The Spillover Effects of Hospital Closures on the Efficiency and Quality of Other Hospitals

Lina D. Song¹, Soroush Saghaian²

¹PhD Program in Health Policy, Harvard University, Cambridge, MA 02138
²Harvard Kennedy School, Harvard University, Cambridge, MA 02138

Problem definition: The recent trend in the U.S. hospital closures can have important impacts on the healthcare sector by changing the operational efficiency and quality of care of the remaining hospitals. We investigate the impact of hospital closures on the surrounding hospitals’ operational efficiency and quality, and study how such hospitals respond to the closure of their neighboring hospital.

Academic/Practical Relevance: Understanding how hospital closures impact the way the remaining hospitals operate can allow policymakers to utilize more effective policy levers in order to mitigate the negative consequences of hospital closures.

Methodology: We analyze more than 14 million inpatient visits made during 11 years to over 3,000 hospitals in the U.S. (before and after various closures), and utilize causal methods to evaluate the spillover effect of hospital closures on the nearby hospitals. We also conduct counterfactual analyses to evaluate policy interventions that could have been used by policymakers.

Results: Hospital closures have both positive and negative spillover effects. When a hospital closes, its nearby hospitals improve their operational efficiency. However, they do so via a speed-up response (i.e., by reducing their service durations to accommodate the increased demand) instead of an effort to lower their average bed idle time. This speed-up response negatively affects some important aspects of the care quality provided, including the 30-day mortality rate. The spillover effect of a hospital closure is highly heterogeneous: hospitals in markets where patients have limited choices of hospitals (e.g., less competition, fewer resources) and hospitals that are more desirable (e.g., high-quality, urban, teaching, and large) tend to experience greater spillover effects.

Managerial Implications: Our analyses suggest two effective policy levers: (a) bailing out specific hospitals (e.g., rural or less desirable than neighbors) from potential closures, and (b) eliminating the speed-up response of specific hospitals (e.g., rural or more desirable hospitals). In addition to helping policymakers, our results help hospital administrators: our findings help them to better understand the consequences (or the absence) of their strategic responses to a neighboring hospital closure, and thereby, adopt more suitable management strategies.

Key words: Hospital Closures; Healthcare Operations Management; Healthcare Quality; Hospital Efficiency

History: Version: 2020-01-13

1. Introduction

A substantial number of U.S. hospitals have closed in the past decade (Kaufman et al. 2016, Friedman et al. 2016, MedPAC 2017). Such closures have occurred widely across the nation affecting a large number of people (see, e.g., Figure 1) (MedPAC 2017). Since U.S. hospitals are facing multiple challenges, including decreasing demand for inpatient services and elevated financial pressures
and uncertainty over the payment structure, the number of hospital closures is expected to rise (Kaufman et al. 2016, Bazzoli et al. 2014, Wertheim and Lynn 1993, Navigant 2019).

The increasing risk of hospital closures has fueled a debate on the need to implement policies that financially support, or “bail out,” hospitals that are on the verge of closure. Proponents of such policies argue that hospital closures can aggravate the already prevalent barriers to access (Hsia et al. 2012, Liu et al. 2014, Buchmueller et al. 2006). Others point out that hospital closures are advantageous, and hence, should not be prevented: there are many inefficient or low-quality hospitals in the current system, and closing such hospitals can improve the healthcare system (Keeler and Ying 1996).

An essential but missing piece of information in this debate that can yield a better understanding of the implications of hospital closures is how they affect the efficiency and/or quality of care of the remaining hospitals. After a hospital closes, patients and payers have to rely on the remaining hospitals for care delivery. Therefore, hospital closures may alter patient demand and patient mix for the remaining hospitals, and the remaining hospitals may adjust their care delivery processes accordingly. However, this potential spillover effect of hospital closures on the remaining hospitals has been largely ignored in the literature. In this paper, we examine this potential spillover effect: we study whether and how hospital closures impact (a) the operational efficiency, and (b) the quality of care in nearby hospitals, and shed light on their important societal implications.

Evidence shows that, after a hospital closes, the nearby hospitals experience an increase in patient demand (Lindrooth et al. 2003, Capps et al. 2010). Whether this increased demand translates into an improvement in operational efficiency, however, is unknown. This is because the remaining hospitals can behave strategically in response to the change in demand. For example, they may expand their capacity instead of serving more patients with their current capacity level. Even if a hospital treats more patients without increasing capacity, it can accommodate the increase in demand with various other mechanisms (e.g., by increasing service or bed utilization rate).

Similarly, the impact of hospital closures on the quality of care delivered at other hospitals is unclear. On the one hand, hospital closures may improve the quality of the remaining hospitals. This can occur for at least two reasons. First, the remaining hospitals may put greater effort into improving their quality to be able to stay in the competition and prevent themselves from being closed. Second, due to learning-by-doing and related positive effects of volume on quality known as “volume-outcome effect” or “productivity spillover” (Ramanarayanan 2008, Chandra and Staiger 2007, Birkmeyer et al. 2002), the increase in patient volume may allow the remaining hospitals to
improve their quality. On the other hand, hospital closures may reduce the quality of the remaining hospitals. This can happen, for example, if the remaining hospitals experience high congestions and care delivery delays due to the increase in patient demand, or if they decide to cut some of their value-adding treatment processes to accommodate the demand spike (Haas et al. 2018, Chan et al. 2016). Thus, in addition to studying the consequences of closures on the operational efficiency of the surrounding hospitals, we provide some evidence and insights into whether and how hospital closures affect the quality of care of the remaining hospitals.

1.1. Framework

We focus on the U.S. short-term acute-care hospitals that serve patients with acute severe injury or episodes of illness. To investigate the impact of hospital closures on the operational efficiency of the nearby hospitals (our first goal), we define operational efficiency as a measure of how much output is produced per input. Specifically, we consider throughput per bed (i.e., the average number of patients served per bed per unit of time) as our measure of operational efficiency. We focus on beds as the main hospital resource, given that empty beds are the major contributors to the low operational efficiency of U.S. hospitals (Keeler and Ying 1996, Gaynor and Anderson 1995). Next, to identify the mechanism through which a change in operational efficiency might occur, we investigate changes in bed utilization and service duration separately. To study the impact of hospital closures on quality of care in the remaining hospitals (our second goal), we consider measures such as patient experience, 30-day readmission rates, and 30-day mortality rates, all of which are widely-accepted measures for hospital care quality (Benbassat and Taragin 2000, Tourangeau et al. 2007). For both our first and second goals, we also identify the heterogeneous effects of closures by the market and hospital characteristics, and generate insights into variations in the closure effect that might elevate care disparities in the society. Finally, we perform counterfactual analyses to shed light on policies that, had been implemented by policymakers or hospital administrators, could yield significant societal gains.
1.2. Data and Empirical Challenges

There are several empirical challenges for studying the impact of hospital closures. First, although studies have examined rural hospital closures or closures in specific geographic areas (Kaufman et al. 2016, Lindrooth et al. 2003, Capps et al. 2010, Carroll 2019, Gujral and Basu 2019), no public data keeps track of all U.S. hospital closures. Thus, we have independently identified closed hospitals through our own research and various validation steps. As an example, we first identified potentially closed hospitals through Medicare Provider of Service (POS) data and Medicare fee-for-service (FFS) claims data and verified each closure separately through multiple sources including local news, state department documents, or findings from research institutions. Next, we used a nationally representative, multi-year patient, hospital, and area level data to improve the generalizability of our findings. Specifically, we used the 20% sample of Medicare FFS claims data, which cover the majority of the U.S. population in ages 65 and older. We then linked the claims dataset to our hospital level data (from POS, Hospital Compare, and the Hospital Consumer Assessment of Healthcare Providers and Systems, or HCAPHS) and area level data (from Area Health Resource Files, or AHRF, and the Centers for Medicare & Medicaid Services (CMS) State/County/Plan Enrollment Data). Our final dataset, after combining these data sources (and cleaning them), includes over 14 million inpatient visits made during 11 observation years to 146 closed and 3,299 open hospitals (Table 1).

Another empirical challenge in studying the impact of hospital closures is that hospital closure can be correlated with both the patient characteristics and the market structure. Such endogeneity can bias the estimate of the hospital closure effect on the outcomes of our interest. To overcome this challenge, we use an extensive set of covariates, including patient level clinical and demographic information, as well as hospital and area level information such as provider supply, concentration, and the insurance market structure. As a main empirical strategy, we utilize the substantial geographic variation and the timing of U.S. hospital closures along with a multilevel panel data to make use of difference-in-difference (DID) analysis with the hospital, area, and year fixed effects. The DID analysis has enabled us to use both cohort and time dimensions, and thereby, adjust for time-invariant unobserved confounders. We use a matched sample that improves the comparability of the comparison groups. In addition to our primary analysis, we check the robustness of our findings by making use of various mechanisms, including an instrumental variable (IV) analysis. Our IV analysis utilizes the state level variation in the decision to expand Medicaid after the Affordable Care Act (ACA) to address the potential time-varying unobserved heterogeneity.
Table 1: Overview of Our Final Data

<table>
<thead>
<tr>
<th></th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitals</td>
<td>3,299</td>
</tr>
<tr>
<td>Closed hospitals</td>
<td>146</td>
</tr>
<tr>
<td>Patients</td>
<td>4,645,532</td>
</tr>
<tr>
<td>Inpatient visits</td>
<td>14,147,180</td>
</tr>
<tr>
<td>Number of observation years</td>
<td>11</td>
</tr>
</tbody>
</table>

1.3. Main Findings

Our results show that hospital closures have both positive and negative effects on the nearby hospitals—and hence, the healthcare sector as a whole. On the positive side, when a hospital closes, the nearby hospitals experience an increase in patient volume, which translates into an improvement in their operational efficiency (i.e., the number of patients treated per unit of capacity). Interestingly, however, we find that their bed utilization rates remain relatively constant, whereas their service duration significantly decreases. This implies that a speed-up behavior (i.e., improving efficiency via reducing service duration) rather than an effort to lower the average bed idle time is the primary mechanism through which nearby hospitals improve their operational efficiency. Because of such behavior, the gains in efficiency have negative implications on some (but not all) aspects of care quality. In particular, we find that when a hospital closes, the 30-day mortality rate of its nearby hospitals substantially increases. We also observe that the effect of hospital closures is highly heterogeneous and largely depends on market and hospital characteristics. For example, closure effects are concentrated in the markets where patients have limited choice of hospitals (e.g., in rural areas). The effect also highly varies by the characteristics of the neighbor hospital. Specifically, hospitals with more desirable characteristics (e.g., large, teaching, urban, or high-quality hospitals) experience greater spillover effects.

Taken together, our results suggest that although hospital closures are effective at improving the operational efficiency of the remaining (and nearby) hospitals, these hospitals do not necessarily improve their operational efficiency in the most desirable way. In particular, the improvement in operational efficiency—serving more patients per unit of capacity—is not due to more effective use of beds, but rather due to spending less time on each patient. Spending less time on each patient may not necessarily be undesirable if it eliminates non-value adding procedures. However, our analyses show that it typically translates to a higher 30-day mortality rate. This suggests that at least some of the value-adding procedures are eliminated as a result of the speed-up behavior. Furthermore, our results on the heterogeneous effect of closures suggest that hospital closures widen social disparities: the adverse consequences, unfortunately, fall disproportionately among hospitals in areas with a limited choice of hospitals. It also increases the efficiency gap between the hospitals
that are either able to accommodate greater patient volume or are more likely to be demanded by patients and the rest.

Finally, our counterfactual analyses aimed at providing policy recommendations indicate that targeted versions of (a) bailouts, and (b) monitoring and regulations of service durations are effective policy levers. For example, bailing out the hospitals in markets with limited patient choice (e.g., rural hospitals) or those that have less desirable characteristics than their neighbors typically yields larger gains in quality than bailing out other hospitals. Focusing monitoring and regulation efforts on the hospitals in markets with limited patient choice or more desirable hospitals such as urban, non-profit, and teaching hospitals post-closure can also help policymakers take advantage of the positive effects of hospital closures and mitigate their negative consequences. However, we conclude with a caveat that these targeted strategies typically come at a higher financial cost and can yield wider social disparities than their non-targeted versions.

1.4. Main Contributions

The contributions of our study are three-fold. First, our results have important policy implications for the ongoing debate on the U.S. healthcare delivery system reform. Recently, some policymakers have argued against the payment mechanisms that support inefficient or financially unsustainable hospitals that rely on government subsidies such as the Medicaid Disproportionate Share Hospitals (DSH) payments for hospitals with high proportions of uncompensated care, or the Critical Access Hospital (CAH) status for hospitals with a high fixed cost. Yet, the limited evidence on the implications of hospital closures has impeded clear policymaking. We draw rigorous evidence from nationally representative data and provide actionable policy suggestions. Second, our results help hospital administrators to understand the implications of nearby hospitals’ closures better, and accordingly revise their management strategies that may adopt in response to such closures. Third, by studying how hospitals respond to demand shocks, our work contributes to the operations management literature that aims at understanding organizational level responses to demand spikes.

2. Conceptual Framework and Hypotheses

2.1. Changes in Operational Efficiency

We first examine if the closure of a hospital results in an improvement in the operational efficiency of its neighboring hospitals. As noted earlier, studies show that hospital closures can result in increased demand at the remaining hospitals in the area (Lindrooth et al. 2003, Capps et al. 2010, Bazzoli et al. 2003). Yet, it is not clear whether the increase in demand is consistent nationally.
More importantly, it is not clear whether this increase in demand translates to improved operational efficiency at nearby hospitals. Thus, we start by using our data to test the following:

**Hypothesis 1.** Closure of a hospital increases the operational efficiency of the nearby hospitals.

Our first hypothesis will enable us to test whether or not the closures have an increasing overall impact on the operational efficiency of the neighboring hospitals. However, it will not generate a detailed understanding of whether such an impact is favorable or unfavorable. For instance, improvement in operational efficiency might not stem from increasing bed utilization rate (i.e., accommodating the increased demand by lowering bed idle times). Given the low bed utilization rate of U.S. hospitals and the high cost of empty hospital beds, reducing bed idle time is a socially desirable way of improving operational efficiency.

In particular, conceptualizing a hospital as a general queueing system with $s$ beds that play the role of servers, the throughput per bed $\lambda/s$—our measure of operational efficiency—can be expressed as $\lambda/s = \rho + \mu$, where $\rho$ is the bed utilization, and $\mu$ is the service rate such that $1/\mu$ is the expected service duration. This implies that if $s$ remains constant—hospitals typically cannot change their number of beds in the short term for various reasons, including the lengthy regulatory processes such as Certificate of Need (CON)—an increase in throughput corresponds to either (a) an increase only in $\rho$, (b) a decrease only in $1/\mu$, or (c) both an increase in $\rho$ and a decrease in $1/\mu$. Thus, to gain a deeper understanding of potential mechanisms through which operational efficiency might be affected, we form two more hypotheses: one with respect to bed utilization ($\rho$) and one with respect to service duration ($1/\mu$). We next discuss each of these separately.

**Bed Utilization.** U.S. hospitals have a wide range in their existing bed utilization rate. Although this can be partially described by differences in average patient demand, there is a good level of discretion on a hospital to set its bed occupancy rate for reasons beyond current average demand (Joskow 1980, Green and Nguyen 2001). Similarly, post-closure, nearby hospitals may accommodate the increased demand by either increasing their bed utilization rate or keeping their utilization rate the same, and instead responding to the closure in a different way. To gain insights into the effect of hospital closure on the bed utilization rate, we examine the following:

**Hypothesis 2.** Closure of a hospital increases the bed utilization rate of the nearby hospitals.

**Service Duration.** As an alternative strategy in response to increased patient demand, hospitals may decrease their service duration. Operations management literature suggests that servers can be strategic about the service duration under financial and nonfinancial motivations, and alter their behavior based on the characteristics of the queue (Cachon and Zhang 2007, Debo et al.
2008, Hopp et al. 2007, Jouini et al. 2008, Tan and Netessine 2014, Oliva and Sterman 2001). In particular, the visibility of the queue length or the server occupancy rate can encourage a speed-up behavior by servers (KC and Terwiesch 2012, Batt and Terwiesch 2012, Shunko et al. 2017). In our setting, providers do not have full visibility of the entire queue. Furthermore, unlike the above-mentioned literature, we are interested in studying changes at the organization (i.e., hospital) level as opposed to individual servers. Hence, it is not clear whether and to what extent the overall average service duration at a nearby hospital will change. To examine whether or not a hospital (as a whole) responds to the increased demand caused by a nearby hospital closure by decreasing its service duration, we test the following:

**Hypothesis 3.** Closure of a hospital decreases the service duration of the nearby hospitals.

### 2.2. Changes in Quality of Care

As noted earlier, it is unclear whether the effect of hospital closures on the quality of care of the neighboring hospital is positive or negative. On the one hand, congested hospital systems are typically more vulnerable to provider errors and often less able to respond to a patient promptly. Hence, a more congested hospital can result in an added risk to patient safety (KC and Terwiesch 2009, Weissman et al. 2007, Haas et al. 2018) as well as reduced patient satisfaction (Thompson et al. 1996). On the other hand, an increase in patient volume can result in a better quality of care from learning and specialization. This is due to the well-established “volume-outcome” or “productivity spillover” effects, which refer to the fact that healthcare providers improve their quality of care and patient outcomes with increased experience (Birkmeyer et al. 2002, Ramanarayanan 2008, Chandra and Staiger 2007).

To gain a deeper understanding of the effect of hospital closures on the quality of care delivered at nearby hospitals, we mainly focus on two dimensions of quality: patient experience and patient health outcomes. We examine the first dimension—patient experience—using a national survey of the inpatient care experience. For the second dimension—patient health outcomes—we measure 30-day readmission rates as well as 30-day mortality rates, the two widely-used outcome measures for inpatient services (Benbassat and Taragin 2000, Tourangeau et al. 2007). These enable us to test the following:

**Hypothesis 4.** Closure of a hospital changes the patient experience of the nearby hospitals.

**Hypothesis 5.** Closure of a hospital changes the 30-day readmission rate of the nearby hospitals.
Hypothesis 6. Closure of a hospital changes the 30-day mortality rate of the nearby hospitals.

3. Related Studies

Our work is related to the stream of literature that examines the relationship between the provider market structure (e.g., the provider supply and the regulatory environment) and the efficiency of the healthcare delivery system (Bates et al. 2006, Rosko and Mutter 2014). Within this stream, some studies have investigated the impact of changes in the supply due to an exit of providers from the market on the remaining providers’ efficiency. Lindrooth et al. (2003), Capps et al. (2010) show that urban hospital closure improves the economic efficiency (i.e., costs per admission) of nearby hospitals through an increase in inpatient admissions, suggesting the existence of economies of scale. Yet, Hodgson et al. (2015) suggests that hospital closures do not generate economies of scale, but merely shifts the high-cost patients to the remaining hospitals. A recent study provides a nuanced perspective by arguing that there are positive or negative economies of scale effect depending on the type of services offered at a hospital (Freeman et al. 2018). Our study is related to this stream of research: we examine the mechanisms through which hospitals’ responses to a sudden change in patient demand affect the economies of scale. However, unlike the studies mentioned above, instead of focusing on cost measures, we study the implications on operational efficiency and quality of care.

Our study is also relevant to the literature on the impact of provider market structure on healthcare quality. Among studies in this vein, our work mainly contributes to those that examine how the reduction in healthcare resources affects the quality of care. A body of literature has studied the impact of hospital or emergency department closure on access to care or health outcomes for the population in the area (Joynt et al. 2015, Buchmueller et al. 2006, Capps et al. 2010, Hsia et al. 2012, Liu et al. 2014, Carroll 2019, Gujral and Basu 2019). These studies show mixed evidence on the impact of hospital closures on patient outcomes. For example, Buchmueller et al. (2006) shows that hospital closures increase mortality from heart attacks and unintentional injuries, whereas Joynt et al. (2015) shows that there are no significant changes in mortality. Gujral and Basu (2019) and Carroll (2019) show that rural hospital closures resulted in an increase in mortality for time-sensitive conditions. While Buchmueller et al. (2006), Carroll (2019), Gujral and Basu (2019) demonstrate that hospital closures can have a negative consequence on patient health, they focus on the dimension of access by measuring the relationship between the increased travel distance (as a result of hospital closures) or loss of primary hospital and patient health. Unlike these studies that focus on the effect of hospital closure on quality as a result of the changes in access, we examine how
patients who do not directly lose access are affected through responses of the remaining hospitals. Notably, our findings have implications both for the patients who directly lose access as well as those who are indirectly affected through the spillover effects that occur at nearby hospitals.

4. Data and Study Sample

4.1. Data
We obtained patient, hospital, and area level information by linking various data sources that include information for our study period (years 2005-2015). Our hospital level information is from Medicare Cost Reports and POS data. The data are collected from hospitals that serve Medicare patients and contain information on facility characteristics, healthcare use, and cost. For patient level information, we used a panel data of FFS Medicare inpatient claims. Medicare inpatient claims provide information on all FFS inpatient services use, the types of procedures performed, and diagnosis through the International Classification of Diseases (ICD-9) and Healthcare Common Procedure Coding System (HCPCS). We benefitted from the Medicare Beneficiary Summary Files (BSF) to identify individual beneficiary’s sociodemographic information. We made use of the Hospital Compare data from the CMS for hospital quality, the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) for patient experience, AHRF for the county level information on health services resources and demographic information, and CMS State/County/Plan Enrollment Data File for the insurance enrollment information.

4.2. Identifying Hospital Closures
We defined hospital closure as ceasing to deliver short-term general hospital services rather than the changes in the ownership or physical appearance of a hospital because we focus on the capacity pooling effect of hospital beds. If a hospital remained in the same physical location but ceased to provide short-term acute care and converted to a different use such as an emergency department, rehabilitation facility, or long-term care facility, we regarded it as closure. However, absent such changes, if a hospital merely changed its name or ownership but stayed in the same physical location, we considered the hospital to be in operation.

To identify potential closures, we first used Medicare POS data. We regarded a hospital as potentially closed if it did not appear in the POS data after a specific year. Because virtually all short-term acute care hospitals in the United States, except military hospitals, participate in the Medicare program, discontinuation of data or claims submission to the CMS suggests it is highly likely that the hospital has experienced changes in operating and/or ownership status. Next, we identified a hospital as potentially closed if its number of hospitalizations from Medicare inpatient
claims has dropped to zero. Among the list of potential closures that met either of our criteria, we then excluded the hospitals that are not short-term acute care hospitals. Finally, we systematically searched and validated each hospital’s operating status through multiple external sources, including local news, state department documents, or a list of rural hospital closures from other research institutions. These sources include a list of rural hospital closures from the University of North Carolina Rural Health Research and Policy Analysis Center and Becker’s Hospital Review. In a few cases where the evidence was not available or definitive, we called the hospital directly. We were eventually able to confirm the operating status of all hospitals on our list. Figure 1 of the Online Appendix shows the detailed steps through which we determined hospital closures.

Table 2 shows the comparison of hospitals that were closed versus those that remained open during our study period. In general, the pre-closure characteristics of closed hospitals indicate clear signs of difficulty: compared to hospitals that did not close, closed hospitals are in more competitive markets (measured by Herfindahl-Hirschman Index). They are less likely to be teaching or nonprofit, more likely to be rural, and more likely to have low operational efficiency and/or quality of care. Their patients are slightly older, less likely to be White, and tend to be poorer and sicker.

4.3. Study Sample

Our study sample of hospitals included all the Medicare-participating U.S. hospitals that were in operations throughout our study period. Among our sample, we considered hospitals that are either in the same or the nearest zip codes within the same state as the closed hospital as “treated,” and the rest as “control.” We defined the nearby hospital this way because our data shows that the majority of the patients that originally visited closed hospitals select the hospitals in the nearest zip codes. In our robustness checks, we use hospital referral regions (HRRs) instead of zip codes to identify nearby hospitals. In all analyses, we removed the closed hospitals from our study sample.

For each hospital in the treatment group, we defined the index year as the year of a nearby hospital’s closure. We excluded the information from the index year to account for the noise during the transition period. We then compared outcomes pre- and post-closure. To be consistent, we limited our observations to five years post-closure since (a) we do not have enough observation years beyond that for the majority of closure index years and (b) there are other contemporaneous changes in the market such as additional closures or entry of hospitals beyond five years. However, in our sensitivity analysis, we analyzed the effect of using other durations for post-closure observation
Table 2: Hospital and Patient Characteristics of the Open and Closed Hospitals

<table>
<thead>
<tr>
<th></th>
<th>Open (n)</th>
<th>Closed (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitals (n)</td>
<td>3,299</td>
<td>146</td>
</tr>
<tr>
<td>Inpatient visits (n)</td>
<td>14,044,827</td>
<td>102,353</td>
</tr>
<tr>
<td>Patients (n)</td>
<td>4,593,211</td>
<td>52,321</td>
</tr>
</tbody>
</table>

**Hospital characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Open (%)</th>
<th>Closed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSH (%)</td>
<td>80.5</td>
<td>78.3</td>
</tr>
<tr>
<td>Teaching hospitals (%)</td>
<td>32.3</td>
<td>22.7</td>
</tr>
<tr>
<td>Ownership – nonprofit (%)</td>
<td>61.5</td>
<td>34.0</td>
</tr>
<tr>
<td>Ownership – private (%)</td>
<td>22.6</td>
<td>10.9</td>
</tr>
<tr>
<td>Ownership – public (%)</td>
<td>15.9</td>
<td>49.7</td>
</tr>
<tr>
<td>Rural (%)</td>
<td>33.4</td>
<td>59.7</td>
</tr>
<tr>
<td>Discharges (n)</td>
<td>9,724.4</td>
<td>3,466.7</td>
</tr>
<tr>
<td>Beds (n)</td>
<td>196.7</td>
<td>100.3</td>
</tr>
<tr>
<td>Operational efficiency</td>
<td>45.9</td>
<td>32.2</td>
</tr>
<tr>
<td>Bed utilization (%)</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>Patient experience (%)</td>
<td>67.4</td>
<td>63.3</td>
</tr>
</tbody>
</table>

**Market characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Open (%)</th>
<th>Closed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMO Penetration (%)</td>
<td>17.0</td>
<td>13.9</td>
</tr>
<tr>
<td>Herfindahl-Hirschman Index</td>
<td>1,451.4</td>
<td>1,317.0</td>
</tr>
<tr>
<td>Population (n)</td>
<td>1,041,501</td>
<td>706,902.6</td>
</tr>
<tr>
<td>Unemployment (%)</td>
<td>9.0</td>
<td>9.3</td>
</tr>
<tr>
<td>Poverty (%)</td>
<td>16.1</td>
<td>18.2</td>
</tr>
<tr>
<td>Under Age 65 (%)</td>
<td>83.0</td>
<td>82.5</td>
</tr>
<tr>
<td>Hosp (/10,000)</td>
<td>0.27</td>
<td>0.39</td>
</tr>
</tbody>
</table>

**Patient characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Open (%)</th>
<th>Closed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg age (years)</td>
<td>78.98</td>
<td>79.79</td>
</tr>
<tr>
<td>Sex – male (%)</td>
<td>42.08</td>
<td>37.33</td>
</tr>
<tr>
<td>Race – White (%)</td>
<td>87.06</td>
<td>81.44</td>
</tr>
<tr>
<td>Race – Black (%)</td>
<td>8.45</td>
<td>13.23</td>
</tr>
<tr>
<td>Dual-eligibles (%)</td>
<td>17.29</td>
<td>26.11</td>
</tr>
<tr>
<td>Comorbidity score</td>
<td>3.37</td>
<td>3.64</td>
</tr>
<tr>
<td>Chronic conditions (n)</td>
<td>23.72</td>
<td>21.94</td>
</tr>
</tbody>
</table>

Note. DSH indicates Disproportionate Share Hospitals. Dual-eligibles indicate Medicare-Medicaid dual eligibles. All differences in covariates between the two groups were statistically significant at p-value < 0.05.

years (see Table 8). Figure 2 in the Online Appendix shows our timeline of hospital closures for index years as well as pre- and post-closure observations.

For our patient level analysis, we considered the study population to be the FFS Medicare beneficiaries who paid at least one visit to the hospitals in our study sample. To improve the comparability, we further restricted the patient population to those who were aged 65 or older, did not have a disability, and were entitled to Medicare due to age (Although Medicare eligibility age is 65, Medicare also covers a small fraction of people under 65 with disabilities). Because patient’s treatment status was based on the treatment status of the hospital they visited, a patient was allowed to be in both treatment and control groups if s/he made multiple visits to both treatment and control hospitals. We excluded transfers to or from another hospital, admissions for rehabilitation, and emergency department visits that did not result in inpatient admissions from the analysis. The methods we used for identifying these patients (e.g., through ICD-9 codes) are explained in the Online Appendix. As noted earlier and shown in Table 1, there were a total of 4,645,532 patients in our final sample with a total of 14,147,180 inpatient visits during our study period (i.e., across 11 observation years) made to 3,299 open and 146 closed hospitals.

### 4.4. Dependent Variables

Our dependent variables include three operational measures (operational efficiency, bed utilization, and service duration) and three quality measures (patient experience, 30-day readmission, and 30-day mortality). Table 3 shows the definition of these variables, data sources used, and their corresponding hypotheses. For operational efficiency and bed utilization, we used the yearly average to address the seasonal and weekly variations in patient demand. For service duration, we used
data on length of stay, but excluded observations with values greater than 30 days, as our goal is to examine the impact of closures on short-term acute care. For measuring 30-day readmission, we considered inpatient claims that were within 30 days of a previous hospitalization’s discharge date. For measuring 30-day mortality, we obtained death information from the Medicare denominator files and calculated the time to death as the number of days between the index discharge date and the date of death. We linked a hospitalization to an incidence of 30-day mortality if death was present within 30 days of discharge.

To study patient experience, we used the HCAHPS survey data collected by the CMS. HCAHPS is a national publicly reported survey for patients’ perceptions of their hospital experience and is obtained by asking discharged patients questions about their hospital stay (see, e.g., Manary et al. (2013)). We used the overall patient’s rating (1 for lowest and 10 for highest) as a primary outcome measure. Because we do not observe individual level responses, we defined a hospital’s overall rating as the total percentage of patients who gave “high” ratings (rating of 9 or 10). We also examined the secondary outcomes from each of the nine core questions about patients’ hospital experiences. These questions include communication with nurses and doctors, the responsiveness of hospital staff, the cleanliness and quietness of the hospital, pain management, communication about medications, discharge information, and whether the patient would recommend the hospital (see, e.g., Goldstein et al. (2005)). We excluded the hospitals that received fewer than 100 survey responses in a given year for this part of the analysis. For measuring overall hospital quality, we used Hospital Compare data (CMS 2018), which provides more comprehensive measures of hospital quality than HCAHPS (see Section 4.5).

4.5. Independent Variables

Table 4 shows the definition and data sources for the independent variables we used to control for potential confounders and examine effect heterogeneity. These variables can be classified into three categories: patient characteristics, hospital characteristics, and area characteristics.

**Patient Characteristics.** To control for patient heterogeneity, we included demographic characteristics such as age, gender, race, a reason for Medicare eligibility, and the Medicare-Medicaid dual eligibility which is often used as a proxy for low-income status. We obtained the total number of chronic conditions a beneficiary had in the previous year from the chronic conditions segment of the Medicare BSF. We calculated patient comorbidities using the Charlson Comorbidity Index (Elixhauser Comorbidity classification) from the patient’s diagnosis history. The scores range from 0 (lowest severity) to 21 (highest severity) and capture the presence of 30 comorbidities (Elixhauser
Table 3: Definition of Outcome Variables and Data Sources

<table>
<thead>
<tr>
<th>Outcome variable</th>
<th>Level</th>
<th>Definition</th>
<th>Source</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational efficiency</td>
<td>Hospital</td>
<td>Total number of patients discharged per hospital bed per year.</td>
<td>Medicare Cost Report, POS</td>
<td>1</td>
</tr>
<tr>
<td>Bed utilization</td>
<td>Hospital</td>
<td>Total bed days used out of available bed days per year.</td>
<td>Medicare Cost Report, POS</td>
<td>2</td>
</tr>
<tr>
<td>Service duration</td>
<td>Patient</td>
<td>Days between admission and discharge date.</td>
<td>Medicare inpatient claims</td>
<td>3</td>
</tr>
<tr>
<td>Patient experience</td>
<td>Hospital</td>
<td>Proportion of patients who rated the total experience of his/her inpatient visit as high (9 or 10 out of 10) from the HCAHPS survey.</td>
<td>HCAHPS</td>
<td>4</td>
</tr>
<tr>
<td>Readmission</td>
<td>Patient</td>
<td>A binary variable indicating the presence of another hospitalization within 30 days of discharge.</td>
<td>Medicare inpatient claims</td>
<td>5</td>
</tr>
<tr>
<td>Mortality</td>
<td>Patient</td>
<td>A binary variable indicating the presence of death within 30 days of discharge.</td>
<td>Medicare inpatient claims</td>
<td>6</td>
</tr>
</tbody>
</table>

et al. 1998). Using these scores allows us to control for the variation in patient health. We categorized the admission into three types, i.e., emergent, urgent, and elective, according to the admission type variable on claims. We also divided the admissions into 15 clinical categories based on the primary diagnosis codes. The details of these categories are explained in the Online Appendix.

**Hospital Characteristics.** We obtained relevant hospital characteristics including the size, academic status, ownership, location, quality, and funding status. Academic status was identified by whether the hospital received any payment from the Graduate Medical Education (GME) program or Indirect Medical Education (IME) program, which pays hospitals for education and training. A hospital was defined as rural if its zip code based Rural-Urban Commuting Area code was greater than 4, or if it was designated as a Critical Access Hospital (CAH), following previous literature (see, e.g., Hart et al. (2005)). We also included an indicator for receiving payment for the Disproportionate Share Hospital (DSH) payments program, which funds hospitals that treat a greater proportion of needy patients. Hospital quality was measured from the publicly available Hospital Compare data provided by the CMS, which draws detailed information on hospital quality from multiple sources, including hospital submitted electronic health records, surveys, and Medicare claims data. The data include 57 quality measures across seven areas of quality and provide an overall rating as well as quality ratings on different dimensions of care. We used the overall rating, which can have values from one star (5.7% of the total hospitals) to five stars (7.36% of the total hospitals).

**Area Characteristics.** We included various market level factors that could influence the operations of hospitals. First, differences in care delivery and quality may exist between the areas with varying degrees of managed care and FFS insurance plan types (see, e.g., Miller and Luft
Table 4: Definition of Independent Variables and Data Sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Numeric, from 64 and up.</td>
<td>Medicare BSF</td>
</tr>
<tr>
<td>Gender</td>
<td>Binary, male or female.</td>
<td>Medicare BSF</td>
</tr>
<tr>
<td>Race</td>
<td>Factor, White, Black, Hispanic, Asian, or others.</td>
<td>Medicare BSF</td>
</tr>
<tr>
<td>Medicare entitlement</td>
<td>Factor, age, disability, or both.</td>
<td>Medicare BSF</td>
</tr>
<tr>
<td>Medicaid eligibility</td>
<td>Binary, dual or non-dual.</td>
<td>Medicare BSF</td>
</tr>
<tr>
<td>Chronic conditions</td>
<td>Numeric, from 0 to 27.</td>
<td>Medicare BSF</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>Numeric, from 0 (least severe) to 21 (most severe).</td>
<td>Medicare inpatient, outpatient claims</td>
</tr>
<tr>
<td>Admission type</td>
<td>Factor, emergent, urgent, or elective.</td>
<td>Medicare inpatient claims</td>
</tr>
<tr>
<td>Diagnosis type</td>
<td>Factor, 15 clinically meaningful categories.</td>
<td>Medicare inpatient claims</td>
</tr>
<tr>
<td><strong>Hospital characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beds</td>
<td>Numeric, greater than 0</td>
<td>Medicare Cost Report, POS</td>
</tr>
<tr>
<td>Academic status</td>
<td>Binary, teaching or non-teaching hospital.</td>
<td>Medicare Cost Report, POS</td>
</tr>
<tr>
<td>Ownership</td>
<td>Factor, nonprofit, private, or public.</td>
<td>Medicare Cost Report, POS</td>
</tr>
<tr>
<td>Location</td>
<td>Binary, rural or urban.</td>
<td>Medicare Cost Report, POS</td>
</tr>
<tr>
<td>Quality</td>
<td>Numeric, from 1 (lowest) to 5 (highest).</td>
<td>Hospital Compare data</td>
</tr>
<tr>
<td>DSH</td>
<td>Binary, DSH or non-DSH.</td>
<td>Medicare Cost Report, POS</td>
</tr>
<tr>
<td><strong>Area characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managed care penetration</td>
<td>Numeric, from 0 (no penetration) to 1 (full penetration).</td>
<td>State/County/Plan Enrollment Data</td>
</tr>
<tr>
<td>Herfindahl-Hirschman Index</td>
<td>Numeric, greater than 0</td>
<td>Medicare inpatient claims</td>
</tr>
<tr>
<td>Provider supply</td>
<td>Numeric, number of providers per 10,000 persons.</td>
<td>AHRF</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>Numeric, from 0 (full employment) to 1 (full unemployment).</td>
<td>AHRF</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>Numeric, from 0 (no poverty) to 1 (full poverty).</td>
<td>AHRF</td>
</tr>
<tr>
<td>Population under age 65</td>
<td>Numeric, from 0 (none under 65) to 1 (all under 65)</td>
<td>AHRF</td>
</tr>
</tbody>
</table>

(1997), Baicker et al. (2013), and the references therein). Thus, we included the yearly county level penetration rate of Medicare managed care plans by calculating the proportion of beneficiaries enrolled in any Medicare Advantage (i.e., Medicare’s managed care type plans, a type of health insurance plans that actively manages cost and quality) out of total Medicare beneficiaries each year. Second, to control for the changes in the degree of market competition, we constructed the Herfindahl-Hirschman Indices (HHIs) for hospitals—a standard measure of concentration—for each market (defined as HRRs) per year. Third, the level of the provider supply may affect bed utilization and operational efficiency. Thus, we adjusted for the area level provider supply, such as the total number of primary care physicians and acute care hospitals per 10,000 persons from AHRF. Lastly, to adjust for any macro level socio-demographic factors, we controlled for the proportion of population unemployed, in poverty, or aged 16 or older for each county from AHRF.

5. Main Empirical Analysis

5.1. Empirical Strategy Overview

Our main empirical strategy is a DID approach with hospital, market, and year fixed effects to examine the changes in hospital and patient outcomes before and after a hospital closure event. This approach allows for controlling observed and unobserved heterogeneity between the treatment and control group that is constant over time. If the parallel trend assumption is met, DID analysis can provide a causal interpretation of the treatment effect. We used a fixed effects model instead
of a random effects model, because the hospital or market effects are likely correlated with the observed patient or hospital characteristics. We used hospital level instead of patient level fixed effect since a large proportion of patients had only one hospital visit. We used a robust standard error clustered at hospital-year to account for the correlation of error terms.

We employed the following model for testing hospital level outcomes (hypotheses 1, 2 and 4):

$$Y_{jt} = \alpha_1 X_{jt} + \alpha_2 Z_{jt} + \beta \text{POSTCLOSURE}_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + \epsilon_{jt}.$$  \hspace{1cm} (1)

To test the patient level outcomes (hypotheses 3, 5, and 6), we utilized the following model:

$$Y_{ijt} = \alpha_1 X_{it} + \alpha_2 Z_{jt} + \beta \text{POSTCLOSURE}_{jt} + \text{HOSPITAL}_j + \text{AREA}_i + \text{YEAR}_t + \epsilon_{ijt}.$$  \hspace{1cm} (2)

In both models (1) and (2), $Y$ represents the outcome variables, POSTCLOSURE is a binary variable that indicates that the observation is made in the post-closure year for the treated group. HOSPITAL, AREA, and YEAR represent the hospital, area, and year fixed effects. For the hospital-level model (Equation 1), there is no area fixed effect since hospitals do not change their locations. $X$ is a vector of patient characteristics, including age, gender, race, Medicare entitlement, Medicaid dual-eligibility, chronic conditions, comorbidity, admission type, and diagnosis type. $Z$ is a vector of area characteristics, including managed care penetration (i.e., the proportion of patients who own managed care insurance plans per market), HHI, provider supply measures, and socioeconomic measures. $\epsilon$ is the error term. Indices $i$, $j$, and $t$ represent a patient, a hospital, and a year, respectively. Bold notation is used to represent vectors.

To improve the comparability of our treatment and control groups, we made use of matching in our main analyses. Specifically, we first estimated the propensity score of being in the treatment and control groups using a logistic regression model where we employed the following matching criteria separately: (a) hospital characteristics only (size, academic status, ownership, location, funding, and quality) and (b) hospital characteristics along with patient characteristics (age, gender, race, dual status, and admission type). We then utilized the nearest-neighbor matching method without replacement. Because of the imbalance between the numbers of hospitals in the treatment and control group, matching resulted in a fewer number of study sample: 243 hospitals for hospital-only based matching and 239 hospitals for hospital and patient-based matching. The balance statistics of the matched groups can be found in Table 1 of the Online Appendix. As part of our robustness checks, we repeat our analyses without matching (see section 7). Finally, to gain further confidence, we revise models (1) and (2), and examine the temporal trend of closure effects by separately measuring our outcome variables for each observation year.
5.2. Assumptions
The main assumption of our fixed effect DID model is that conditioned on the unobserved fixed differences by groups, each observation-specific error term is uncorrelated with the explanatory variables in all periods (i.e., strict exogeneity holds). We control for multiple dimensions of time-varying proxies for health, socio-economic status, and market characteristics such as insurance penetration and competition level in our analyses to address potential violations of the strict exogeneity assumption. However, there might still be two major threats to this assumption.

First, there can be a patient-level selection that is correlated with hospital closures. For example, sicker patients may have chosen a particular hospital and have contributed to its closure, and then their influx to neighbor hospital after closure could have contributed to increased adverse outcomes. Since our controls (e.g., our proxies for health) might have not fully captured this endogeneity, we perform our analyses after eliminating the patients of closed hospitals from our sample. Furthermore, we note that this endogenous selection process is rather static because underlying clinical or socio-economic differences for healthcare are typically stable in the short term (see also Fiscella et al. (2000)). These give us confidence that our results are not affected by such a patient-level selection process.

Another critical assumption in our DID analyses is the parallel trend, which posits that the differences between the treatment and control groups are constant over time. The assumption is not formally testable, but we show that the pre-treatment outcomes do have a parallel trend (Figure 3 of the Online Appendix). To gain further confidence, we also test for the common trend between the treatment and the control group before the treatment year by including the interaction term of treatment variable with each pre-treatment year. We find that the pre-treatment trends are not significantly different between the two groups prior to the treatment year (Table 2 of the Online Appendix).

Finally, because we cannot completely verify the extent to which these unmeasurable aspects bias our results, we use an instrumental variable (IV) approach as part of our robustness checks (see Section 7.2). This IV approach further mitigates the concerns mentioned above and gives us assurance about the validity of our results.

6. Results and Discussions
6.1. Summary Statistics
An average hospital in our data serves 9,724 patients per year, with the operational efficiency of 46 (patients per bed per year), and bed utilization of 54% (see Table 3 in the Online Appendix). The
average service duration, 30-day readmission rate, and 30-day mortality rate are 4.81 days, 16%, and 6%, respectively, which are consistent with the existing literature (see, e.g., Joynt et al. (2011), Bueno et al. (2010)). Compared to the hospitals in the control group, hospitals in the treatment group are more likely to be teaching, public, and urban (Table 5). They are also more likely to be located in a competitive market, which is consistent with our results that hospitals in competitive markets are more likely to close. Patients in the treatment group are slightly older, less likely to be male or White race, more likely to be low-income, and sicker (Table 5). We adjust these differences in hospital and patient characteristics in our DID analysis.

6.2. Average Effect

Figure 2 and Table 6 show the DID results of our main model for our hospital and patient level outcomes, respectively. The full results are presented in Table 4 of the Online Appendix. Our results indicate that hospitals improve their operational efficiency after the closure of nearby hospitals by 1.6 additional discharges per bed (equivalent to 3.5% increase) in the post-closure years, and this change is statistically significant (hypothesis 1). To test if the efficiency gain is driven by an increase in patient volume as opposed to a change in hospitals’ capacity, we separately examine the changes in volume and capacity. We see a substantial increase in patient volume, but no significant increase in the number of beds (see Table 5 of the Online Appendix). While the bed utilization rate also increases by about 2.1% in the post-closure years, this increase is not statistically significant (hypothesis 2). The log length of stay of the remaining hospitals, however, decreases by a statistically significant amount of 2.3%, which translates to one out of every five patients being released a day earlier (hypothesis 3).

There is no statistically significant change in overall patient experience ratings after hospital closures (hypothesis 4). When the ratings for each of the nine domain for the patient experience is examined separately, two out of nine domains (doctor communication and staff explanation) show a significant reduction in quality (Table 6 of the Online Appendix). Hospital closures do not impact the 30-day readmission rate of their nearby hospitals (hypothesis 5). The 30-day mortality rate of such hospitals, however, increases by a statistically significant amount of 3.8%, which translates to an additional 2.3 deaths per 1,000 patients per year (hypothesis 6).

Overall, we find evidence of efficiency improvement following a closure event as measured by the number of patients served per bed per unit of time. Although hospitals improve their efficiency, such an improvement is not due to an increase in bed utilization (lower bed idle times): a decrease in service duration—a speed-up behavior—is the main reason behind the efficiency improvement.
Table 5: Hospital, Market, and Patient Characteristics in the Treatment and Control Groups

<table>
<thead>
<tr>
<th>Hospital characteristics</th>
<th>Treatment</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitals (n)</td>
<td>158</td>
<td>3,141</td>
</tr>
<tr>
<td>Hospitals (hospital-year)</td>
<td>1,009</td>
<td>30,656</td>
</tr>
<tr>
<td>Patients (n)</td>
<td>175,694</td>
<td>4,537,801</td>
</tr>
<tr>
<td>Inpatient visits (n)</td>
<td>418,175</td>
<td>13,626,652</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Market characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg HMO penetration rate (%)</td>
</tr>
<tr>
<td>Avg Herfindahl-Hirschman Index</td>
</tr>
<tr>
<td>Avg total population (n)</td>
</tr>
<tr>
<td>Avg unemployment rate (%)</td>
</tr>
<tr>
<td>Avg poverty rate (%)</td>
</tr>
<tr>
<td>Avg under age 65 (%)</td>
</tr>
<tr>
<td>Avg hospitals (/10,000)</td>
</tr>
</tbody>
</table>

Note. DSH indicates Disproportionate Share Hospital. Dual-eligible hospitals indicate Medicare-Medicaid dual eligible beneficiaries. For hospital level variables, all differences in covariates between the two groups were statistically significant at p-value < 0.001, except DSH (p-value 0.21), avg. bed utilization (p-value 0.10), avg. unemployment (p-value 0.07), and avg. hospital (p-value 0.29). For patient level variables, all differences in covariates between the two groups were statistically significant at p-value < 0.001.

Table 6: Difference-in-Differences Estimates: Average Effect of Hospital Closure

<table>
<thead>
<tr>
<th>Hypothesis Outcomes</th>
<th>(1) Operational efficiency</th>
<th>(2) Bed utilization</th>
<th>(3) Service duration</th>
<th>(4) Patient experience</th>
<th>(5) Readmission</th>
<th>(6) Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearby hospital's closure</td>
<td>1.589* (0.666)</td>
<td>0.011 (0.006)</td>
<td>−0.115** (0.036)</td>
<td>−0.487 (0.357)</td>
<td>0.0020 (0.0016)</td>
<td>0.0023** (0.0008)</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.80</td>
<td>0.87</td>
<td>0.84</td>
<td>0.84</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Number of observations</td>
<td>4,145</td>
<td>4,145</td>
<td>1,623,079</td>
<td>2,969</td>
<td>1,623,079</td>
<td>1,623,079</td>
</tr>
</tbody>
</table>

Note. All models include hospital, area, and year fixed effects. Standard errors are in parentheses. Standard errors are robust and clustered at the hospital and the year levels. *p < 0.05; **p < 0.01; ***p < 0.001.

in the remaining hospitals. Importantly, indicating both a decrease in service duration and an increase in 30-day mortality rate, our results suggest that the remaining hospitals likely respond to the increased demand caused by a nearby hospital closure through eliminating some value-added care steps.

6.3. Heterogeneous Effect by Market and Closure Characteristics

The spillover effect of closure may depend on the market environment. For example, in a competitive market, hospital closure could be driven by market forces. To generate further insights into the
role of market characteristics, we stratified our study sample based on characteristics such as the level of competition (measured by HHI) and resources (measured by the number of inpatient hospitals per population). Figure 3 (and Table 7 of the Online Appendix) shows that the neighbor hospitals in less competitive or lower resource markets (i.e., where there are fewer remaining choice of hospitals after a hospital closes) experience a greater increase in efficiency, greater reduction in service duration, and a greater increase in adverse patient outcomes. This suggests that bailing out hospitals in areas where patients have fewer choices (e.g., rural areas) can be an effective strategy in mitigating the negative consequences of closures on quality of care while bailing out hospitals in opposite markets might be a more effective strategy in improving the overall operational efficiency.

We test the consequence of utilizing these and similar levers that can be used by policymakers in Section 8.

6.4. Heterogeneous Effect by Hospital Characteristics

We next examined the heterogeneous spillover effect of closures based on neighbor hospitals' characteristics (academic status, quality, location, and size). Our results are presented in Figure 4 (see also Table 8 of the Online Appendix) and show that, after a hospital closes, the neighbor hospitals that are generally considered to be more desirable (e.g., teaching, high quality, urban, and large) experience a significant increase in efficiency. In contrast, less desirable (e.g., non-teaching, low quality, rural, and small) hospitals do not experience any significant changes in efficiency. This suggests that the effect of hospital closures is not uniform even in the same market; it depends largely on hospital characteristics. This offers further opportunities for policymakers to target their policies not only at specific markets but also at specific hospitals (see Section 8 for our detailed policy recommendations). Furthermore, we note that the more desirable hospitals that gained efficiency tend to be the ones that also experienced a significant increase in 30-day mortality. Since the gain
in efficiency is mainly due to a speed-up behavior, this suggests that the speed-up response of the nearby hospitals—particularly by the more desirable hospitals—is the main driving force behind the increase in 30-day mortality. In Table 9 of the Online Appendix, we formally test the mechanism that drives the increase in 30-day mortality. The results support our main finding: hospitals that accommodate the increased demand following a nearby hospital closure by improving their efficiency do so through a speed-up behavior, which most likely involves the elimination of some value-added care delivery procedures.

7. Robustness

7.1. Robustness Check Using Instrumental Variable (IV) Analysis
As noted in Section 5.2, our fixed effects model may not fully address the time-varying unobservable confounders that can bias our results. Therefore, we incorporated an IV analysis in our DID design (i.e., instrumented difference-in-differences, or DDIV) by identifying an IV that can account for unmeasured confounders (De Chaisemartin and D'HaultfŒuille 2017, Duflo 2001). Specifically, we made use of the state level variations in the decision to expand Medicaid as an instrument that influences the likelihood of hospital closures but is unlikely to be correlated with our outcome variables. The Affordable Care Act (ACA) originally intended to expand Medicaid coverage to low-income adults, but the provision was ruled coercive by the supreme court. Therefore, each state could choose to expand or not expand Medicaid, which created a variation in Medicaid eligibility by state. Evidence shows that the expansion is associated with improved hospital financial performance and a lower likelihood of hospital closure (Lindrooth et al. 2018, Blavin 2016). Using these facts, we specified our first-stage equation as:
Figure 4: Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure on (a) Operational Efficiency (Left) and (b) 30-Day Mortality (Right) by Hospital Characteristics

Note. Each dot indicates the size of the DID coefficient. The DID coefficient was estimated by stratifying the sample into the corresponding hospital characteristics. All outcomes are scaled as percent changes from the mean. Grey lines depict the 95% confidence intervals around the coefficient of the DID variable. Standard errors in parentheses are robust and clustered at the hospital and the year levels.

\[
\text{POSTCLOSURE}_{jt} = \delta_1 X_{jt} + \delta_2 Z_{jt} + \mu \text{MEDICAID}_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + \nu_{jt},
\]  

(3)

where MEDICAID\(_{jt}\) denotes whether hospital \(j\)'s state expanded Medicaid in year \(t\) and POSTCLOSURE\(_{jt}\) indicates whether the neighbor hospital of hospital \(j\) is closed in year \(t\). Our second-stage equation for hospital level outcomes is

\[
Y_{jt} = \alpha_1^{IV} X_{jt} + \alpha_2^{IV} Z_{jt} + \beta^{IV} \text{POSTCLOSURE}_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + \epsilon_{jt},
\]

(4)

and for patient level outcomes is

\[
Y_{ijt} = \alpha_1^{IV} X_{it} + \alpha_2^{IV} Z_{jt} + \beta^{IV} \text{POSTCLOSURE}_{jt} + \text{HOSPITAL}_j + \text{AREA}_i + \text{YEAR}_t + \epsilon_{ijt},
\]

(5)

where POSTCLOSURE is the estimated value from the first-stage equation (3), and \(\beta^{IV}\) is the impact of hospital closures on outcome variables adjusting for the selection using the instrument.

The key assumptions for our IV approach are: (1) the instrument does not affect the outcome except through treatment (exclusion restriction), and (2) the instrument is associated with the treatment variable (Hudson et al. 2017). Available studies in the literature suggest that Medicaid expansion—our IV—is strongly correlated with our treatment variable (see, e.g., Lindrooth et al. (2018)), and hence, assumption (2) holds. Our direct tests on the level of correlation between Medicare expansion and our treatment variable further confirm this (see Table 7).

However, unlike assumption (2), we cannot directly test assumption (1). Several studies suggest that although Medicaid expansion is associated with the changes in payer mix and financial margins of the hospitals, it does not impact their overall use or patterns of inpatient care (Pines et al. 2016, Freedman et al. 2017, Pickens et al. 2018). Nevertheless, to gain further confidence, we tested if Medicaid expansion is associated with changes in the hospital level outcomes among hospitals that
Table 7: Instrumental Variable Estimates

<table>
<thead>
<tr>
<th>First stage</th>
<th>Second stage</th>
<th>(3) Bed utilization</th>
<th>(4) Service duration</th>
<th>(5) Patient experience</th>
<th>(6) Readmission</th>
<th>(7) Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicaid expansion</td>
<td>-0.013**</td>
<td>0.94**</td>
<td>-0.13***</td>
<td>-2.1</td>
<td>-0.007</td>
<td>0.002***</td>
</tr>
<tr>
<td>Nearby hospital’s closure</td>
<td>(0.004)</td>
<td>(0.0049)</td>
<td>(0.010)</td>
<td>(1.63)</td>
<td>(0.012)</td>
<td>(0.0011)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.51</td>
<td>0.76</td>
<td>0.35</td>
<td>0.08</td>
<td>0.35</td>
<td>0.05</td>
</tr>
<tr>
<td>Number of observations</td>
<td>24,440</td>
<td>35,315</td>
<td>35,315</td>
<td>9,830,849</td>
<td>35,315</td>
<td>9,830,849</td>
</tr>
<tr>
<td>Weak identification (F-stat)</td>
<td>91.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The F-statistics test for identifying the weak instrument is based on (Stock and Yogo 2002). The rule of thumb suggests that a first stage F-statistic below 10 indicates the presence of weak instruments. All models include hospital, area, and year fixed effects. Standard errors are in parentheses. Standard errors are robust and clustered at the hospital and the year levels. *p < 0.05; **p < 0.01; ***p < 0.001.

are in the control group. Our results suggest that assumption (2) most likely holds for hospital level outcomes (see Table 10 of the Online Appendix). For patient level outcomes, we note that because our study population is Medicare beneficiaries who have already had insurance coverage, the expansion of Medicaid is unlikely to affect their care patterns.

In addition to the two above-mentioned assumptions required for an IV approach, the DDIV approach that we employ requires two more assumptions: (3) in the absence of treatment, the difference between the treatment and control group is constant over time, and hence, shows a parallel trend, and (4) the effect of the instrument is monotone (Hudson et al. 2017). Assumption (4) is well-satisfied, since once a hospital closes, it will stay closed throughout the study period by our definition of a closure. To verify the parallel trend assumption in (3), we examined the pre-expansion trends of the treatment group’s outcomes by expansion status. Our results indicate that the parallel trend assumption most likely holds. However, it should be noted that our IV estimate is applicable only to the hospitals whose behaviors are influenced by the IV. In addition, the instrument generates a variation at the state level, so within-state unobservable differences between the treatment and the control groups may still remain. Despite these, our IV analysis helps to validate the findings from our primary analysis and provide us with a useful additional robustness check mechanism.

Table 7 shows the result of the first stage and second stage regression estimates (Equations (3)-(5)). The first column shows that Medicaid expansion is a significant predictor of fewer closures of nearby hospitals. The subsequent columns show the impact of hospital closures on the nearby hospital’s operational efficiency and quality. The fact that the results of our IV analysis and the DID analysis are consistent gives us confidence that our results are fairly robust.
7.2. Robustness Check Using Temporal Effects

We further examined the robustness of our results by studying the temporal trend of closure effects (i.e., by separately measuring the effect for each observation year). We employed the following model for testing hospital level outcomes (hypotheses 1, 2, and 4):

\[ Y_{jt} = \alpha_1 X_{jt} + \alpha_2 Z_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + \beta_1 \text{POSTYEAR1}_{jt} + \beta_2 \text{POSTYEAR2}_{jt} + \beta_3 \text{POSTYEAR3}_{jt} + \beta_4 \text{POSTYEAR4}_{jt} + \beta_5 \text{POSTYEAR5}_{jt} + \epsilon_{jt}. \]  

(6)

To test patient level outcomes (hypotheses 3, 5, and 6), we utilized the following model:

\[ Y_{ijt} = \alpha_1 X_{it} + \alpha_2 Z_{jt} + \text{HOSPITAL}_j + \text{AREA}_i + \text{YEAR}_t + \beta_1 \text{POSTYEAR1}_{jt} + \beta_2 \text{POSTYEAR2}_{jt} + \beta_3 \text{POSTYEAR3}_{jt} + \beta_4 \text{POSTYEAR4}_{jt} + \beta_5 \text{POSTYEAR5}_{jt} + \epsilon_{ijt}. \]  

(7)

In both models (1) and (2), POSTYEAR1, ···, POSTYEAR5 are binary variables that indicate that the observation is made in each of the post-closure years for the treated group. All other variables are the same as our main equations.

Our results (see Figure 4 of the Online Appendix) show that the effects of hospital closures discussed earlier persist over time, though their magnitudes typically decrease. This diminishing effect could be due to other contemporaneous market or policy changes that mask the closure effect. For example, many markets have experienced either an opening or an additional closure of hospitals within five years of closure. Alternatively, the diminishing effect could suggest that hospitals get accustomed to the increased demand over time. Nevertheless, the fact that we observe the same directional effects for every post-closure year as those in our main analyses gives us further confidence that our results are robust during our entire study period.

7.3. Other Robustness Checks

To gain further confidence in the validity of our results, we also performed various other robustness checks. Here, we first describe them and then discuss the resulting observations.

**Unmatched Sample.** We repeated our analyses using the original sample data instead of using the matched sample.

**Treatment Variable.** We repeated our analyses using an alternative area-based definition of the treatment variable. Specifically, we assigned all hospitals located in the same healthcare market that experienced at least one closure to the treatment group, where the market was defined as an HRR.

**Patients of Closed Hospitals.** Because the patients of closed hospitals are poorer and sicker (Table 2), it is possible that the unobservable changes in patient composition contribute to the
increase in mortality. Thus, we removed the patients of closed hospitals from the study population and re-estimated the closure effect.

**Travel Distance.** The changes in the travel distance as a result of closure may have contributed to the increased mortality. Thus, we repeated our analyses after including the average distance between patients and hospitals (using patients’ and hospitals’ zip code centroids).

**Observations Years.** As another robustness checks, we included additional observation years for the hospital level outcomes. Specifically, we repeated our analyses after including observational years 2001-2004 and 2016-2017, which we removed from our main analysis for comparability with the observation years for the patient level data (2005-2015).

**Placebo Test.** We randomly assigned hospital closures to some hospitals and patients in the control group and re-estimated the model with the placebo-treated group. We also assigned the treatment to hospitals in the years before the closure year as a placebo treatment and again re-estimated the model.

All of the changes resulted in observations that are reasonably consistent with our primary analysis (Table 8). In particular, the results of our various robustness checks indicate that our main findings are fairly robust: hospital closures increase the operational efficiency of nearby hospitals, but have negative consequences on some aspects of quality of care, including the mortality rate. When the market based treatment definition was used, service duration decreased marginally (p-value < 0.1), likely since more hospitals (especially the ones farther away from the closed hospitals) are included in the treatment group. When the patients of closed hospitals were removed, despite removing about 24% of the patients in the treatment group, service duration decreased with statistical significance, and mortality increased marginally. The results of the placebo test showed no significant closure effect on all outcomes. Furthermore, the estimates of the rest of the outcomes were all consistent with our primary results. Finally, the fact that removing the patients of closed hospitals and adding the changes in distance does not affect the spillover effect strengthens the interpretation of the speed-up behavior as the driver of adverse patient outcomes.

### 8. Policy Implications

Our results suggest that there are at least two policy levers that could be utilized by policymakers and hospital administrators to harness the positive spillover effect of hospital closures and/or mitigate their negative consequences. First, the fact that the effect of hospital closure depends on the market and hospital characteristics indicates that bailing out hospitals that have specific characteristics or are in particular markets can have a strong impact. Second, policies that can
eliminate the speed-up behavior (e.g., appropriate monitoring and regulations against reductions in service durations) post a nearby hospital closure can be beneficial. In order to provide clear policy recommendations based on our results, we now perform various counterfactual analyses and examine the effectiveness of these policy levers (had they been utilized).

8.1. **Policy Lever 1: Selective Hospital Closures and Bailouts**

For policymakers, knowing “which” hospitals they should close or bail out under considerations of efficiency and quality can be highly informative. For example, there has been an ongoing debate on cutting the Medicaid DSH payment program or reforming the CAH status for financial support—two programs that, roughly speaking, try to bail out specific hospitals and prevent them from potential closures. While these programs help to support safety-net hospitals that are not viable under market competition forces (Neuhausen et al. 2014, Bazzoli et al. 2014), they have been controversial for being costly and also negatively affecting the efficiency of the healthcare system.
Based on the current distribution of the characteristics of U.S. hospitals, we examined hypothetical scenarios of closures to estimate the spillover effect their neighbors (and thus, the healthcare system) experience. First, for market-based scenarios, we estimated the expected increase in operational efficiency and 30-day mortality to the neighboring hospitals when a hospital in a market with a low (high) competition or a low (high) resource closes. We divided the markets into an equal number of high versus low competition and high versus low resources markets and randomly chose one hospital in each market (\(N = 306\)) as a closing hospital, resulting in \(N = 153\) closures. Second, for hospital-based scenarios, we examined the scenarios of closures based on their relative desirability (teaching, high-quality, urban, and large) compared to their neighbors. We randomly selected \(N = 153\) hospitals such that the characteristics of our interest are either more or less desirable relative to their neighbors. We estimated the effect of hospital closures on the neighbor hospitals’ efficiency and 30-day mortality using the estimated parameters from the Equations (1) and (2), holding all other characteristics (other than neighbor’s closures) constant. Finally, we calculated the mean changes in outcomes.

Figure 5(a) shows the estimated increases in operational efficiency and 30-day mortality under different closure scenarios compared to the case of no hospital closures. Overall, our results indicate that there is no dominant strategy that improves both efficiency and patient outcomes, as the increase in efficiency tends to be present with an increase in mortality. There is also wide heterogeneity in the magnitude of policy effect, which suggests that choosing specific types of hospitals to bail out can make a substantial difference in overall impact. If greater weight is placed on reducing the adverse consequence in mortality as opposed to improving efficiency, our results recommend bailing out hospitals that are in the areas with fewer choices of hospitals (e.g., in rural areas) or those that have less desirable characteristics than their neighbors.

8.2. Policy Lever 2: Selective Elimination of Speed-up Behavior

Current studies on the interventions related to the speed of medical care focus on reducing the service time for time-sensitive conditions such as stroke care in emergency department settings (Fonarow et al. 2011, Meretoja et al. 2012). Limited data exist, however, on interventions that can slow down the service to conserve value-added care. In the absence of such data, we focused on estimating the maximum achievable benefits from eliminating the speed-up behavior. Our empirical findings suggest that the hospitals that speed up increase their efficiency by reducing the service duration while keeping the bed utilization rate constant. Thus, we examined the hypothetical scenario when hospitals respond to the increase in patient demand by increasing their bed utilization
Figure 5: Counterfactual Results Under (a) Policy Lever 1: Selective Hospital Closures and Bailouts (Left) and (b) Policy Lever 2: Selective Elimination of Speed-up Behavior (Right)

Note. The black shapes indicate the hospital characteristics, and the blue shapes indicate the market characteristics.

rate instead of changing their service duration. Because the changes in service duration affect the 30-day mortality, we also estimate the reduction in mortality as a result of eliminating the speed-up behavior.

Using our main model, we first predicted the changes in service duration in the absence of the speed-up behavior and translated it into the number of additional patient days for each hospital by multiplying the changes in service duration with the annual patient volume. Using the predicted bed days and service duration, we then re-estimated the impact of the closure on bed utilization and mortality. Our results show that without the speed-up behavior, a hospital’s closure will, on average, increase its neighbors’ bed utilization by 2.2%, and reduce their 30-day mortality rate by 0.08%.

Next, we considered the cases where only certain types of hospitals based on the market or hospital characteristics are targeted to eliminate their speed-up behavior. Figure 5(b) shows the potential gains in bed utilization and the reduction in mortality via policy interventions aimed at eliminating the speed-up behavior (e.g., monitoring and regulating service durations) compared to the status quo where hospitals speed up their services. We observe that in implementing policy lever 2, targeting specific markets where patients have limited choices of hospitals (e.g., rural) or targeting specific hospitals such as urban, non-profit, and teaching (as opposed to rural, for-profit, private, or non-teaching hospitals) is a dominant strategy. This means that policymakers can focus their monitoring and regularization efforts of service durations on these types of markets or hospitals so as to gain the best results.
9. Limitations

Our study has several limitations. First, our outcome variables are based on two different data sources, where the data on patient outcomes is a subset of the data on hospital outcomes. However, we find that these two data sources provide consistent findings. Specifically, both datasets consistently suggest evidence of the speed-up behavior. Moreover, although patient mortality was measured only among Medicare patients, the patient experience that was measured by considering both Medicare and non-Medicare patients showed similar evidence of quality reduction.

A second limitation of our study is the existence of contemporaneous policy changes and their effects, especially after the enactment of the ACA. For example, hospitals’ readmission rates are likely affected by other ongoing changes from the payment reforms such as the Hospital Readmission Reduction Program—a pay-for-performance program that penalizes hospitals with too many readmissions. We adjusted for year-specific shocks and estimated the year-specific treatment effect to mitigate this concern, but the effect of policy changes may vary in a way that is unsolvable to us.

Finally, as we discussed in Section 5.2, our results are limited by the limitations of the DID method we utilized. We employed several alternative strategies to address such limitations, including an IV approach to examine the time-varying omitted variable bias on the hospital level outcomes. Although the results of the IV analysis on hospital outcomes are consistent with our main results, the consistency is valid to an extent the assumptions of IV analysis is valid. We also included an extensive set of covariates in our model, performed a variety of robustness checks, and examined the assumptions on strict exogeneity and the parallel trend. Overall, our supplementary analyses and various sensitivity tests give us confidence that our results are reasonably robust. Nevertheless, future research can further verify our findings.

10. Conclusion

We examined how an exit of a hospital from a market affects the remaining hospitals’ operational efficiency and quality. Our results indicate that in response to the increase in patient demand, nearby hospitals improve their operational efficiency. However, this improvement in operational efficiency is not due to better utilization of resources but is instead due to a speed-up behavior as a response to the increase in demand. This speed-up behavior allows the remaining hospitals to serve more patients with their current level of resources. There is, however, an important negative consequence on some aspects of quality of care, especially an increase in the 30-day mortality rate. Furthermore, the spillover effect of hospital closures is heterogeneous and is stronger when the
nearby hospitals have more desirable characteristics (e.g., large, teaching, urban, or high-quality hospitals) or are in markets where patients have limited choices of hospitals (e.g., in rural areas).

Our empirical findings and counterfactual analyses suggest that targeted versions of some policies can be effective in harnessing the positive impacts of hospital closures and mitigating their negative consequences. For example, we find that (a) bailing out hospitals that are in areas where patients have limited choices of hospitals or those that have less desirable characteristics than their neighbors, and (b) reducing the speed-up of the hospitals in areas where patients have limited choices of hospitals or that have specific characteristics (e.g., urban, non-profit, and teaching) could be effective policies. Our results can be helpful for the current policy debates on the rural hospital closures by showing that the targeted policy interventions that invest in rural hospitals can be effective. It should be noted, however, that our study has focused on understanding the spillover effect of closures (and thus these policies) in the dimensions of efficiency and quality, but policymakers should also consider other dimensions such as cost and equity. For example, while our study has not focused on the cost dimension, our estimates of the cost based on the cost of increased length of stay from previous studies (Bartel et al. 2014, Taheri et al. 2000) suggest that the average intervention cost for large hospitals can be up to five times greater than that of small hospitals (Figure 5 of the Online Appendix). Thus, it is likely that our policy recommendations are more costly than some other potential policies. As such, we emphasize that policymakers need to weigh the advantages and disadvantages of our policy recommendations carefully and accordingly adopt the best strategy.

In addition to policymakers, our findings can also help hospital administrators to adopt suitable strategies in response to a neighboring hospital closure. From hospitals’ perspectives, the surge of patient demand as a result of a nearby hospital’s closure may present as an opportunity for improving profit margins. In particular, because Medicare pays for inpatient services mainly based on diagnosis-related groups (DRGs) that classify patients of similar clinical characteristics and costs rather than the length of stay, the strategy of speeding up to treat more patients might maximize the hospital’s revenue in short-term. Our results, however, point out that such a strategy can adversely affect their hospitals’ long-term sustainability. In light of the recent payment reforms that emphasize the role of hospital quality outcomes (The Centers for Medicare & Medicaid Services 2017) and the growing role for hospitals’ quality outcomes on patients’ choice (Saghafian and Hopp 2019, 2020), hospital administrators should be aware that deterioration in key quality measures as a result of speed-up can result in a loss of patient share for their hospital.
References


Author: Hospital Closures, Efficiency, and Quality

Article submitted to : manuscript no.


