* **26**(4) 1104–1110.

Kaboli, P. J., J.T. Go, J. Hockenberry, J.M. Glasgow, S.R. Johnson, G.E. Rosenthal, M.P. Jones, M. Vaughan-Sarrazin. 2012. Associations between reduced hospital length of stay and 30-day readmission rate and mortality: 14-year experience in 129 veterans affairs hospitals. *Annals of internal medicine* **157**(12) 837–845.

Kanter, Robert K, John R Moran. 2007. Hospital emergency surge capacity: an empiric new york statewide study. *Annals of emergency medicine* **50**(3) 314–319.

Kc, D., B. R. Staats, F. Gino. 2013. Learning from my success and from others' failure: Evidence from minimally invasive cardiac surgery. *Management Science* **59**(11) 2435–2449.

Kc, D., C. Terwiesch. 2009. Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science* **55** 1486–1498.

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

Kinjo, K., H. Sato, I. Shiotani, T. Kurotobi, Y. Ohnishi, E. Hishida, D. Nakatani, H. Mizuno, Y. Yamada, S. Fukui, M. Fukunami. 2003. Variation during the week in the incidence of acute myocardial infarction: increased risk for japanese women on saturdays. *Heart* **89**(4) 398–403.

Kuntz, L., R. Mennicken, S. Scholtes. 2015. Stress on the ward: Evidence of safety tipping points in hospitals. *Management Science* **61**(4) 754–771.

Lassen, K., M. M. E. Coolsen, K. Slim, F. Carli, J. E. de Aguilar-Nascimento, M. Schäfer, R. W. Parks, K. C. H. Fearon, D. N. Lobo, N. Demartines, M. Braga, O. Ljungqvist, C. H. C. Dejong. 2013. Guidelines for perioperative care for pancreaticoduodenectomy: Enhanced recovery after surgery (eras) society recommendations. *World Journal of Surgery* **37**(2) 240–258.

Manfredini, R., I. Casetta, E. Paolino, O. la Cecilia, B. Boari, E. Fallica, E. Granieri. 2001. Monday preference in onset of ischemic stroke. *The American Journal of Medicine* **111**(5) 401–403.

Miller, T. E., J. K. Thacker, W. D. White, C. Mantyh, J. Migaly, J. Jin, A. M. Roche, E. L. Eisenstein, R. Edwards, K. J. Anstrom, R. E. Moon, T. J. Gan. 2014. Reduced length of hospital stay in colorectal surgery after implementation of an enhanced recovery protocol. *Anesthesia & Analgesia* **118**(5) 1052–1061.

Murphy, K. M., R. H. Topel. 2006. The value of health and longevity. *Journal of Political Economy* **114** 871–904.

National Quality Forum. 2009. Measure Submission and Evaluation Worksheet 5.0: Pulmonary Project: Hospital 30-day, All-cause, Risk-standardized Mortality Rate (RSMR) following Pneumonia Hospitalization. Tech. Rep. 0468.

Nichols, G.A., K. Reynolds, M.G. Butler, T.M. Kimes, W.W. Chan. 2014. Association of acute heart failure length of stay with subsequent re-hospitalization and all-cause mortality among patients with preserved vs. reduced ejection fraction. *Circulation* **130** A11307.

Powell, A., S. Savin, N. Savva. 2012. Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *Manufacturing & Service Operations Management* **14**(4) 512–528.

Reynolds, K., M. G. Butler, T. M. Kimes, A. G. Rosales, W. Chan, G. A. Nichols. 2015. Relation of acute heart failure hospital length of stay to subsequent readmission and all-cause mortality. *The American Journal of Cardiology* **116**(3) 400 – 405.

Rinne, S. T., B. F. Collins, M. Perkins, D. H. Au, C.-F. Liu. 2014. Weekend discharges after copd hospitalizations are not associated with a higher risk of readmission. *American Journal of Respiratory and Critical Care Medicine* **189**(American Thoracic Society 2014 International Conference, Meeting Abstracts) A3717.

Rosenbaum, P. R., D. B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* **39**(1) 33–38.

Ryan, Andrew M., James F. Burgess, Christopher P. Tompkins, Stanley S. Wallack. 2009. The relationship between medicare's process of care quality measures and mortality. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* **46**(3) 274–290.

Ryan, J. W., E. D. Peterson, A. Y. Chen, M.T. Roe, E. M. Ohman, C. P. Cannon, P. B. Berger, J. F. Saucedo, E. R. DeLong, S. L. Normand, C. V. Pollack Jr, D. J. Cohen. 2005. Optimal timing of intervention in non-ST-segment elevation acute coronary syndromes: insights from the CRUSADE (Can Rapid risk stratification of Unstable angina patients Suppress ADverse outcomes with Early implementation of the ACC/AHA guidelines) Registry. *Circulation* **112** 3049–3057.

Silverman, Elaine, Jonathan Skinner. 2004. Medicare upcoding and hospital ownership. *Journal of Health Economics* **23**(2) 369–389.

Spielberg, C., D. Falkenhahn, S. N. Willich, K. Wegscheider, H. Völler. 1996. Circadian, day-of-week, and seasonal variability in myocardial infarction: comparison between working and retired patients. *American Heart Journal* **132**(3) 579–585.

Taheri, P. A., D. A. Butz, L. J. Greenfield. 2000. Length of stay has minimal impact on the cost of hospital admission. *Journal of the American College of Surgeons* **191**(2) 123–130.

The Henry J. Kaiser Family Foundation. 2014. Hospital adjusted expenses per inpatient day. URL http://kff.org/other/state-indicator/expenses-per-inpatient-day/. Accessed July 3, 2014.

Thiele, R. H., K. M. Rea, F. E. Turrentine, C. M. Friel, T. E. Hassinger, B. J. Goudreau, B. A. Umapathi, I. L. Kron, R. G. Sawyer, T. L. Hedrick, T. L. McMurry. 2015. Standardization of care: Impact of an enhanced recovery protocol on length of stay, complications, and direct costs after colorectal surgery. *Journal of the American College of Surgeons* **220**(4) 430 – 443.

Varnava, A. M., J. E. Sedgwick, A. Deaner, K. Ranjadayalan, A. D. Timmis. 2002. Restricted weekend service inappropriately delays discharge after acute myocardial infarction. *Heart* **87** 216–219.

Williams, T. A., K.M. Ho, G.J. Dobb, J.C. Finn, M. Knuiman, S.A.R. Webb. 2010. Effect of length of stay in intensive care unit on hospital and long-term mortality of critically ill adult patients. *British journal of anaesthesia* **104**(4) 459–464.

Wong, H., R. C. Wu, G. Tomlinson, M. Caesar, H. Abrams, M. W. Carter, D. Morra. 2009. How much do operational processes affect hospital inpatient discharge rates? *Journal of Public Health* **31** 546–553.

Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*. MIT press.

Appendix

A. Addressing Censoring Biases - Details

In Section 3.3, we proposed an adjustment to our outcome variable to address censoring biases caused by decreasing mortality risks. In this section, we use simulation models to demonstrate that our approach will result in conservative estimates of the true treatment effect.

A.1. Simulation Setup

We now describe our simulation setup, which we use to explore the impact of our adjustments on our estimates. We simulate 1,000,000 hospital stays. For each hospital stay we use a beta distribution to simulate *time to death*, t_{death}, and a negative binomial distribution to simulate *time to discharge*, $t_{discharge}$. These two random variables are independent of each other. The distributions and parameters have been carefully chosen to best represent what we observe in the data.[12] *Length of stay*, los, is defined as the minimum of t_{death} and $t_{discharge}$. If $t_{discharge} \geq t_{death}$, patients die during hospital stay and do not experience their simulated time of discharge. If $t_{discharge} < t_{death} < los + 30$, patients die within 30 days of discharge.

Increasing los has no effect on mortality in the current setup; it only affects whether the patient dies in hospital or after discharge. In order to introduce the effect of length of stay, we randomly select $p\%$ of the patients[13] and modify their length of stay and 30-day post-discharge mortality according to the following procedure:

$$los' = los - 1,$$
$$\mathbb{1}(\text{30-day post-discharge mortality}) = \mathbb{1}(min(Pr(\text{30-day post-discharge mortality} \mid los) \times \alpha, 1) > U_{survive})$$

where Pr(30-day post-discharge mortality $\mid los$) are empirically derived (see Figure A.1), $U_{survive}$ is a uniform random number between 0 and 1, and $\alpha \geq 1$ is a design parameter which dictates the impact of shorter LOS. That is, for the $p\%$ of the patients that are randomly selected, their los is reduced by 1 day. Multiplying Pr(30-day post-discharge mortality $\mid los$) by α introduces the increase in mortality risk due to being discharged a day early.[14] If $\alpha = 1$, the decreased length of stay has no effect. For $\alpha > 1$, the decreased length of stay increases mortality risk. We use a Probit model to estimate the impact of los on $\mathbb{1}$(30-day post-discharge mortality).

Above, we proposed two 'adjustments' to our outcome variable to address two potential sources of biases. We apply them to our simulation to examine their effectiveness. Specifically, we 1) let $\mathbb{1}(\text{30-day post-discharge mortality}) = 0$ for patients who die during hospital stay and 2) use Pr(30-day post-discharge mortality $\mid los$) instead of Pr(30-day post-discharge mortality $\mid los'$) to define $\mathbb{1}$(30-day post-discharge mortality) for the randomly selected patients, which ensures that we consider the same time window for mortality as the patients whose length of stay did not change.

Note that in this simulation setup, we are able to avoid any unobservable confounders. Thus, while our original setting has three potential sources of bias, our simulation setup here focuses only on the two censoring biases.

[12] We choose the distributions and parameters so that the mortality risk declines over time (see Figure 1) and the average length of stay, the in-hospital mortality rate and the 30-day post-discharge mortality rate are similar to what we observe in the data. See Appendix Figure A.1 for details.

[13] In the simulated data, about 0.3% of the patients has $los = 0$. Because we cannot reduce their los, we do our patient selection only among the patients with $los > 0$.

[14] If a selected patient has $los - 1 = los' = t_{death}$, we treat the patient as an in-hospital death.

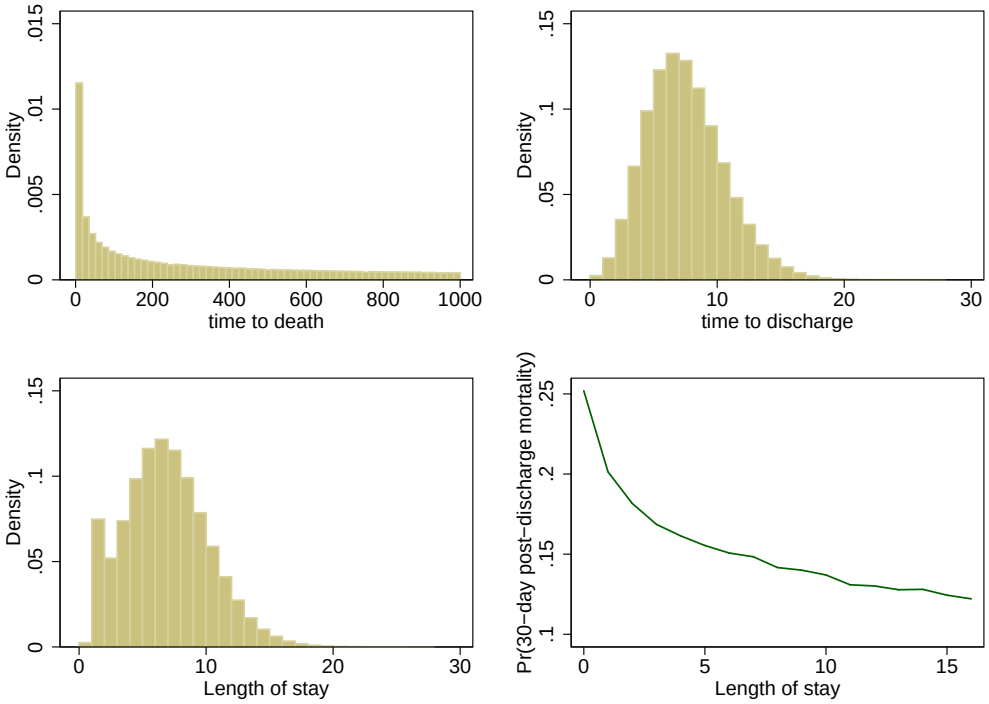


Figure A.1 **Simulation Setup: A beta distribution with shape parameters $a = 0.4$ and $b = 1$ is used to simulate time to death (`STATA command ceil(rbeta(0.4,1)*1000)` and a negative normal distribution with $n = 20$ and $p = 20/27$ is used to simulate time to discharge (`STATA command rnbinomial(20,20/27)`). Length of stay is defined as the minimum of time to death and time to discharge. Histograms of time to death, time to discharge and length of stay and the graph of Pr(30-day post-discharge mortality | length of stay) are shown. All based on 1,000,000 simulated observations.**

A.2. Simulation Results

The first table in Table A.1 presents the regression results when none of our proposed adjustments are made; that is, both the in-hospital death bias and decreasing mortality hazard bias are present. The coefficient of los is negative and statistically significant even when there is no effect of LOS on 30-day post-discharge mortality, $\alpha = 1$, which suggests that without addressing the two biases, we might be capturing a spurious effect of length of stay on mortality.

The second table in Table A.1 presents the regression results when both of our proposed adjustments are made. Now the coefficient of los is positive and statistically significant when there is no effect of LOS on 30-day post-discharge mortality, i.e. $\alpha = 1$. This suggests that our two proposed adjustments are likely to over-adjust for the two potential biases. Thus, if we apply our two adjustments and still find a negative coefficient for length of stay, it is likely a conservative estimate for the true magnitude of the treatment effect. When there is an effect ($\alpha > 1$) and as the effect size increases, the coefficient estimates move in the negative direction, capturing the increasing treatment effect. It also shows that there exists parameter regimes where the estimated coefficient is negative ($p = 20\%$ and $\alpha = 3$), suggesting that our proposed adjustments are not so large as to always completely mask a treatment effect when it exists.

Table A.1 Simulation Results: 30-Day Mortality

In-hospital death bias present & Decreasing mortality hazard bias present									
p	5%	5%	5%	10%	10%	10%	20%	20%	20%
α	1	2	3	1	2	3	1	2	3
los	-0.024***	-0.028***	-0.032***	-0.026***	-0.033***	-0.040***	-0.032***	-0.043***	-0.053***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	-0.834***	-0.773***	-0.716***	-0.814***	-0.701***	-0.596***	-0.764***	-0.565***	-0.383***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Num. of obs.	866284	866284	866284	865871	865871	865871	864990	864990	864990
Pseudo R^2	0.003	0.004	0.005	0.003	0.005	0.007	0.005	0.008	0.013

In-hospital death bias addressed by coding these patients $y_i = 0$ & Decreasing mortality hazard bias addressed by using an adjusted time-window									
p	5%	5%	5%	10%	10%	10%	20%	20%	20%
α	1	2	3	1	2	3	1	2	3
los	0.013***	0.010***	0.008***	0.012***	0.007***	0.003***	0.009***	0.002***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-1.169***	-1.123***	-1.078***	-1.163***	-1.072***	-0.989***	-1.141***	-0.975***	-0.830***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Num. of obs.	1000000	1000000	1000000	1000000	1000000	1000000	1000000	1000000	1000000
Pseudo R^2	0.001	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000

Note. Probit model where the dependent variable is $\mathbb{1}$(30-day post-discharge mortality). Standard errors in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In Table A.2, we provide additional regression results where we apply each of the two adjustments separately. We find that both adjustments make the coefficient estimates more positive, but the over-adjustment comes from addressing the in-hospital death bias.

From these simulation results, we conclude that while there are still biases in our coefficient estimates when introducing these adjustments, they are now in the direction which makes it more difficult to estimate the treatment effect we are interested in. That is, the coefficients are biased in the positive direction, while we hypothesize that the treatment effect will result in a negative coefficient estimate.

Table A.2 Simulation Results 2: 30-Day Mortality

In-hospital death bias addressed by coding these patients $y_i = 0$ & Decreasing mortality hazard bias present									
p	5%	5%	5%	10%	10%	10%	20%	20%	20%
α	1	2	3	1	2	3	1	2	3
los	0.012***	0.009***	0.007***	0.010***	0.005***	0.001*	0.004***	-0.002***	-0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-1.165***	-1.116***	-1.069***	-1.145***	-1.052***	-0.967***	-1.096***	-0.933***	-0.788***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Num. of obs.	1000000	1000000	1000000	1000000	1000000	1000000	1000000	1000000	1000000
Pseudo R^2	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000

In-hospital death bias present & Decreasing mortality hazard bias addressed by using an adjusted time-window									
p	5%	5%	5%	10%	10%	10%	20%	20%	20%
α	1	2	3	1	2	3	1	2	3
los	-0.023***	-0.027***	-0.031***	-0.024***	-0.031***	-0.037***	-0.026***	-0.038***	-0.048***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Constant	-0.838***	-0.782***	-0.728***	-0.835***	-0.725***	-0.623***	-0.818***	-0.617***	-0.437***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Num. of obs.	866284	866284	866284	865871	865871	865871	864990	864990	864990
Pseudo R^2	0.003	0.003	0.004	0.003	0.004	0.006	0.003	0.006	0.010

Note. Probit model where the dependent variable is $\mathbb{1}$(30-day post-discharge mortality). Standard errors in parentheses. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

B. Additional Tables and Figures

Table A.3 Data Selection

Sample	Observations	% prior	% initial
All Admissions in 2000-2011, except for Dec 2011 admits/discharges	186,472,400	NA	100.0%
Excluding overlapping admissions	159,674,712	85.6%	85.6%
Excluding post-acute care	133,056,603	83.3%	71.4%
Excluding stays with hospital transfers	118,681,435	89.2%	63.6%
Excluding those in facilities not paid under PPS	112,949,185	95.2%	60.6%
Excluding stays that are not FFS	111,499,586	98.7%	59.8%
Excluding non-AMI patients	2,457,542	2.2%	1.3%
Excluding those admitted within 30 days of prior admission's discharge	2,365,168	96.2%	1.3%
Excluding hospitals with less than 25 visits	2,357,080	99.7%	1.3%
Excluding patients with inpatient service related discharge destinations	2,272,453	96.4%	1.2%
Excluding non-elderly admissions	2,033,282	89.5%	1.1%
Excluding those that left against medical advice	2,025,086	99.6%	1.1%
Excluding those with unknown race or not residing in the US	2,018,185	99.7%	1.1%
Excluding elective patients (including unknown elective status)	1,899,920	94.1%	1.0%
Excluding same day discharge	1,899,914	100.0%	1.0%
Excluding cost outliers	1,828,145	96.2%	1.0%
Excluding length of stays $\geq$ the 99th percentile (20 days)	1,808,889	98.9%	1.0%

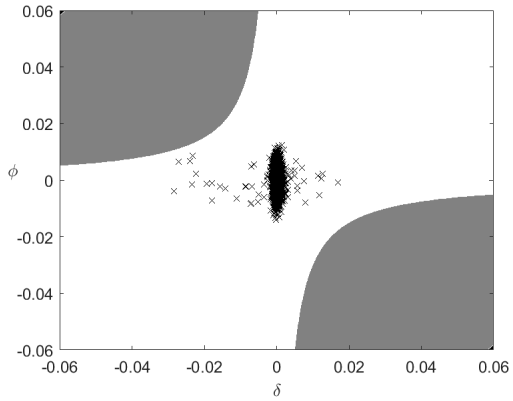
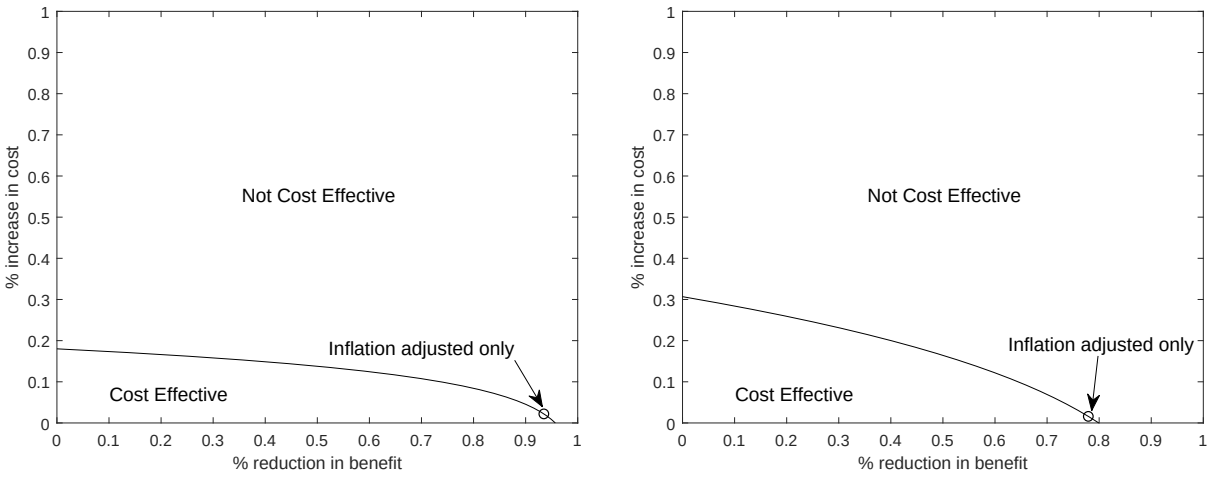
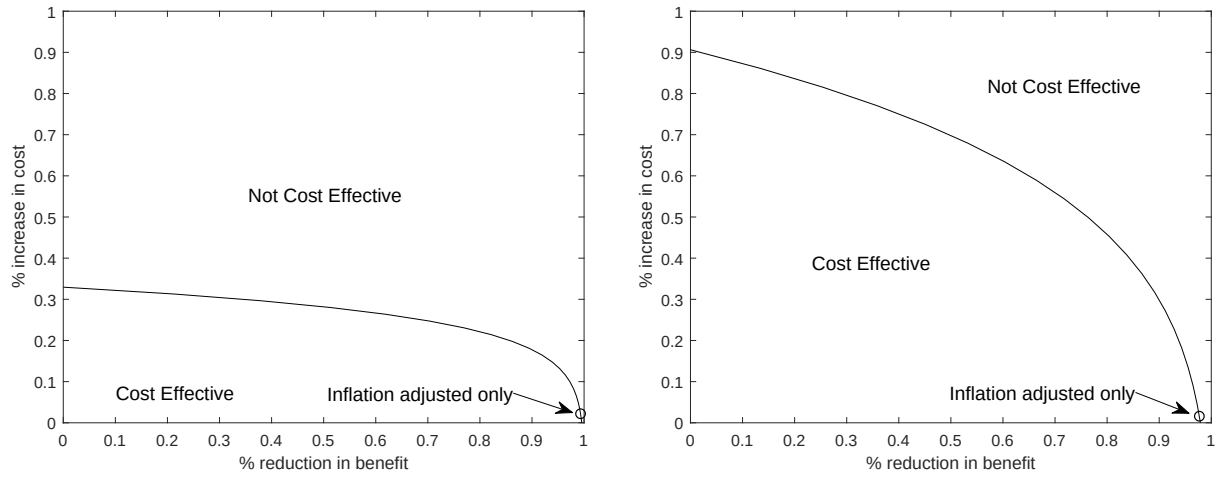


Figure A.2 **Sensitivity Analysis for the 30-day mortality model using predicted Saturday discharge as an IV. (1) White area represents regime for δ and ϕ where the estimates for θ are negative and statistically significant at the 5% level. (2) Gray area represents regime for δ and ϕ where the estimates for θ are statistically not different than 0 at the 5% level. (3) Black area represents regime for δ and ϕ where the estimates for θ are positive and statistically significant at the 5% level. x's represent δ and ϕ values for observed covariates.**



(a) Taheri et al. (2000) costs (b) The Henry J. Kaiser Family Foundation (2014) costs

Figure A.3 **Robustness of cost-effectiveness for keeping patients in the hospital an additional day, when treatment effect is based on the all days of the week IV (Table 4 Col 4). Baseline benefits are given by the value of living an additional month as estimated in Murphy and Topel (2006).**



(a) Taheri et al. (2000) costs (b) The Henry J. Kaiser Family Foundation (2014) costs

Figure A.4 **Robustness of cost-effectiveness for keeping patients in the hospital an additional day, when treatment effect is based on the predicted Saturday discharge IV (Table 4 Col 4). Baseline benefits are given by the value of living an additional month as estimated in Murphy and Topel (2006).**

Table A.4: Full results for column 4 in Table 3

	Coef.	Robust Std. Err.	p-value		Coef.	Robust Std. Err.	p-value
Second stage				First stage			
log(los)	-0.17	0.07	0.015	MonTue admit	-0.05	0.00	0.000
DRG (before 2007)				DRG (before 2007)			
105	-0.05	0.27	0.844	105	-0.04	0.04	0.267
106	0.09	0.08	0.252	106	0.06	0.01	0.000
107	0.09	0.06	0.134	107	0.03	0.01	0.005
108	0.09	0.10	0.357	108	-0.06	0.02	0.001
109	0.10	0.12	0.412	109	-0.07	0.01	0.000
110	0.25	0.07	0.000	110	-0.60	0.01	0.000
111	0.04	0.14	0.772	111	-0.58	0.02	0.000
112	0.11	0.09	0.190	112	-0.74	0.01	0.000
113	1.16	0.12	0.000	113	0.03	0.04	0.425
114	0.77	0.18	0.000	114	-0.05	0.05	0.236
115	0.49	0.07	0.000	115	-0.35	0.01	0.000
116	0.06	0.08	0.447	116	-0.75	0.01	0.000
117	0.63	0.28	0.023	117	-0.63	0.09	0.000
118	0.46	0.16	0.005	118	-0.49	0.04	0.000
119	0.92	0.42	0.026	119	-0.45	0.18	0.014
120	0.85	0.07	0.000	120	-0.26	0.02	0.000
121	0.54	0.08	0.000	121	-0.72	0.01	0.000
122	0.40	0.08	0.000	122	-0.73	0.01	0.000
124	0.53	0.21	0.012	124	-0.68	0.05	0.000
125	0.04	0.28	0.885	125	-0.81	0.04	0.000
144	0.49	0.10	0.000	144	-0.84	0.02	0.000
145	0.50	0.15	0.001	145	-0.90	0.04	0.000
468	0.71	0.07	0.000	468	-0.14	0.02	0.000
476	0.47	0.20	0.022	476	0.05	0.04	0.285
477	0.52	0.09	0.000	477	-0.37	0.02	0.000
478	0.63	0.08	0.000	478	-0.31	0.02	0.000
479	0.10	0.38	0.794	479	-0.54	0.06	0.000
483	0.84	0.12	0.000	483	0.12	0.03	0.000
514	0.27	0.11	0.015	514	-0.18	0.02	0.000
515	0.51	0.12	0.000	515	-0.25	0.03	0.000
516	0.06	0.08	0.444	516	-0.76	0.01	0.000
517	-0.11	0.38	0.778	517	-0.79	0.04	0.000
518	0.47	0.45	0.289	518	-0.52	0.15	0.001
525	-0.06	0.38	0.879	525	-0.49	0.08	0.000
526	-0.05	0.08	0.502	526	-0.80	0.01	0.000
527	0.25	0.41	0.536	527	-0.79	0.04	0.000
535	0.13	0.08	0.101	535	-0.18	0.01	0.000
541	0.63	0.24	0.009	541	0.31	0.07	0.000
542	0.79	0.22	0.000	542	0.12	0.08	0.116
547	-0.07	0.07	0.361	547	0.09	0.01	0.000
551	0.34	0.08	0.000	551	-0.33	0.01	0.000
553	0.67	0.08	0.000	553	-0.29	0.02	0.000
555	0.06	0.08	0.458	555	-0.77	0.01	0.000
557	-0.10	0.08	0.235	557	-0.81	0.01	0.000
DRG (after 2007)				DRG (after 2007)			
3	0.41	0.20	0.045	3	0.25	0.07	0.001
4	1.20	0.20	0.000	4	0.28	0.09	0.002
216	0.15	0.10	0.112	216	0.19	0.02	0.000
217	0.11	0.16	0.501	217	0.24	0.03	0.000
219	0.04	0.38	0.919	219	0.14	0.08	0.099
220	0.74	0.45	0.102	220	0.12	0.12	0.326
222	0.36	0.09	0.000	222	-0.02	0.02	0.380
223	-0.34	0.21	0.103	223	-0.29	0.02	0.000

226	0.57	0.16	0.000	226	-0.12	0.03	0.000
227	-0.05	0.41	0.902	227	-0.44	0.06	0.000
228	-0.49	0.34	0.145	228	0.22	0.03	0.000
229	-0.29	0.35	0.408	229	0.15	0.03	0.000
231	0.16	0.13	0.215	231	0.21	0.02	0.000
232	-0.07	0.16	0.658	232	0.20	0.02	0.000
233	0.08	0.08	0.270	233	0.28	0.02	0.000
234	-0.04	0.08	0.587	234	0.20	0.02	0.000
235	0.37	0.16	0.022	235	0.22	0.03	0.000
236	-0.07	0.25	0.792	236	0.12	0.02	0.000
237	0.47	0.07	0.000	237	-0.41	0.02	0.000
238	0.28	0.09	0.002	238	-0.59	0.02	0.000
239	1.08	0.25	0.000	239	0.38	0.06	0.000
242	0.54	0.08	0.000	242	-0.18	0.02	0.000
243	0.32	0.10	0.001	243	-0.31	0.02	0.000
244	0.29	0.15	0.054	244	-0.43	0.03	0.000
245	0.99	0.27	0.000	245	-0.38	0.08	0.000
246	0.25	0.07	0.001	246	-0.54	0.02	0.000
247	-0.20	0.09	0.018	247	-0.88	0.02	0.000
248	0.36	0.07	0.000	248	-0.49	0.02	0.000
249	-0.04	0.08	0.607	249	-0.83	0.02	0.000
250	0.51	0.08	0.000	250	-0.46	0.02	0.000
251	0.04	0.09	0.636	251	-0.81	0.02	0.000
252	0.69	0.08	0.000	252	-0.16	0.02	0.000
253	0.42	0.14	0.003	253	-0.40	0.04	0.000
254	0.19	0.41	0.645	254	-0.67	0.07	0.000
255	0.89	0.36	0.014	255	0.24	0.09	0.006
256	1.00	0.54	0.064	256	-0.02	0.29	0.950
258	0.58	0.28	0.040	258	-0.31	0.08	0.000
260	1.09	0.42	0.010	260	-0.14	0.15	0.376
261	0.92	0.55	0.093	261	-0.27	0.18	0.122
264	0.85	0.09	0.000	264	-0.16	0.03	0.000
280	0.73	0.07	0.000	280	-0.56	0.02	0.000
281	0.43	0.08	0.000	281	-0.75	0.02	0.000
282	0.26	0.09	0.004	282	-0.90	0.02	0.000
286	0.35	0.29	0.236	286	-0.45	0.08	0.000
314	0.72	0.11	0.000	314	-0.63	0.04	0.000
315	0.59	0.14	0.000	315	-0.80	0.04	0.000
316	0.55	0.16	0.001	316	-0.87	0.05	0.000
981	0.77	0.10	0.000	981	0.06	0.03	0.040
982	0.69	0.21	0.001	982	-0.15	0.06	0.015
983	0.84	0.37	0.023	983	-0.68	0.12	0.000
984	0.88	0.33	0.008	984	0.18	0.09	0.055
987	0.76	0.12	0.000	987	-0.08	0.04	0.042
988	0.16	0.23	0.491	988	-0.26	0.06	0.000
Elixhauser 1	0.23	0.01	0.000	Elixhauser 1	0.22	0.00	0.000
Elixhauser 2	0.04	0.01	0.000	Elixhauser 2	0.07	0.00	0.000
Elixhauser 3	0.04	0.01	0.000	Elixhauser 3	0.05	0.00	0.000
Elixhauser 4	-0.01	0.01	0.495	Elixhauser 4	0.02	0.00	0.000
Elixhauser 5	0.05	0.01	0.000	Elixhauser 5	-0.01	0.00	0.004
Elixhauser 6	-0.13	0.00	0.000	Elixhauser 6	-0.04	0.00	0.000
Elixhauser 7	0.57	0.02	0.000	Elixhauser 7	0.23	0.01	0.000
Elixhauser 8	0.17	0.01	0.000	Elixhauser 8	-0.01	0.00	0.000
Elixhauser 9	0.08	0.01	0.000	Elixhauser 9	0.09	0.00	0.000
Elixhauser 10	0.03	0.00	0.000	Elixhauser 10	0.01	0.00	0.000
Elixhauser 11	0.07	0.01	0.000	Elixhauser 11	0.12	0.00	0.000
Elixhauser 12	-0.06	0.01	0.000	Elixhauser 12	-0.02	0.00	0.000
Elixhauser 13	0.19	0.01	0.000	Elixhauser 13	0.05	0.00	0.000
Elixhauser 14	0.13	0.02	0.000	Elixhauser 14	-0.03	0.01	0.000

Elixhauser 15	-0.05	0.02	0.047	Elixhauser 15	0.06	0.01	0.000
Elixhauser 16	0.23	0.13	0.073	Elixhauser 16	0.07	0.04	0.055
Elixhauser 17	0.25	0.02	0.000	Elixhauser 17	0.04	0.01	0.000
Elixhauser 18	0.57	0.01	0.000	Elixhauser 18	0.05	0.01	0.000
Elixhauser 19	0.32	0.01	0.000	Elixhauser 19	0.03	0.00	0.000
Elixhauser 20	0.03	0.01	0.036	Elixhauser 20	0.03	0.00	0.000
Elixhauser 21	0.08	0.01	0.000	Elixhauser 21	0.11	0.00	0.000
Elixhauser 22	-0.13	0.01	0.000	Elixhauser 22	0.00	0.00	0.088
Elixhauser 23	0.45	0.01	0.000	Elixhauser 23	0.21	0.00	0.000
Elixhauser 24	0.20	0.01	0.000	Elixhauser 24	0.13	0.00	0.000
Elixhauser 25	0.05	0.02	0.028	Elixhauser 25	0.23	0.00	0.000
Elixhauser 26	-0.04	0.02	0.013	Elixhauser 26	0.11	0.00	0.000
Elixhauser 27	0.00	0.02	0.999	Elixhauser 27	0.09	0.01	0.000
Elixhauser 28	0.06	0.03	0.066	Elixhauser 28	0.22	0.01	0.000
Elixhauser 29	0.17	0.02	0.000	Elixhauser 29	0.15	0.01	0.000
Elixhauser 30	0.10	0.01	0.000	Elixhauser 30	0.04	0.00	0.000
Elixhauser 31	-0.10	0.01	0.000	Elixhauser 31	0.02	0.00	0.000
Age spline 1	0.02	0.00	0.000	Age spline 1	0.01	0.00	0.000
Age spline 2	0.03	0.00	0.000	Age spline 2	0.01	0.00	0.000
Age spline 3	0.03	0.00	0.000	Age spline 3	0.00	0.00	0.000
Age spline 4	0.02	0.00	0.000	Age spline 4	-0.01	0.00	0.000
Female	-0.01	0.01	0.064	Female	0.05	0.00	0.000
Race				Race			
Black	-0.04	0.01	0.000	Black	0.06	0.00	0.000
Hispanic	-0.03	0.02	0.082	Hispanic	0.02	0.00	0.000
Other	-0.10	0.01	0.000	Other	0.02	0.00	0.000
Surgery	-0.12	0.01	0.000	Surgery	0.12	0.00	0.000
Intensive care use				Intensive care use			
No	0.02	0.02	0.379	No	-0.16	0.01	0.000
General	0.02	0.02	0.428	General	-0.02	0.01	0.001
Surgery	0.03	0.03	0.177	Surgery	-0.03	0.01	0.000
Medical	0.02	0.02	0.372	Medical	-0.01	0.01	0.338
Intermediate	0.02	0.02	0.416	Intermediate	0.03	0.01	0.000
Adm month				Adm month			
2	0.00	0.01	0.807	2	0.00	0.00	0.939
3	-0.02	0.01	0.050	3	-0.01	0.00	0.003
4	-0.03	0.01	0.001	4	-0.03	0.00	0.000
5	-0.03	0.01	0.001	5	-0.04	0.00	0.000
6	-0.03	0.01	0.000	6	-0.04	0.00	0.000
7	-0.02	0.01	0.024	7	-0.04	0.00	0.000
8	-0.03	0.01	0.001	8	-0.04	0.00	0.000
9	-0.02	0.01	0.030	9	-0.04	0.00	0.000
10	-0.01	0.01	0.188	10	-0.03	0.00	0.000
11	-0.03	0.01	0.014	11	-0.03	0.00	0.000
12	-0.02	0.01	0.092	12	-0.04	0.00	0.000
Adm year				Adm year			
2001	0.02	0.01	0.071	2001	-0.02	0.00	0.000
2002	0.04	0.01	0.000	2002	-0.03	0.00	0.000
2003	0.06	0.01	0.000	2003	-0.04	0.00	0.000
2004	0.07	0.01	0.000	2004	-0.06	0.00	0.000
2005	0.08	0.01	0.000	2005	-0.08	0.00	0.000
2006	0.08	0.01	0.000	2006	-0.11	0.00	0.000
2007	0.06	0.02	0.000	2007	-0.15	0.01	0.000
2008	0.04	0.02	0.100	2008	-0.18	0.01	0.000
2009	0.03	0.03	0.183	2009	-0.22	0.01	0.000
2010	0.01	0.03	0.635	2010	-0.26	0.01	0.000
2011	-0.02	0.03	0.533	2011	-0.33	0.01	0.000
Hospital FE (over 3000)				Hospital FE (over 3000)			
smallest coef	-0.99	0.35	0.005	smallest coef	-0.74	0.22	0.001

largest coef	1.21	0.38	0.002	largest coef	0.52	0.06	0.000
Constant	-2.28	0.19	0.000	Constant	2.04	0.02	0.000