



Data-Driven Healthcare Operations Management: The Role of Provider Market and Public Policy

Citation

Song, Dahye (Lina). 2020. Data-Driven Healthcare Operations Management: The Role of Provider Market and Public Policy. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37366000>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

© 2020 Dahye (Lina) Song

All rights reserved.

Dissertation Advisor:
Professor Soroush Saghafian

Author:
Dahye (Lina) Song

**Data-Driven Healthcare Operations Management:
The Role of Provider Market and Public Policy**

Abstract

Evidence of misallocated resources and misaligned incentives suggests that the U.S. healthcare delivery system operates at suboptimal production efficiency. This dissertation aims to identify effective policy levers to improve the process and service of the U.S. healthcare delivery system by studying the interactions of public policy and the provider market structure. It focuses on the two prominent ongoing changes in the provider market: hospital closure and hospital-physician vertical integration.

Chapter 1 examines the spillover effects of hospital closures on the efficiency and quality of other hospitals. The recent trend of U.S. hospital closures can have a broad impact on the healthcare system by changing the operation of the remaining hospitals. By analyzing Medicare claims data, I find that when a hospital closes, its neighbor hospitals benefit from an increase in volume and operational efficiency. However, I find evidence that the neighbor hospitals do so via a speed-up response (i.e., by reducing their service durations to accommodate the increased demand), instead of improving their bed utilization rate. I further show that such speed-up behavior negatively affects some aspects of the care quality, including an increase in the 30-day mortality rate.

Chapter 2 identifies an optimal financial allocation decision for U.S. hospitals under access-quality trade-off. Supporting rural (as opposed to urban) hospitals can improve access at the risk of reducing care quality. I study allocation decisions for the two hospital support programs currently under debate—the Disproportional Share Hospital (DSH) payment and the Critical Access Hospital (CAH) program—where one primarily targets

urban hospitals and the other rural hospitals. Using a self-selection model that incorporates patients' preferences for geography and quality, I analyze the characteristics of an optimal allocation that improves social aggregate utility under budget constraints. Based on the result, I suggest a threshold-based decision rule and evaluate the performance of our policy against the current payment policies through simulation analyses.

Chapter 3 studies how the financial integration between physicians and hospitals ("vertical integration") affects the care quality of a medical procedure (colonoscopy for colorectal cancer screening and diagnosis) for which the patient outcome is highly sensitive to the quality of the procedure. By examining the recent acquisitions of specialist practices by hospitals and the FFS Medicare patients, I find that integration significantly reduces outcome-related quality, such as post-procedure complications. I further show that such changes are due to the resource constraints that integrated physicians face that drive them to reduce some recommended process-based care. I suggest that carefully designing the incentive structures of the administered payment system for integrated providers can mitigate adverse consequences.

Contents

Abstract	iii
Acknowledgments	xi
1 The Spillover Effects of Hospital Closures on the Efficiency and Quality of Other Hospitals	1
1.1 Introduction	1
1.1.1 Main Findings	3
1.1.2 Main Contributions	5
1.2 Conceptual Framework and Hypotheses	5
1.2.1 Changes in Operational Efficiency	5
1.2.2 Changes in Quality of Care	7
1.3 Related Studies	8
1.4 Data and Study Sample	10
1.4.1 Data	10
1.4.2 Identifying Hospital Closures	10
1.4.3 Study Sample	11
1.4.4 Dependent Variables	13
1.4.5 Independent Variables	15
1.5 Main Empirical Analysis	17
1.5.1 Empirical Strategy Overview	17
1.5.2 Assumptions	19
1.6 Results and Discussions	20
1.6.1 Summary Statistics	20
1.6.2 Average Effect	20
1.6.3 Heterogeneous Effect by Market and Closure Characteristics	23
1.6.4 Heterogeneous Effect by Hospital Characteristics	23
1.7 Robustness	25
1.7.1 Robustness Check Using Instrumental Variable (IV) Analysis	25
1.7.2 Robustness Check Using Temporal Effects	28
1.7.3 Other Robustness Checks	29

1.8	Policy Implications	30
1.8.1	Policy Lever 1: Selective Hospital Closures and Bailouts	31
1.8.2	Policy Lever 2: Selective Elimination of Speed-up Behavior	32
1.9	Limitations	34
1.10	Conclusion	35
2	The Allocation of Funds in Healthcare: Which Hospitals to Support?	38
2.1	Introduction	38
2.1.1	Policy Background	40
2.1.2	Research Questions	42
2.1.3	Summary of Main Findings and Contributions	43
2.2	Literature Review	45
2.3	The Model	47
2.4	Analysis	50
2.4.1	Quality Investment Decisions	50
2.4.2	Hospital Closures	60
2.4.3	Summary and heuristic	63
2.5	Simulation Experiments	64
2.5.1	Data and Study Population	65
2.5.2	Variable Definitions	65
2.5.3	Parameter Estimation	67
2.6	Policy Simulation	69
2.6.1	Policy Scenarios	69
2.6.2	Strategies to Redistribute Funds	72
2.6.3	Hospital Closure Scenarios	73
2.7	Conclusions and Policy Recommendations	76
3	The Impact of Vertical Integration on Physician Behavior and Healthcare Delivery: Evidence from Gastroenterology Practices	79
3.1	Introduction	79
3.1.1	Policy Context and Settings	81
3.1.2	Challenges and Framework	82
3.1.3	Main Findings and Contributions	84
3.2	Conceptual Framework	85
3.2.1	Background in GI Practices	85
3.2.2	Theories of Vertical Integration	85
3.3	Literature	87
3.4	Data and Study Setting	88
3.4.1	Data	88

3.4.2	Measuring Vertical Integration	89
3.4.3	Study Population and Comparison Groups	91
3.5	Variable Definitions and Descriptions	92
3.5.1	Outcome Variables	92
3.5.2	Independent Variables	95
3.5.3	Descriptive Statistics	96
3.6	Main Empirical Strategy	98
3.6.1	Overview	98
3.6.2	Main Assumptions	98
3.7	Results and Discussions	102
3.7.1	Average Effect	102
3.7.2	Drivers of Main Results	105
3.7.3	Heterogeneous Effect	109
3.8	Robustness Checks	111
3.8.1	Effect on Patient Experience	111
3.8.2	Changes in Patient Risk Composition	112
3.8.3	Coding and Gaming Behavior	114
3.8.4	Measuring Integration	114
3.8.5	Other Statistical Tests	114
3.9	Conclusion and Policy Recommendation	115
References		119
Appendix A Appendix to Chapter 1		132
A.1	Additional Methods	132
A.1.1	Independent Variables	132
A.1.2	Cost Estimation for Two Policy Levers	133
A.2	Tables and Figures	133
Appendix B Appendix to Chapter 2		141
B.1	Proofs	141
Appendix C Appendix to Chapter 3		144
C.1	Additional Methods	144
C.1.1	Measuring Quality	144

List of Tables

1.1	Overview of Our Final Data	3
1.2	Hospital and Patient Characteristics of the Open and Closed Hospitals . . .	12
1.3	Definition of Outcome Variables and Data Sources	14
1.4	Definition of Independent Variables and Data Sources	16
1.5	Hospital, Market, and Patient Characteristics in the Treatment and Control Groups	21
1.6	Difference-in-Differences Estimates: Average Effect of Hospital Closure . . .	22
1.7	Instrumental Variable Estimates	27
1.8	Robustness Checks Results	37
2.1	Data Sources	43
2.2	Summary of Hospital Characteristics	67
2.3	Parameters Used in Our Model	68
2.4	Summary of Parameters Used in heuristic for Hospital Investment	73
3.1	Overview of Our Dataset	82
3.2	Quality Measures and Definitions	93
3.3	Definition of Independent Variables and Data Sources	96
3.4	Patient Characteristics and Outcome Variables by Integration Status	97
3.5	Regression Result: Process-Related and Outcome-Related Quality	104
3.6	Regression Result: Operational Efficiency and Spending	104
3.7	Difference-in-Differences Estimates: Heterogeneous Effect of Integration . . .	109
3.8	Robustness Checks Results	113
3.9	Summary of the Impact of Vertical Integration	116
A.1	Summary of Hospital and Patient Characteristics, Matched Sample, (a) Hospi- tal Characteristics Only (Top) and (b) Both Hospital and Patient Characteristics (Bottom)	134
A.2	Summary Statistics of Hospital and Patient Variables	135
A.3	Full Difference-In-Differences Results	136
A.4	Regression Results: Volume and Capacity	137

A.5	Difference-in-Differences Estimates: Average Effect of Hospital Closure on Patient Experience	137
A.6	Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure by Market Characteristics	137
A.7	Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure by Hospital Characteristics	138
A.8	Difference-in-Differences Estimates: Effect of Medicaid Expansion on Hospital Outcomes	138

List of Figures

1.1	Hospital Closures in the U.S. (2006–2014)	2
1.2	Difference-in-Differences Estimates: Average Effect of Hospital Closure . . .	22
1.4	Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure by (a) Market Competitiveness (Left) and (b) Market Resource Level (Right)	24
1.6	Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure on (a) Operational Efficiency (Left) and (b) 30-Day Mortality (Right) by Hospital Characteristics	25
1.8	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)	33
2.1	Log Distance to Hospital 1 and Hospital 2	54
2.2	Cumulative Distribution of Mean a_1 for all Hospital Pairs	55
2.3	heuristic for Investment Decision between Two Hospitals	64
2.4	Cost-Effectiveness Plane for Baseline Policy Scenarios	71
2.5	Changes in Social Welfare (Mean per Patient Utility) under Various Investment Scenarios	74
2.6	Cost-Effectiveness Plane for Hospital Closures (N=303)	75
3.1	Trend of GI Physician Integration (2009-2015)	83
3.3	Distribution of $INTEG_{jt}$	91
3.4	Trends of Process-Related and Outcome-Related Quality	100
3.5	Difference-in-Differences Estimates: Average Effect of Integration	103
3.6	Difference-in-Differences Estimates: Average Effect	107
3.7	Average Effect of Integration on Patient Composition, Scaled into Percent Changes	113
A.1	The General Procedure for Identifying Hospital Closures.	139
A.2	Timeline of Hospital Closures	140
A.3	Difference-in-Differences Estimates: Average Effect of Hospital Closure by Year After Closure	140

Acknowledgments

I would like to thank my dissertation committee, Soroush Saghafian, Mary Beth Landrum, and Joseph Newhouse, for their insightful feedback, mentorship, patience, and guidance. I am also grateful for helpful advice on research and beyond from Richard Zeckhauser, John Hsu, Vicki Fung, Chin Hur, Xabier García de Albéniz, Miguel Hernan, Ken Freedberg, Robert Huckman, Ankur Pandya, Jane Kim, Mark Shepard, Meredith Rosenthal, and Kathy Schwartz. My program cohort, fellow members of the decision science track, and colleagues from the MGH Mongan Institute also provided invaluable support throughout this process. I am indebted to Vanessa Azzone from the Harvard Health Care Policy Department, Jean Roth from NBER, and Maggie Price from MGH for their help in accessing data, Debbie Whitney and Colleen Yout for their help in navigating through the Ph.D. program, and Joe Newhouse, John Hsu, and Katherine Baicker for granting access to various data for research. This research was supported by the grant from the Agency for Healthcare Research and Quality (R36HS026086-01). The content is solely my responsibility and does not necessarily represent the official views of AHRQ. Lastly, this dissertation could not have been completed without the love and encouragement from my family—especially my parents, my sister Amanda and her family, my husband Jinhyuk, and my daughter Adrienne.

Chapter 1

The Spillover Effects of Hospital Closures on the Efficiency and Quality of Other Hospitals¹

1.1 Introduction

A substantial number of U.S. hospitals have closed in the past decade (Kaufman *et al.*, 2016; Friedman *et al.*, 2016; Medicare Payment Advisory Commission (US), 2017). Such closures have occurred widely across the nation affecting a large number of people (see, e.g., Figure 1.1) (Medicare Payment Advisory Commission (US), 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, G. J. *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

¹Co-authored with Soroush Saghaian

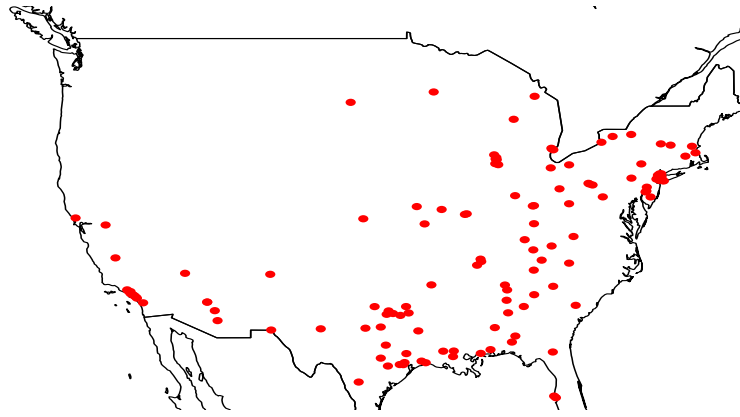


Figure 1.1: *Hospital Closures in the U.S. (2006–2014)*

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 chapter, 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.

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

Table 1.1: *Overview of Our Final Data*

	Number of observations
Hospitals	3,299
Closed hospitals	146
Patients	4,645,532
Inpatient visits	14,147,180
Number of observation years	11

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.

1.1.1 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.

1.1.2 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 Medicare 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.

1.2 Conceptual Framework and Hypotheses

1.2.1 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, R. C.

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 λ/s —our measure of operational efficiency—can be expressed as $\lambda/s = \rho * \mu$, where ρ is the bed utilization, and μ 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 ρ , (b) a decrease only in $1/\mu$, or (c) both an increase in ρ 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 (ρ) 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 Serman, 2001). In particular, the visibility of the queue length or the server occupancy rate can encourage a speed-up behavior by servers (KC, D. S. and Terwiesch, C., 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.*

1.2.2 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.*

1.3 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, M. D. and Mutter, R. L., 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, R. C. *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, C., 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, C. (2019) show that rural hospital closures resulted in an increase in mortality for time-sensitive conditions. While Buchmueller *et al.* (2006); Carroll, C. (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.

1.4 Data and Study Sample

1.4.1 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.

1.4.2 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 Appendix shows the detailed steps through which we determined hospital closures.

Table 1.2 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.

1.4.3 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

Table 1.2: *Hospital and Patient Characteristics of the Open and Closed Hospitals*

	Open	Closed		Open	Closed
Hospitals (n)	3,299	146	<i>Market characteristics</i>		
Inpatient visits (n)	14,044,827	102,353	MA Penetration (%)	17.0	13.9
Patients (n)	4,593,211	52,321	Herfindahl-Hirschman Index	1,451.4	1,317.0
<i>Hospital characteristics</i>			Population (n)	1,041,501.0	706,902.6
DSH (%)	80.5	78.3	Unemployment (%)	9.0	9.3
Teaching hospitals (%)	32.3	22.7	Poverty (%)	16.1	18.2
Ownership – nonprofit (%)	61.5	34.0	Under Age 65 (%)	83.0	82.5
Ownership – private (%)	22.6	10.9	Hosp (/10,000)	0.27	0.39
Ownership – public (%)	15.9	49.7	<i>Patient characteristics</i>		
Rural (%)	33.4	59.7	Avg age (years)	78.98	79.79
Discharges (n)	9,724.4	3,466.7	Sex – male (%)	42.08	37.33
Beds (n)	196.7	100.3	Race – White (%)	87.06	81.44
Operational efficiency	45.9	32.2	Race – Black (%)	8.45	13.23
Bed utilization (%)	0.5	0.4	Dual-eligibles (%)	17.29	26.11
Patient experience (% good)	67.4	63.3	Comorbidity score	3.37	3.64
			Chronic conditions (n)	23.72	21.94

Note. DSH indicates Medicare 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.

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 years (see Table 1.8). Figure 2 in the 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 Appendix. As noted earlier and shown in Table 1.1, 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.

1.4.4 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 1.3 shows the definition of these variables and data sources used. 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.

Table 1.3: *Definition of Outcome Variables and Data Sources*

Outcome variable	Level	Definition	Source
Operational efficiency	Hospital	Total number of patients discharged per hospital bed per year.	Medicare Cost Report, POS
Bed utilization	Hospital	Total bed days used out of available bed days per year.	Medicare Cost Report, POS
Service duration	Patient	Days between admission and discharge date.	Medicare inpatient claims
Patient experience	Hospital	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.	HCAHPS
Readmission	Patient	A binary variable indicating the presence of another hospitalization within 30 days of discharge.	Medicare inpatient claims
Mortality	Patient	A binary variable indicating the presence of death within 30 days of discharge.	Medicare inpatient claims

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).

1.4.5 Independent Variables

Table 3.3 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 *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 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 Medicare 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

Table 1.4: *Definition of Independent Variables and Data Sources*

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

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 (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.

1.5 Main Empirical Analysis

1.5.1 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):

$$Y_{jt} = \alpha_1 \mathbf{X}_{jt} + \alpha_2 \mathbf{Z}_{jt} + \beta \text{POSTCLOSURE}_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + \epsilon_{jt}. \quad (1.1)$$

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

$$Y_{ijt} = \alpha_1 \mathbf{X}_{it} + \alpha_2 \mathbf{Z}_{jt} + \beta \text{POSTCLOSURE}_{jt} + \text{HOSPITAL}_j + \text{AREA}_i + \text{YEAR}_t + \epsilon_{ijt}. \quad (1.2)$$

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 1.1), there is no area fixed effect since hospitals do not change their locations. $\mathbf{X}$ is a vector of patient characteristics, including age, gender, race, Medicare entitlement, Medicaid dual-eligibility, chronic conditions, comorbidity, admission type, and diagnosis type. $\mathbf{Z}$ 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. ϵ 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 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.

1.5.2 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. Yet, 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.

1.6 Results and Discussions

1.6.1 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 2 in the 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, K. E. *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 1.5). 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 1.5). We adjust these differences in hospital and patient characteristics in our DID analysis.

1.6.2 Average Effect

Figure A.3 and Table 1.6 show the DID results of our main model for our hospital and patient level outcomes, respectively. The full results are presented in Table 3 of the 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 4 of the 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

Table 1.5: Hospital, Market, and Patient Characteristics in the Treatment and Control Groups

	Treatment	Control		Treatment	Control
Hospitals (n)	158	3,141	<i>Patient demographics</i>		
Hospitals (hospital-year)	1,009	30,656	Avg age (years)	79.25	78.97
Patients (n)	175,694	4,537,801	Sex – male (%)	41.34	42.10
Inpatient visits (n)	418,175	13,626,652	Race – White (%)	82.49	87.20
<i>Hospital characteristics</i>			Race – Black (%)	11.65	8.36
DSH (%)	82.1	80.4	Race – Asian (%)	1.98	1.33
Teaching hospitals (%)	46.2	31.9	Race – Hispanic (%)	1.98	1.71
Ownership – nonprofit (%)	62.1	61.5	Dual-eligibles (%)	18.93	17.24
Ownership – private (%)	18.3	22.8	Avg comorbidity score	3.62	3.36
Ownership – public (%)	19.5	15.8	Avg chronic conditions (n)	24.02	23.71
Rural (%)	24.4	33.7	<i>Patient clinical</i>		
Avg Star rating	2.8	3.0	Admissions – circulatory (%)	20.02	19.64
Avg discharges (n)	12,129.3	9,645.3	Admissions – digestive (%)	5.20	5.44
Avg beds (n)	245.1	195.1	Admissions – endocrine (%)	2.64	2.29
Avg operational efficiency	44.0	46.0	Admissions – infectious (%)	46.66	46.14
Avg bed utilization (%)	0.6	0.5	Admissions – injury (%)	6.67	6.92
Avg patient experience (% good)	66.0	67.4	Admissions – musculoskeletal (%)	6.35	7.42
<i>Market characteristics</i>			Admissions – nervous (%)	1.27	1.05
Avg MA penetration rate (%)	19.4	16.9	Admissions – respiratory (%)	5.19	5.40
Avg Herfindahl-Hirschman Index	1,186.2	1,460.4	Avg service duration (days)	5.14	4.80
Avg total population (n)	1,520,613	1,024,876	Avg readmission (%)	17.44	16.03
Avg unemployment rate (%)	9.2	9.0	Avg mortality (%)	5.96	6.38
Avg poverty rate (%)	17.0	16.1			
Avg under age 65 (%)	83.8	83.0			
Avg hospitals (/10,000)	0.26	0.27			

Note. DSH indicates Disproportionate Share Hospital. Dual-eligibles 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 .

earlier (hypothesis 3).

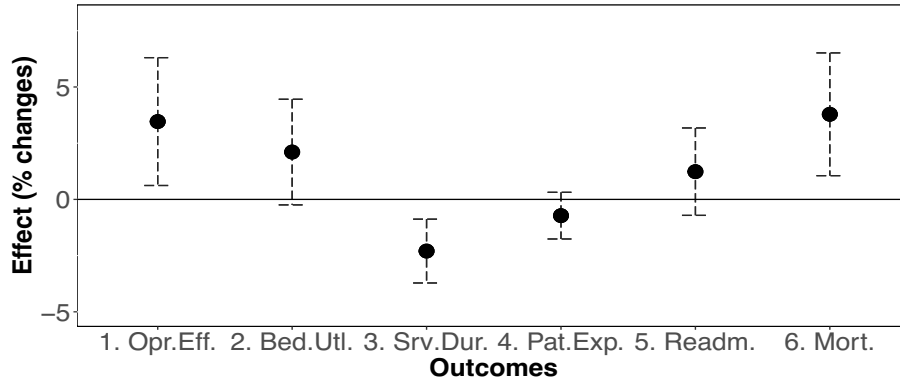
There is no statistically significant change in overall patient experience ratings after hospital closures (hypothesis 4). When the ratings for each of the nine domain for the patient experience is examined separately, two out of nine domains (doctor communication and staff explanation) show a significant reduction in quality (Table 5 of the 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

Table 1.6: *Difference-in-Differences Estimates: Average Effect of Hospital Closure*

Hypothesis Outcomes	(1) Operational efficiency	(2) Bed utilization	(3) Service duration	(4) Patient experience	(5) Readmission	(6) Mortality
Nearby hospital's closure	1.589* (0.666)	0.011 (0.006)	−0.115** (0.036)	−0.487 (0.357)	0.0020 (0.0016)	0.0023** (0.0008)
Adjusted R^2	0.80	0.87	0.08	0.84	0.05	0.03
Number of observations	4,145	4,145	1,623,079	2,969	1,623,079	1,623,079

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. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

**Figure 1.2:** *Difference-in-Differences Estimates: Average Effect of Hospital Closure*

Note. All effects are scaled as changes in percentages. Each dot indicates the size of the DID coefficient. Grey lines depict the 95% confidence intervals around the coefficient of the DID variable. Standard errors are robust and clustered at the hospital and the year levels.

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 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 by eliminating some value-added care steps.

1.6.3 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 1.4 (and Table 6 of the Appendix) shows that the neighbor hospitals in less competitive or lower resource markets (i.e., where there are fewer remaining choices 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.

1.6.4 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 A.7 (see also Table 7 of the 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

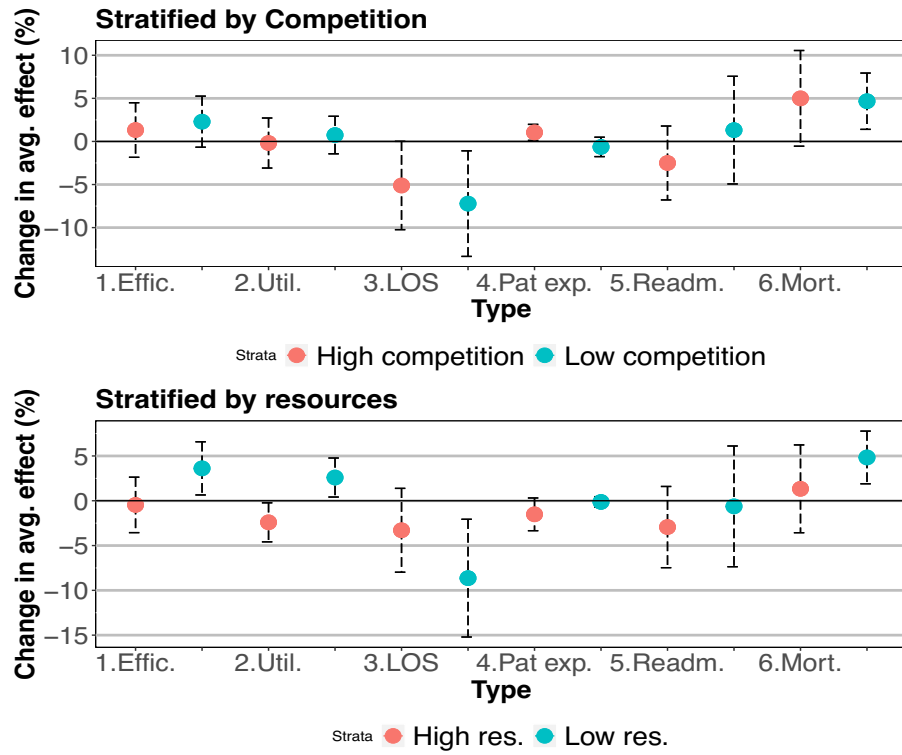


Figure 1.4: Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure by (a) Market Competitiveness (Left) and (b) Market Resource Level (Right)

Note. High competition markets are defined as having the HHI greater than the median HHI of all markets. High resource markets are defined as having the resource level (inpatient hospitals per population) greater than the median of all markets. Each dot indicates the size of the DID coefficient for each of the six hypotheses. 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.

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.

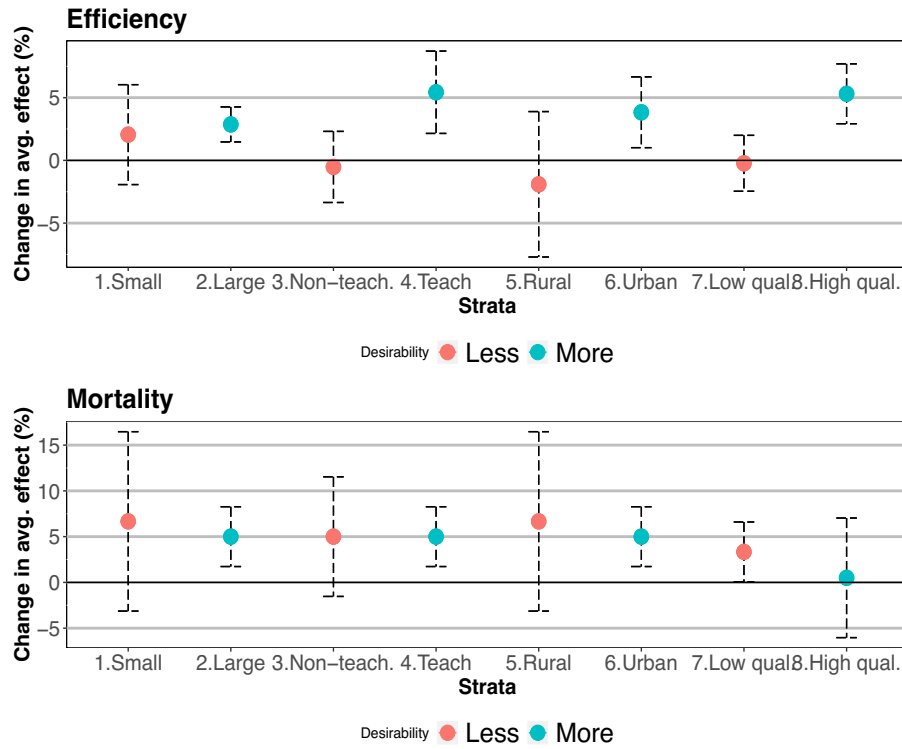


Figure 1.6: Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure on (a) Operational Efficiency (Left) and (b) 30-Day Mortality (Right) by Hospital Characteristics

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.

1.7 Robustness

1.7.1 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:

$$\text{POSTCLOSURE}_{jt} = \delta_1 \mathbf{X}_{jt} + \delta_2 \mathbf{Z}_{jt} + \mu \text{MEDICAID}_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + v_{jt}, \quad (1.3)$$

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

$$Y_{jt} = \alpha_1^{IV} \mathbf{X}_{jt} + \alpha_2^{IV} \mathbf{Z}_{jt} + \beta^{IV} \widehat{\text{POSTCLOSURE}}_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + \epsilon_{jt}^{IV}, \quad (1.4)$$

and for patient level outcomes is

$$Y_{ijt} = \alpha_1^{IV} \mathbf{X}_{it} + \alpha_2^{IV} \mathbf{Z}_{it} + \beta^{IV} \widehat{\text{POSTCLOSURE}}_{jt} + \text{HOSPITAL}_j + \text{AREA}_i + \text{YEAR}_t + \epsilon_{ijt}, \quad (1.5)$$

where $\widehat{\text{POSTCLOSURE}}$ is the estimated value from the first-stage equation (1.3), and β^{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 1.7).

However, unlike assumption (2), we cannot directly test assumption (1). Several studies

Table 1.7: Instrumental Variable Estimates

	First stage (1) Medicaid expansion	Second stage (2) Operational efficiency	(3) Bed utilization	(4) Service duration	(5) Patient experience	(6) Readmission	(7) Mortality
State Medicaid expansion	-0.013** (0.004)						
Nearby hospital's closure		0.94** (0.36)	0.01 (0.0049)	-0.13*** (0.010)	-2.1 (1.63)	-0.007 (0.012)	0.002*** (0.0011)
Adjusted R^2	0.51	0.76	0.35	0.08	0.35	0.05	0.03
Number of observations	24,440	35,315	35,315	9,830,849	35,315	9,830,849	9,830,849
Weak identification (F-stat)	91.62						

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. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

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 are in the control group. Our results suggest that assumption (2) most likely holds for hospital level outcomes (see Table 8 of the 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 1.7 shows the result of the first stage and second stage regression estimates (Equations (1.3)-(1.5)). 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.

1.7.2 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):

$$Y_{jt} = \alpha_1 \mathbf{X}_{jt} + \alpha_2 \mathbf{Z}_{jt} + \text{HOSPITAL}_j + \text{YEAR}_t + \beta_1 \text{POSTYEAR1}_{jt} + \beta_2 \text{POSTYEAR2}_{jt} + \beta_3 \text{POSTYEAR3}_{jt} + \beta_4 \text{POSTYEAR4}_{jt} + \beta_5 \text{POSTYEAR5}_{jt} + \epsilon_{jt}. \quad (1.6)$$

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

$$Y_{ijt} = \alpha_1 \mathbf{X}_{it} + \alpha_2 \mathbf{Z}_{jt} + \text{HOSPITAL}_j + \text{AREA}_i + \text{YEAR}_t + \beta_1 \text{POSTYEAR1}_{jt} + \beta_2 \text{POSTYEAR2}_{jt} + \beta_3 \text{POSTYEAR3}_{jt} + \beta_4 \text{POSTYEAR4}_{jt} + \beta_5 \text{POSTYEAR5}_{jt} + \epsilon_{ijt}. \quad (1.7)$$

In both models (1) and (2), POSTYEAR1, $\dots$, 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 3 of the 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.

1.7.3 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 1.2), 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 1.8). 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 ($p\text{-value} < 0.1$), 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.

1.8 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).

1.8.1 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, G. J. *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 (1.1) and (1.2), holding all other characteristics (other than neighbor's closures) constant. Finally, we calculated the mean changes in outcomes.

Figure 1.8(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.

1.8.2 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 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

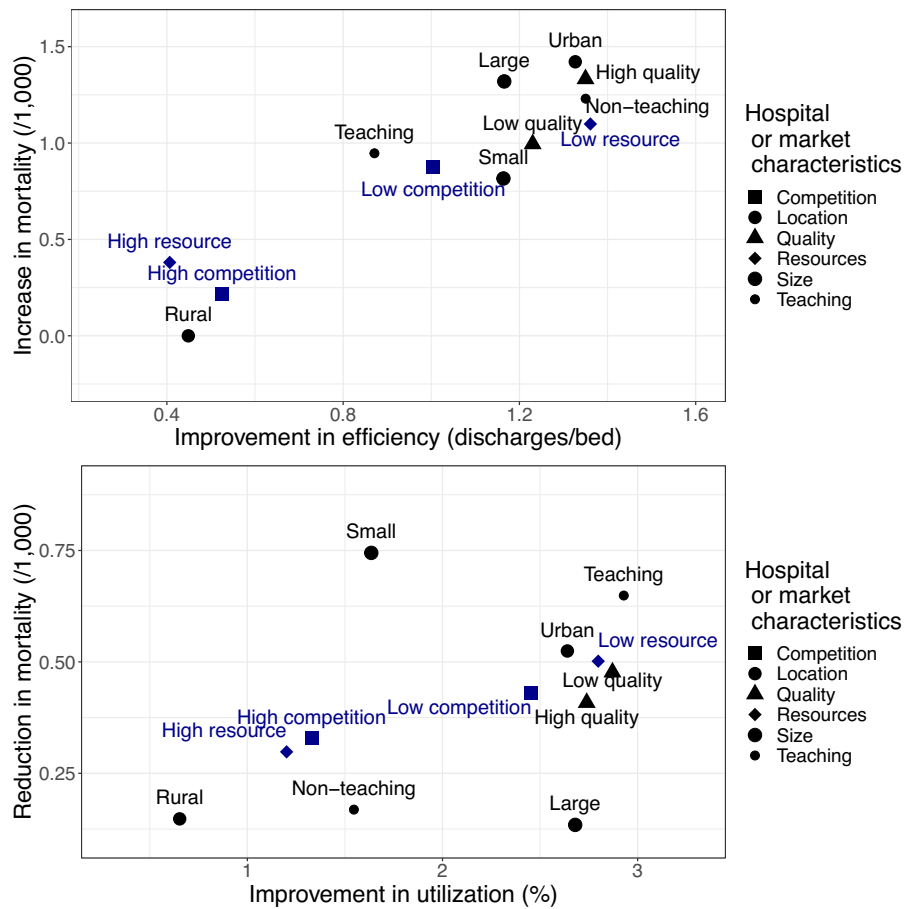


Figure 1.8: Counterfactual Results Under (a) Policy Lever 1: Selective Hospital Closures and Bailouts (Left) and (b) Policy Lever 2: Selective Elimination of Speed-up Behavior (Right)

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

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 1.8(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.

1.9 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 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.

1.10 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 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. 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; Saghafian, S. and Hopp, W. J., 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.

Table 1.8: Robustness Checks Results

Hypothesis Outcomes		(1) Operational efficiency	(2) Bed utilization	(3) Service duration	(4) Patient experience	(5) Readmission	(6) Mortality
Unmatched sample	Nearby hospital's closure	2.84** (1.01)	0.01 (0.01)	-0.11** (0.04)	-0.87 (0.63)	0.001 (0.004)	0.002* (0.001)
	Adjusted R^2	0.75	0.85	0.08	0.82	0.04	0.03
	Number of observations	25,813	25,813	13,539,638	19,534	13,539,638	13,539,638
Treatment variable – HRR defined	Nearby hospital's closure	1.61*** (0.47)	0.007 (0.005)	-0.006† (0.004)	0.362 (0.50)	0.002 (0.001)	0.002* (0.001)
	Adjusted R^2	0.79	0.88	0.10	0.88	0.04	0.04
	Number of observations	16,601	16,601	3,103,077	5,112	3,103,077	3,103,077
Remove patients of – closed hospitals	Nearby hospital's closure	NA ¹	NA ¹	-0.11** (0.04)	NA ¹	0.003 (0.004)	0.001† (0.001)
	Adjusted R^2	NA	NA	0.08	NA	0.04	0.03
	Number of observations	NA	NA	13,398,447	NA	13,398,447	13,398,447
Travel distance	Nearby hospital's closure	1.85* (0.82)	0.014 (0.007)	-0.18*** (0.041)	-0.67 (0.56)	-0.0003 (0.005)	0.003* (0.001)
	Adjusted R^2	0.78	0.86	0.08	0.84	0.04	0.03
	Number of observations	28,757	28,757	10,905,790	20,882	10,905,790	10,905,790
Study sample – years 2001-2017	Nearby hospital's closure	1.87* (0.94)	0.011 (0.010)	NA ²	-0.13 (1.00)	NA ²	NA ²
	Adjusted R^2	0.78	0.86	NA	0.84	NA	NA
	Number of observations	28,983	28,983	NA	20,996	NA	NA
Placebo test – random treatment	Nearby hospital's closure	-0.13 (0.31)	0.002 (0.003)	0.012 (0.01)	0.14 (0.44)	-0.0003 (0.001)	0.001 (0.001)
	Adjusted R^2	0.78	0.86	0.08	0.84	0.04	0.03
	Number of observations	28,212	28,212	10,711,050	20,389	10,711,050	10,711,050

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. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; †A marginal increase at p -value < 0.10 . ¹Removing the patients of closed hospitals only affects the patient-level analyses. ²Including the hospital level data only affects the hospital-level analyses.

Chapter 2

The Allocation of Funds in Healthcare: Which Hospitals to Support?¹

2.1 Introduction

One critical challenge the U.S. healthcare system is facing is allocating the funds wisely to satisfy various dimensions of healthcare, where there often is a trade-off among objectives of access, quality, and cost. In particular, with increasing pressure to address the rising healthcare cost, there has been significant attention on how to allocate funds to the U.S. hospitals efficiently (currently, inpatient care cost takes up more than 30% of the total healthcare expenditure) (Centers for Medicare & Medicaid Services, 2018). Policymakers have been re-examining the efficiency of existing hospital payment policies through Medicare, the social insurance for the U.S. elderly, and the largest payer (e.g., 25% of total spending) for hospital care. Through making decisions on how funds are allocated, these payment policies not only address the healthcare cost but also shape the provider market structure both in terms of hospitals' quality in the short term and hospitals' operating status in the long term. Thus, decisions on how to pay for U.S. hospitals should be made wisely with consideration of various dimensions of societal consequences.

¹Co-authored with Soroush Saghafian

Two Medicare hospital payment programs for the hospitals that serve vulnerable populations have recently been under consideration for reform. These are Medicare Disproportional Share Hospital (DSH) payment for the hospitals that treat a greater share of low-income patients, and the Critical Access Hospital (CAH) program for the rural hospitals with high fixed costs. The safety-net hospitals or rural hospitals are often at a disadvantage under the current payment system that Medicare usually pays hospitals, i.e., the inpatient prospective payment system (IPPS) currently employed by the Centers for Medicare & Medicaid Services (CMS). This is because the payment to hospitals is made based on prospectively set rates per procedure, whereas these hospitals have a higher proportion of uncompensated care or small sizes. Thus, DSH and CAH programs provide extra support for hospitals that are more likely to experience financial difficulties associated with serving a more vulnerable population or achieving economies of scale. Numerous evidence shows that these policies have been successful at (a) improving the quality of care through funding for extra resources and (b) preventing the closure of these essential safety-net hospitals that vulnerable population relies on by improving their financial health (Baicker, K. and Staiger, D. O., 2005; Gowrisankaran *et al.*, 2018; Baicker, K. and Staiger, D. O., 2005; Li *et al.*, 2007; Neuhausen *et al.*, 2014; Casey *et al.*, 2015; Cole *et al.*, 2014; Graves, 2012; Carroll, C., 2019). More recently, however, questions have been raised on the adequacy of the level of funding and the efficiency of these programs for achieving such objectives (Barnes and Harp, 2018; Nguyen and Sheingold, 2011).

Both programs have received criticism in that the amount of spending exceeds the societal benefit generated by these programs. Thus, the recent healthcare reform and subsequent legislations have made several attempts to reduce the spending on these programs. Yet, substantial debate remains on to what extent the funding should be reduced, and which hospital should bear a greater burden of reduced payment. Furthermore, these programs have important differences in targeting a mutually exclusive set of hospitals: the majority of DSH program funds go to urban, large, and/or teaching hospitals, whereas the majority of CAH funds go to rural hospitals (Fishman and Bentley, 1997). Thus, the decision to

allocate funds between these programs has different implications for society. In particular, the decision embeds an inherent trade-off between access and quality. Allocating funds to the hospitals that are more likely to treat a greater number of patients (e.g., urban areas) will improve the utility of more patients. Yet, such a decision will have a greater detrimental impact on patient access through accelerating rural hospital closures or increasing the travel cost to farther, urban areas.

To ensure that policy decision incorporates various consequences, it is critical to accurately reflect the impact of funding decisions on both access and quality, and ultimately societal welfare, which is embedded in patients' preference. A large body of literature shows that a patient's demand declines with distance and increases with the quality of care (Gutacker *et al.*, 2016; Tai *et al.*, 2004; Capps *et al.*, 2010; McNamara, 1999; Chandra and Staiger, 2007). Thus, to understand the true effect of hospital closure, it is important to quantify the overall impact on societal welfare. For example, when patients who choose the nearby worse quality hospital migrate to the further better quality hospital, they will experience an increase in travel distance, but also receive better quality care. Thus, this may not harm—or even benefit—the overall societal welfare. Although studies have observed and quantified these trade-offs, these preferences have not been directly incorporated into providing the decision rules for identifying optimal investment of funds. In this chapter, we provide a decision-making framework to evaluate the allocation policies both within and between DSH and CAH payment programs in order to improve access, quality, and equity.

2.1.1 Policy Background

Medicare DSH Program. Since 1986, the Medicare DSH payment programs have offset the cost from treating large numbers of uninsured and Medicaid patients for safety-net hospitals.² Currently, 3,109, or about 75%, of U.S. short-term acute care hospitals receive payments, where most (93%) of them are large hospitals in urban areas. The eligibility and

²There are Medicare and Medicaid DSH payment programs. In this chapter, we focus on the Medicare DSH program, where the payments flow directly from the federal government to individual hospitals, whereas the Medicaid DSH program lets each state have significant flexibility to structure and determine the payment.

the size of payments are mostly determined by the Medicare DSH formula that accounts for the inpatient days attributable to Medicare patients with Supplemental Security Income (SSI).³ The Affordable Care Act (ACA) has proposed major changes to the Medicare DSH program by mandating to cut the fund by \$22.1 billion between fiscal years 2014 and 2019, with an anticipation that the insurance expansion would increase DSH program recipient's revenues. The scheduled reduction has subsequently been delayed since first scheduled with rising concerns on its impact on the safety-net hospitals, and the future course of payment cut is still unclear.

CAH Program. The CAH program was established in 1997 to help small and rural hospitals by providing cost-based reimbursement since these hospitals are financially vulnerable under the IPPS. Currently, Medicare pays CAHs 101% of their allowable costs for most services. Although CAH has been successful at preventing rural hospital closures, there have also been concerns about the increasing program cost, the increasing number of CAH payment recipients, and the overall value of the program (Rosko and Mutter, 2010). For example, Medicare's payments to CAHs represented 5% of all Medicare payments to hospitals in 2016. Thus, there has been legislative discussion by CMS to reduce the cost-based reimbursement level from 101% to 100% of the allowable costs.

Another aspect of the policy under debate is the distance requirement. A CAH is required to be located more than a 35-mile drive from another hospital or CAH or more than 15 miles in areas with mountainous terrain or only secondary roads.⁴ Initially, states could waive the requirement as "necessary providers" during 1997-2005, and as a consequence, about two-thirds of all current CAHs that do not meet the distance requirement was exempted through the necessary providers' exemption. Medicare discontinued the exemptions in 2006, but all existing CAHs were allowed to maintain their status (Center for Medicaid and State Operations/Survey and Certification Group, 2007; Levinson and General, 2013). More

³SSI is a means-tested federal welfare program that provides cash assistance to the elderly or disabled population.

⁴Other requirements include having 25 or fewer acute hospital beds, having an average length of stay of fewer than 96 hours, and providing 24-hour emergency services.

recently, there have been discussions on eliminating the existing exemptions.⁵ Such changes will result in a reduction in funding for hospitals that are in rural areas, and may, in turn, aggravate the recent trend in hospital closures (Neuhausen *et al.*, 2014; Kaufman *et al.*, 2016).

2.1.2 Research Questions

The DSH and CAH programs described in the previous section are two main policy levers designed to support different types of hospitals. The former mainly aims at supporting urban hospitals, and the latter primarily provides support for rural hospitals. Motivated by these policy levers, our central question in this chapter is: *which hospitals should be supported?*. In other words, *which hospitals should be targeted by policymakers to receive supportive investments?*

In addressing this question, we note that in addition to considering attributes such as urban versus rural, the government can target hospitals based on their quality (e.g., health outcomes). On the one hand, investing in the quality of hospitals that have lower-than-desired outcomes can help them to become competitive with higher-quality hospitals in attracting patients. On the other hand, it might be better to focus on improving the quality of hospitals that already have good outcomes and even close the low-quality hospitals so that patients do not receive lower-than-desired care.

These decisions affect patient utility in two main ways: (a) quality of care they receive (e.g., health outcomes), and (b) access to care (e.g., difficulty caused by travel distance). Thus, taking a utilitarian approach, we study the impact of such decisions on the overall utility that patients receive and shed light on policies that can improve societal outcomes. In particular, we adopt a patient-level regression-based self-selection model (Roy, 1951; Heckman *et al.*, 1990), in which patients self-select hospitals that provide them with the highest utility. We then examine the policymakers' problem to improve societal outcomes such as social welfare (defined as the expected sum of achieved utilities by all patients) and

⁵In 2011, the Congressional Budget Office (CBO) proposed restricting the eligibility policy for CAH, which depends on both the size and the location of the hospitals. In 2013, the Office of Inspector General recommended the CMS to remove the "necessary provider" exemption, and legislation was submitted in 2015 to prohibit CAH designation for facilities that are less than 10 miles from the nearest hospital.

Table 2.1: Data Sources

Data source	Variables used	Level	Years
FFS inpatient Medicare claims	Hospital visits (choice, diagnosis, procedures)	Patient	2014
Medicare Beneficiary Summary Files	Demographic, socioeconomic, clinical, enrollment	Patient	2014
Hospital Compare	Quality (overall, risk-adjusted readmissions and survival)	Hospital	2015
HCAHPS	Patient experience	Hospital	2014
Medicare Cost Report	Facility characteristics, utilization, financial	Hospital	2014

social inequality (defined as the variance of achieved utilities among all patients).

Finally, we calibrate our models by assembling and utilizing various sources of data on patient and hospital characteristics (see Table 2.1 for our sources of data). This allows us to (a) estimate various parameters of our models by making use of data related to over 300,000 patients who have received care in 1,515 U.S. hospitals, and (b) provide various insights as well as actionable recommendations for policymakers.

2.1.3 Summary of Main Findings and Contributions

Below we summarize our main findings and contributions.

- We show that because rural patients' distance to both rural and urban hospitals are more dispersed than urban patients, patients living in rural areas tend to select the rural, low-quality hospitals even when their average distance to these two types of hospitals are similar. We use this finding to show that, because of such selection, investing in improving the quality of lower-quality rural hospitals can be more welfare-improving than doing the same in higher-quality urban hospitals.
- Our results provide several policy insights. First, we find that reducing the number of hospitals that receive quality improvement investments is more efficient than reducing the amount of investment each hospital receives. Thus, given any fixed amount of money, policies that try to cast a wide net and support many hospitals, but each with a limited amount of money, could be improved if instead they focus on specific hospitals and provide them with more support. Our findings on hospitals that should be targeted for this purpose can help policymakers use any fixed amount of money

more effectively. Second, our findings show that investing in further improving the quality of high-quality hospitals is generally more efficient than doing so in low-quality hospitals. That is, in general, it is better for policymakers to invest where care delivery has “strength” instead of “weakness.” However, there are exceptions to this general finding. For example, consider two hospitals labeled as A and B, where B is in a geographical area with a substantially lower hospital density while having a slightly lower quality compared to A. Then we observe that investing in Hospital B is better than Hospital A. Third, given that there is a correlation between quality (e.g., good versus bad patient outcomes) and geography (e.g., urban versus rural), combining both quality and geography in targeting hospitals for allocation of funds is more efficient than separately targeting hospitals based on these aspects (e.g., the targeting strategy in DSH and CAH programs). Fourth, keeping a greater number of hospitals open, under optimal investment, always generates greater social welfare than closing some, but closing less desirable hospitals (e.g., lower quality or rural) can generate greater social welfare than keeping a greater number of hospitals open, under suboptimal investment. Our findings contribute to the ongoing policy debate on the hospital payment policy by suggesting decision-making rules that can improve societal welfare.

- We make use of our findings and provide a heuristic to guide policymakers in their efforts. The heuristic takes into account the quality and geography of patients to identify the optimal allocation of funds to different hospitals. We first show that in the absence of heuristic, the legislative attempts to reduce DSH or CAH fundings are unlikely to be welfare-improving. Next, we show through simulation analyses calibrated with real-world data that the heuristic results in 2.9% increase in social welfare compared to the original fund reduction strategies. The improvement achieved by the heuristic is mainly due to the fact that it provides a good balance between the underlying quality and access trade-offs. For example, among the 1,515 hospitals in our data set, it targets 980 hospitals for quality-improvement investments, recommends 303 hospitals for closure, and allows other hospitals to operate as they do (i.e., without

any additional support or change).

- Based on our findings and above-mentioned insights, we provide specific policy recommendations by shedding light on ways the current CAH and DSH programs can be improved. Specifically, our results recommend the following adjustments to these programs: (1) restricting CAH eligibility to rural hospitals and reducing DSH support from hospitals in Medicaid expansion states, and (2) incorporating quality of care dimension into CAH payment criteria.

2.2 Literature Review

Healthcare Resource Allocation. Numerous studies in healthcare operations examine how to allocate scarce healthcare resources to improve social welfare in various contexts from communicable diseases, chronic diseases, to inpatient care (Deo *et al.*, 2013; Earnshaw *et al.*, 2002; Su and Zenios, 2005; Bertsimas *et al.*, 2013; Branas *et al.*, 2000; Mills *et al.*, 2018; Khademi *et al.*, 2015). Within the literature, our work is relevant to studies on provider payment, in particular, Medicare’s provider payment policy. Several studies have examined the optimal incentive design for healthcare provider payment, such as the hospice reimbursement (Ata *et al.*, 2013), bundled payment (Gupta and Mehrotra, 2015; Adida *et al.*, 2016), Hospital Readmissions Reduction Program (Zhang *et al.*, 2016), the Medicare Shared Savings Program (Aswani *et al.*, 2019), and the pay-for-performance programs (Bastani *et al.*, 2016). These studies have used various analytic models such as a game-theoretic model (Gupta and Mehrotra, 2015; Zhang *et al.*, 2016) and principal-agent framework (Bastani *et al.*, 2016; Aswani *et al.*, 2019). Our study examines the two hospital payment policies—DSH and CAH program—and focuses on incorporating the patients’ preferences into allocation decisions instead of the hospitals’ strategic behaviors.

Provider Choice. Our work is related to the empirical literature studying the demand for healthcare (Lu and Lu, 2018; Dong *et al.*, 2018; Moscone *et al.*, 2012; Chernew *et al.*, 1998). Various drivers of patient choices, such as distance, quality, insurance, and referral,

have been examined. Chandra and Staiger (2007) uses an economic model where patients make a trade-off between distance and quality to examine the productivity spillover of the providers' care volume. (Gowrisankaran *et al.*, 2018) also uses a model to examine the impact of the CAH program on social welfare. Capps *et al.* (2010); Raval and Rosenbaum (2018); McNamara (1999) also use the choice model to examine how hospital closures or mergers affect social welfare. Our study differs from these studies in that we prescribe the optimal resource allocation that takes into account patients' choice behavior. Our study is similar in vein as Saghafian, S. and Hopp, W. J. (2020), which also uses the distance-quality trade-off of patient choice to examine a public policy question—the role of quality transparency (see, also, Saghafian and Hopp (2019) for some policy-relevant discussions). However, our study differs in that we study the allocation of funds that determines the quality and the operating status of hospitals.

Healthcare Policy. Our study is also related to health and public economics literature, particularly on hospital closures. Streams of literature have examined the welfare impact of hospital closures on distance, hospital characteristics, patient demographic and socioeconomic characteristics, insurance type, and cost (Avdic, 2016; Capps *et al.*, 2010; Weber, 2014). Capps *et al.* (2010) examines the impact of hospital closures/bailouts on society through a patients' choice model that incorporates the distance. It compared the utility loss from hospital closure with cost savings from improved efficiency and found that the cost-savings from efficiency gains outweigh patient costs of increased distances. Weber (2014) measures the welfare gain from ambulatory surgery centers (ASCs) from the patient choice model and finds that closures of ASCs result in welfare loss. Song and Saghafian (2019a) shows that hospitals can experience a negative spillover effect on quality when a neighbor hospital closes. Finally, Gowrisankaran *et al.* (2018) shows that the conversion of hospitals to CAH reduces patient welfare through a reduction in bed sizes, but the benefits from forestalling closure offset the reduction in utility. Our study provides the framework to determine under which conditions the benefits of the closure of quality investment outweigh the costs.

2.3 The Model

There are two levels of decision-makers in the model. First, patients choose hospitals according to their own utility function, which depends on the quality of the hospital, the access convince (i.e., the geographical proximity), and some other patient-specific factors. Next, the social planner makes resource allocation decisions to maximize social welfare (the sum of achieved utilities by all patients) under resource constraints (budgets that represent available funds).

Patient Utility. To model patients' choice behavior, we modify and adopt a framework of self-selection that has been widely used in labor economics (Roy, 1951; Heckman *et al.*, 1990) and more recent studies on patients' choice of providers (Chandra and Staiger, 2007; Saghafian, S. and Hopp, W. J., 2020). In designing a model to capture patient utility, we note that quality and distance are known to be two key elements affecting the patient choice of hospitals (see, e.g., Gutacker *et al.* (2016); Tai *et al.* (2004); Capps *et al.* (2010); McNamara (1999); Chandra and Staiger (2007); Saghafian, S. and Hopp, W. J. (2020), and the references therein).

Suppose we have hospitals j , where $j \in \mathcal{J} \triangleq \{1, 2, \dots, m\}$ and patients i , where $i \in \mathcal{I} \triangleq \{1, 2, \dots, n\}$. For patient i , the utility of choosing hospital j , assuming the hospital is open, is

$$U_i^j = Q^j - \alpha D_i^j + \varepsilon_i, \quad (2.1)$$

where Q^j denotes the quality of hospital j , and D_i^j is a random variable with distribution $\mathcal{F}_{D^j}$ that represents the distance to hospital j , α is a coefficient that represents the trade-off between quality and distance, (e.g., willingness to travel one more mile to receive care from a hospital that has 1% higher risk-adjusted survival rate in a specific treatment). Finally, in (1), ε_i represents other factors (including psychological or behavioral) that may influence the choice of the patient but are not systematic or observable by policymakers.

In our analyses, we assume that Q^j does not depend on the patient (e.g., it is the average quality of hospital j). In Section 5, we make use of our data along with the maximum

likelihood approach to estimate α , and also validate the estimated value against values provided in other studies. We also assume that the patient-specific error term ε_i is not correlated with the hospital. This assumption will hold reasonably well if the impact of across-hospital variations of patient-specific (e.g., psychological or behavioral) on outcomes is likely smaller than the within-hospital variations.

To take the operational status of hospitals into account, we let y^j be a binary variable that indicates the operation status of hospital j , i.e., $y^j = 1$ if hospital j is open, and $y^j = 0$ if closed. We then model the fact that closed hospitals cannot be chosen by patients by assuming that the utility of choosing a closed hospital is a large negative value denoted by $-M \ll 0$. Specifically, we let $\hat{U}_i^j$, denote patient i 's utility of choosing hospital j , incorporating information on whether it is open or closed, where

$$\hat{U}_i^j = U_i^j y^j - M(1 - y^j) = (Q^j - \alpha D_i^j + \varepsilon_i) y^j - M(1 - y^j). \quad (2.2)$$

In this setting, each patient i chooses hospital k , if it provides patient i with the highest value of $\hat{U}_i^j$ (i.e., if $k = \arg\max_j \hat{U}_i^j$).

Hospital Quality. To reflect the fact that a hospital quality can be improved, if policymakers decide to invest in it, we assume

$$Q^j(x^j) = q^j + kx^j, \quad (2.3)$$

where x^j denotes the investment amount in hospital j , q^j is the baseline quality of hospital j , and $k \geq 0$ is a coefficient that reflects the effectiveness of investment in improving quality. Our model is based on the evidence that hospitals have used DSH or CAH funding to improve quality, such as quality assurance activities (Casey *et al.*, 2015; Baicker, K. and Staiger, D. O., 2005). In Section 5, we explain how we obtain this parameter k from the literature.

Social Welfare. We denote social welfare by $\bar{\Pi}$, where $\bar{\Pi}$ is a random variable defined as the mean of *achieved* (i.e. post patient selection) individual utility. Thus, we have

$$\bar{\Pi} = \frac{\sum_{i \in \mathcal{J}} \hat{U}_i^*}{n}, \quad (2.4)$$

where n is the population size of patients, and $\hat{U}_i^*$ is the achieved utility by patient i in the population:

$$\hat{U}_i^* = \max_{j \in \mathcal{J}} \hat{U}_i^j. \quad (2.5)$$

Policymaker's Problem. We assume that a policymaker wants to maximize the expected social welfare defined as:

$$\mathbb{E}[\bar{\Pi}] = \frac{\mathbb{E}[\sum_{i \in \mathcal{J}} \hat{U}_i^*]}{n} = \frac{\sum_{i \in \mathcal{J}} \mathbb{E}[\hat{U}_i^*]}{n}. \quad (2.6)$$

In this setting, the policymaker has two levels of decisions to make: Given an available amount of funds (a budget), which hospitals to invest in (support) and which hospitals to close so that the overall outcomes are improved? In making these decisions, the policymaker needs to consider the quality-access tradeoff mentioned earlier, and target the correct set of hospitals considering this tradeoff.

Formally, the policymaker's problem can be represented as

$$\max_{\underline{x}, \underline{y}} \mathbb{E}[\bar{\Pi}] \quad (2.7)$$

subject to

$$0 \leq \sum_{j \in \mathcal{J}} x^j \leq B \quad (2.8)$$

$$x^j \geq 0 \quad \forall j \in \mathcal{J} \quad (2.9)$$

$$y^j \in \{0, 1\} \quad \forall j \in \mathcal{J}, \quad (2.10)$$

where B denotes the budget, $\underline{x} = (x^j)_j \in \mathcal{J}$, and $\underline{y} = (y^j)_j \in \mathcal{J}$. By definition of optimality, we know that for a fixed number of hospitals $|\mathcal{J}|$, we can find the optimal solution by comparing each closure scenario, conditional on optimal budget allocation for each closure scenario.

2.4 Analysis

We analytically examine the simplest case, where there are only two hospitals, i.e., $\mathcal{J} = \{1, 2\}$. In Section 6, we utilize our data and make use of simulation analyses to generate further insights by considering 1,515 hospitals in the U.S. Our detailed proofs are provided in the Appendix.

To analyze the case with two hospitals, let the vector of random variables denoting the distance to hospitals be

$$\underline{D} = (D^1, D^2). \quad (2.11)$$

Denote by $F_{\underline{D}}$ the joint distribution of $\underline{D}$, where⁶

$$E[\underline{D}] = \underline{\mu} = (\mu^1, \mu^2), \quad (2.12)$$

$$Var(\underline{D}) = \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}. \quad (2.13)$$

In what follows, we denote the four possible closure scenarios of the two hospitals by $\underline{y}_k$ for $k = 1, 2, 3, \dots, 4$: $\underline{y}_1 = (1, 0)$ represents the case when Hospital 1 is open and Hospital 2 is closed; $\underline{y}_2 = (0, 1)$ is when Hospital 1 is closed and Hospital 2 is open; $\underline{y}_3 = (0, 0)$ is when both hospitals are closed; and $\underline{y}_4 = (1, 1)$ represents the case when both hospitals are open. We will first solve for optimal allocation $\underline{x}_k^*$ for each $\underline{y}_k$, and compare $\mathbb{E}[\bar{\Pi}_k^*] = \mathbb{E}[\bar{\Pi}(\underline{x}_k^*, \underline{y}_k)]$ to find the optimal solution $\underline{y}^*$. Since $\mathcal{J} = \{1, 2\}$, we can express the investment vector as $\underline{x} = (x, B - x)$, where x is the investment for Hospital 1. The labeling of hospital matters, as we discuss in the next section.

2.4.1 Quality Investment Decisions

The first three cases of $\underline{y}$, i.e., $\underline{y}_1$, $\underline{y}_2$, and $\underline{y}_3$ are trivial. When only one hospital is open, the optimal strategy is to invest all budget in the open hospital. When both hospitals are closed,

⁶As we will see the correlation between distance to the two hospitals in (13) (induced by the geographical location of hospitals and patients) plays an important role in the analyses.

$\mathbb{E}[\bar{\Pi}_3^*] = -M \ll 0$, so $\underline{y}_3$ cannot be the optimal solution. Thus, we can focus on analyzing the case where both hospitals are open, i.e., $\underline{y}_4$, and compare the optimal social welfare with that when only one of the hospitals is kept open.

When both hospitals are open, we have

$$\mathbb{E}[\bar{\Pi}_4(x)] = \frac{\sum_{i \in \mathcal{J}} \mathbb{E}[\hat{U}_i^*(x)]}{n} = p\mathbb{E}[U_i^1 | U_i^1 > U_i^2] + (1-p)\mathbb{E}[U_i^2 | U_i^1 < U_i^2], \quad (2.14)$$

where p is the probability that a patient chooses Hospital 1, i.e.,

$$p = \Pr[U_i^1 > U_i^2] = \Pr\left[\frac{(q^1 - q^2) + k(2x - B)}{\alpha} \geq D^1 - D^2\right]. \quad (2.15)$$

As (15) shows, the value of p depends on the investment amount x , the difference in baseline qualities and distance to the two hospitals, and some other parameters. To characterize (15) (and hence (14)), we first need some definitions and preliminary results.

Definition 1. We define the following new variables (i) $\bar{D}^j = D^j - \mu^j$ such that $E(\bar{D}^1) = E(\bar{D}^2) = 0$.

(ii) $\tilde{D} = \bar{D}^1 - \bar{D}^2$.

(iii) $c_0 = (q^1 - q^2) - \alpha(\mu^1 - \mu^2)$.

(iv) $c(x) = (q^1 - q^2) + k(2x - B) - \alpha(\mu^1 - \mu^2) = c_0 + k(2x - B)$.

(v) $\sigma^2 = \sigma_{11} + \sigma_{22} - 2\sigma_{12}$.

(vi) $a_1 = \frac{\sigma_{11} - \sigma_{12}}{\sigma^2}$ and $a_2 = a_1 - 1 = \frac{-\sigma_{22} + \sigma_{12}}{\sigma^2}$.

Definition 1 provides various change of variables that allow us to establish clear insights.⁷ In Definition 1, $\bar{D}^j$ is the demeaned random component of distance, where $\text{Var}(\bar{D}^1, \bar{D}^2) = \text{Var}(D^1, D^2) = \Sigma$. $\tilde{D}$ is the stochastic component of each patient's difference in distance to the two hospitals. c_0 is the deterministic difference in baseline qualities and weighted/adjusted distance to the two hospitals, $c(x)$ similar to c_0 but instead of baseline qualities, it considers qualities post-investment. a_1 and a_2 are the scaled difference in variance and covariance of the distance distribution.

⁷These change of variables are primarily motivated by some results in Heckman *et al.* (1990).

Observe that, based on Definition 1, we can rewrite $\bar{D}^j = a_j \tilde{D} + \hat{D}$, where $\hat{D} = a_1 \bar{D}^2 - a_2 \bar{D}^1$. Notably, we have that $E(\tilde{D}) = 0$, $E(\hat{D}) = 0$, and $\tilde{D}$ and $\hat{D}$ are independent by construction. Each $\bar{D}_i^j$ and $\tilde{D}$ are independent and identically distributed random variable with $F_{\bar{D}^j}$ and $F_{\tilde{D}}$, respectively. Using these, we can rewrite the patient utility as

$$U_i^j = (kx^j + q^j) - \alpha(\mu^j + \bar{D}_i^j) + \varepsilon, j \in \mathcal{J}, \quad (2.16)$$

where

$$U_i^1 - U_i^2 = c(x) - \alpha \tilde{D}_i. \quad (2.17)$$

Thus, when both hospitals are open, we have

$$\begin{aligned} \mathbb{E}[\bar{\Pi}_4] &= p\mathbb{E}[U_i^1 | U_i^1 > U_i^2] + (1-p)\mathbb{E}[U_i^2 | U_i^1 < U_i^2] \\ &= p \left[q^1 + kx - \alpha\mu^1 - \alpha a_1 \mathbb{E}[\tilde{D}_i | \tilde{D}_i < c/\alpha] \right] + (1-p) \left[q^2 + k(B-x) - \alpha\mu^2 - \alpha a_2 \mathbb{E}[\tilde{D}_i | \tilde{D}_i > c/\alpha] \right], \end{aligned} \quad (2.18)$$

where

$$p = \Pr[c(x)/\alpha \geq \tilde{D}_i]. \quad (2.19)$$

Definition 2. Consider two hospitals, Hospital A and Hospital B. Hospital A is said to be more desirable at baseline than Hospital B if $c_0 = (q^A - q^B) - \alpha(\mu^A - \mu^B) \geq 0$, and Hospital B is said to be more desirable at baseline than Hospital A otherwise. Similarly, Hospital A is said to be more desirable at investment level x than Hospital B if $c(x) = (q^A(x) - q^B(x) - \alpha(\mu^A - \mu^B)) \geq 0$, and Hospital B is said to be more desirable at investment level x than Hospital A otherwise.

WLOG, for the rest of the analysis, we will label the more desirable hospital at baseline as

Hospital 1.⁸ We also note that $\frac{\partial p}{\partial c} > 0$ for all c , i.e., as the desirability of Hospital 1 increases (e.g., through greater investment), more people choose Hospital 1.

⁸Hospital 1 can be more desirable at baseline than Hospital 2 through either (i) $q^1 > q^2$ and $\mu^1 < \mu^2$, i.e., having both better quality and shorter mean distance; (ii) $q^1 < q^2$ and $\mu^1 \ll \mu^2$, i.e., having worse quality but much shorter mean distance; or (iii) $q^1 \gg q^2$ and $\mu^1 > \mu^2$, i.e., having much better quality and longer mean distance. Thus, in order for the hospital with greater mean distance to be more desirable, it needs to have substantially better quality, whereas the hospital with shorter mean distance can have equivalent or worse quality and still be more attractive.

Definition 3. *Hospital 1 and Hospital 2 are competitive within budget B if there exists a non-negative investment threshold such that $c(x) \leq 0$ for all $x \leq x_{thr}$.*

Given that Hospital 1 is more desirable at the baseline than Hospital 2, the two hospitals are competitive if Hospital 2 becomes the desirable hospital at some feasible investment level (i.e., if their relative desirability changes as a result of an investment that is feasible within the budget). For the rest of the analysis, we assume that Hospital 1 and Hospital 2 are competitive, because the allocation decision when they are not competitive is trivial—it is always optimal to invest in Hospital 1, the more desirable hospital.

To analytically examine the policymaker’s problem of optimizing the social welfare, in what follows we assume that the distance distribution, $F_{\tilde{D}}$, is log concave, a property that holds for a wide range of distributions.⁹ More specifically, Figure 2.1 plots the distribution of each patient’s distance to their two nearest hospitals from our data, overall and stratified by patients in rural or urban areas based on their zip codes.¹⁰ Our data depicted in Figure 2.1 shows that the distribution of patients’ log distance to hospitals is very close to normal. Thus, in our main analysis, we assume that $F_{\tilde{D}}$ is log normal. We next present the definitions and known properties of log concave and normal random variables that we need for our analyses.

Lemma 1. *[Properties of log concavity (Heckman et al., 1990).] If $\tilde{D}$ is log concave, the following properties hold:*

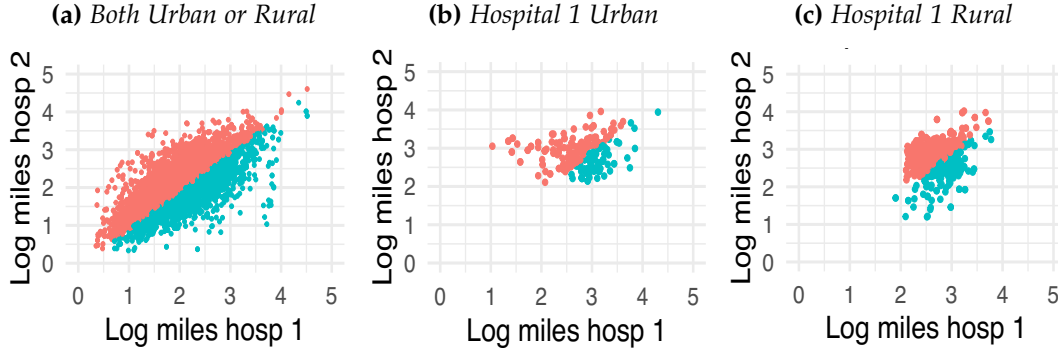
$$(i) \ 0 < \frac{\partial \mathbb{E}[\tilde{D} | \tilde{D} \geq d]}{\partial d} < 1. \quad (ii) \ 0 < \frac{\partial \mathbb{E}[\tilde{D} | \tilde{D} \leq d]}{\partial d} < 1. \quad (iii) \ \frac{\partial \text{Var}[\tilde{D} | \tilde{D} > d]}{\partial d} \leq 0. \quad (iv) \ \frac{\partial \text{Var}[\tilde{D} | \tilde{D} \leq d]}{\partial d} > 0. \quad (v) \ \text{Var}[\tilde{D} | \tilde{D} \geq d] \leq \sigma^2.$$

Lemma 2. *[Properties of normality (Heckman et al., 1990).] Let $\tilde{D}$ be a normal random variable with $\mathbb{E}[\tilde{D}] = 0$ and $\text{Var}(\tilde{D}) = \sigma^2$. Define $\lambda(d) = \mathbb{E}(\tilde{D} | \tilde{D} > d)$, then for $d \in (-\infty, \infty)$ the*

⁹Log concave distributions include normal, exponential, logistic, uniform, chi-squared (with degrees of freedom ≥ 2), gamma (with shape parameter ≥ 1), beta (with both parameters ≥ 1), extreme value, and Weibull densities (with shape parameter ≥ 1).

¹⁰A zip code is defined as rural if the Rural-Urban Commuting Area code is greater than 4. The hospital was also defined as rural if it was designated as a CAH (Hart et al., 2005).

Figure 2.1: Log Distance to Hospital 1 and Hospital 2



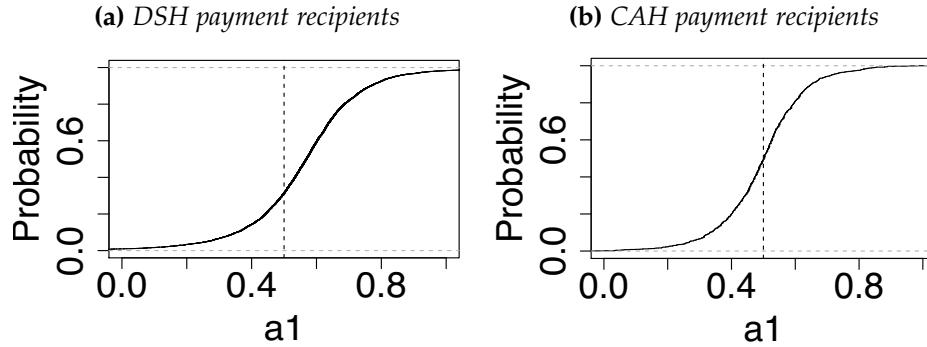
Note. The distance distributions are obtained from our study population. The estimated a_1 are (a) $a_1 = 0.52$, (b) $a_1 = 0.55$, and (c) $a_1 = 0.47$. The choices are hypothetical under the assumption that the differences in desirability between the two hospitals in each patient's choice set is $c_0 = 0.1$. Hosp indicates hospital. Patients who choose Hospital 1 are indicated in red. Patients who choose Hospital 2 are indicated in blue.)

following properties hold:

$$(i) \lambda(d) = \frac{1}{\sqrt{2\pi}} \frac{\exp(-d^2/2)}{\Phi(-d)} > \max\{0, d\}. \quad (ii) \mathbb{E}[\tilde{D}|\tilde{D} < d] = -\mathbb{E}[\tilde{D}|\tilde{D} > -d]. \quad (iii) 0 < \frac{\partial \lambda(d)}{\partial d} = \lambda'(d) = \lambda(d)(\lambda(d) - d) < 1. \quad (iv) 0 < \frac{\partial^2 \lambda(d)}{\partial (d)^2}. \quad (v) \lambda(d)1 < \lambda(d_2) \text{ for } d_1 < d_2. \quad (vi) \lambda'(d_1) < \lambda'(d_2) \text{ for } d_1 < d_2.$$

Definition of Geography. Next, we formally define rural and urban hospitals based on a_1 , the measure of relative dispersion of distance between the two hospitals. When the distance to both hospitals are equally spread out, i.e., $\sigma_{11} = \sigma_{22}$, then $a_1 = 1/2$. When $\sigma_{12} > \sigma_{22}$, $a_1 > 1/2$, i.e., patients' distance to Hospital 1 is relatively more dispersed than the distance to Hospital 2. This is likely the situation when Hospital 1 is more urban compared to Hospital 2. When $\sigma_{12} < \sigma_{11}$, $a_1 < 1/2$, i.e., the distance to Hospital 2 is relatively more dispersed than the distance to Hospital 1. This is likely when Hospital 1 is more rural compared to Hospital 2. Thus, we use a_1 as a measure of geography between the two hospitals, where $a_1 = 1/2$ indicates the patients in the market have an equal amount of rurality/urbanity to both hospitals, $a_1 > 1/2$ indicates Hospital 1 is more urban, and $a_1 < 1/2$ indicates Hospital 1 is more rural.

Figure 2.2: Cumulative Distribution of Mean a_1 for all Hospital Pairs



Our definition of rurality/urbanity using a_1 is empirically supported by our data. The mean a_1 of all U.S. hospital pairs was 0.51. Figure 2.2 shows the cumulative distribution of mean a_1 of all hospital pairs in our sample, stratified by the DSH or CAH program participation status of Hospital 1 (i.e., the more desirable hospital). Consistent with the fact that the majority of DSH payment recipients are urban and CAH payment recipients are rural, the distribution of a_1 was skewed above 0.5 for DSH payment recipients, and below 0.5 for CAH payment recipients. These data support our assumption of using $a_1 = 1/2$ as a cut-off for defining rural/urban hospitals.

Geography and Hospital Choice. We next examined the relationship between the geography and choice of hospital. Figure 2.1 shows patients' choice of hospitals when both hospitals are equally urban or rural; when Hospital 1 is urban and Hospital 2 is rural; or when Hospital 1 rural and Hospital 2 urban, and when Hospital 1 is slightly more desirable at baseline than Hospital 2 ($c_0 = 0.1$). The patients with distance distribution that lies below the line $D_2 = c_0 + D_1$ will choose Hospital 2 (color in blue), and those above the line will choose Hospital 1 (color in red). As expected, patients select into hospitals based on their own geography. For example, patients who are close to both hospitals (i.e., urban patients, in the lower-left corner of the plane) tend to choose the urban hospital, whereas the patients who are far from both hospitals (i.e., rural patients, in the upper right corner of the plane) tend to choose the rural hospitals. More importantly, based on Figure 2.1, we make the

following observations.

Observation 1. *Patients prefer their own geography to the extent that they choose a worse quality hospital. For example, if the urban hospital has a choice between a lower quality urban hospital and a higher quality rural hospital, they still prefer urban hospitals. Such behavior is similar for rural patients, i.e., they prefer lower quality rural hospitals to a higher quality urban hospital.*

Post-Selection Utility Difference. We next examine the difference in expected utility between those who choose Hospital 1 and those who choose Hospital 2. We define this quantity by

$$\tau = \mathbb{E}[U_i^1 | U_i^1 > U_i^2] - \mathbb{E}[U_i^2 | U_i^1 < U_i^2] = c - \alpha a_1 \mathbb{E}[\tilde{D}_i | \tilde{D}_i < c/\alpha] + \alpha a_2 \mathbb{E}[\tilde{D}_i | \tilde{D}_i > c/\alpha], \quad (2.20)$$

and note that a high value for $|\tau|$ implies inequality (in terms of achieved utilities) among those who treated.

Proposition 1. *Suppose $\tilde{D}$ is a random variable with a normal distribution. Define the threshold a_τ as:*

$$a_\tau = \frac{\lambda(c/\alpha) - c/\alpha}{\lambda(c/\alpha) + \lambda(-c/\alpha)}, \quad (2.21)$$

where $\lambda(\cdot)$ is defined in Lemma 3. Then (i) if $a_1 > a_\tau$ then $\tau > 0$ and (ii) if $a_1 < a_\tau$ then $\tau < 0$.

Corollary 1. (i) $0 < a_\tau < 1$ for any c . (ii) When $c = 0$, $a_\tau = 1/2$.

To put the above results into context, first consider the case when Hospital 1 and Hospital 2 have very similar desirability, i.e., when c is close to 0. Based on our observation, urban patients with smaller travel distance are more likely to select into urban hospitals, thus resulting in higher utility occurred at urban hospitals. Similarly, Proposition 1 part (i) indicates that when the more desirable hospital is sufficiently more urban than the less desirable hospital, the patients who choose the hospital will gain a higher utility than those who choose the other. However, when the more desirable hospital is not sufficiently more

urban, or when the more desirable hospital is more rural than, the less desirable hospital, it is possible that the patients who choose the hospital gain lower social welfare, despite choosing a more desirable hospital because it attracts more rural patients with farther travel distance. The result presents an inherent inequality between the patients in the rural and urban areas.

Proposition 2. a_τ is decreasing in c .

As Hospital 1 increasingly becomes more desirable, the aggregate utility gained by its patients becomes larger than the ones achieved by Hospital 2, and the inequality effect from geography decreases. The result indicates that investing in a rural hospital to increase the quality can overcome the inequalities in social welfare achieved by rural versus urban patients. Based on these observations, we next examine the properties of social welfare and optimal investment decision when there are two open hospitals.

Properties of $\mathbb{E}[\bar{\Pi}_4]$ and Optimal Allocation. In the absence of geography differences (e.g., when the patients' distances to Hospital 1 and Hospital 2 are perfectly correlated, or the two hospitals have similar geography resulting in $a_1 = 1/2$), the optimal investment is $x = B$: it is always optimal to invest all the fundings into the hospital that is more desirable at baseline. Next, we examine the case where hospitals have different geographies (rural versus urban). We present the following results regarding the quantity $\mathbb{E}[\bar{\Pi}_4]$, i.e., the social welfare when both hospitals are open. Recall that x denotes the investment amount Hospital 1 (the more desirable hospital at the baseline) receives and x_{thr} is the investment threshold such that investing above it in Hospital 2 will make Hospital 2 more desirable than Hospital 1.

Proposition 3. $\frac{\partial^2 \mathbb{E}[\bar{\Pi}_4]}{\partial x^2} > 0$, i.e., the function $\mathbb{E}[\bar{\Pi}_4]$ is convex for $0 \leq x \leq B$.

Lemma 3. For $0 < x_{thr} < B$, either $\mathbb{E}[\bar{\Pi}_4(x_{thr})] < \mathbb{E}[\bar{\Pi}_4(B)]$ or $\mathbb{E}[\bar{\Pi}_4(x_{thr})] < \mathbb{E}[\bar{\Pi}_4(0)]$.

Theorem 1. When $c_0 > 0$ and the two hospitals are competitive, the optimal investment level in

Hospital 1 denoted by x^* satisfies

$$\mathbb{E}[\bar{\Pi}_4(x^*)] = \max(\mathbb{E}[\bar{\Pi}_4(0)], \mathbb{E}[\bar{\Pi}_4(B)]). \quad (2.22)$$

The above results indicate that the optimal solution is either at $x = 0$ or $x = B$. The implication is that allocating all the funds in one hospital generates greater social welfare than splitting the funds between the two when both hospitals are decided to be kept open. In the context of CAH payment debate, this indicates that instead of reducing the funds to all existing CAH-eligible hospitals, reducing the total number of hospitals receiving the fund by imposing stricter eligibility (e.g., distance requirement). We further examine this idea in our simulation analyses where we make use of real-world data and include information on 1,515 hospitals in the U.S. Our results indicate that this insight gained by analyzing the allocation of funds between two hospitals holds more broadly.

While the above results indicate that it is optimal to invest all the budget in one hospital, they do not answer the following question: which hospital should receive all the budget? Our next result provides the answer.

Theorem 2. Define a threshold a_{inv} on a_1 (relative rurality of Hospitals 1 and 2) as

$$a_{inv} = \frac{w_{inv} + y_{inv}}{y_{inv} + z_{inv}}, \quad (2.23)$$

where the quantities w_{inv} , y_{inv} , and z_{inv} are:

$$w_{inv} = p(B)c(B) - p(0)c(0) - kB \quad (2.24)$$

$$y_{inv} = (1 - p(B))\alpha\lambda(c(B)/\alpha) - (1 - p(0))\alpha\lambda(c(0)/\alpha) \quad (2.25)$$

$$z_{inv} = -\alpha p(B)\lambda(-c(B)/\alpha) + \alpha p(0)\lambda(-c(0)/\alpha) \quad (2.26)$$

When the two hospitals are competitive, the investment decision between Hospital 1 and 2 depends

on a_{inv} and a_1 such that we have:

$$x^* = \begin{cases} B & \text{if } y_{inv} + z_{inv} > 0, a_1 \leq a_{inv} \\ 0 & \text{if } y_{inv} + z_{inv} > 0, a_1 \geq a_{inv} \\ B & \text{if } y_{inv} + z_{inv} < 0, a_1 \geq a_{inv} \\ 0 & \text{if } y_{inv} + z_{inv} < 0, a_1 \leq a_{inv} \end{cases} \quad (2.27)$$

To provide intuition for the decision rule, we can interpret w_{inv} as the net utility gain by moving the marginal patients from less to more desirable hospitals as a result of the quality investment. Furthermore, we can think of z_{inv} as the net disutility from the changes in distance as a result of moving the marginal patients via quality investment, and y_{inv} as the net disutility from the patients moving in the opposite direction. When w_{inv} is substantially greater than y_{inv} , the threshold on a_1 becomes high, i.e., it becomes optimal to invest in the hospital that is more desirable at baseline. However, when w_{inv} is substantially smaller than y_{inv} , the threshold becomes low, i.e., it becomes optimal to invest in the hospital that is less desirable at baseline. When the threshold is in the middle, the decision rule favors investing in the rural hospital.

We also note that $y_{inv} + z_{inv}$, i.e., the net disutility of moving patients from one to another hospital, explains why the allocation rule favors rural hospitals. First, the quantity represents the cost of bypassing a hospital, i.e., choosing the farther hospital for better quality. The cost of bypassing behavior from rural patients is higher than those of urban patients, and as a result, $y_{inv} + z_{inv}$ is more likely to be negative. In particular, when the more desirable hospital is much more urban compared to the other, bypassing cost will occur such that it outweighs the cost of additional quality gain. This is why, in certain cases, it is optimal to invest in the less desirable, rural hospital.

Corollary 2. Let $\bar{q} = q^1 - q^2$. $\frac{\partial a_{inv}}{\partial \bar{q}} > 0$ if $y_{inv} + z_{inv} > 0$ and $\frac{\partial a_{inv}}{\partial \bar{q}} < 0$ if $y_{inv} + z_{inv} < 0$.

Finally, the corollary above suggests that as the baseline desirability between the two hospitals becomes large, the threshold for investing in Hospital 1 (the more desirable

hospital at the baseline) increases, i.e., the geography effect diminishes and it becomes more likely that investing in the more desirable hospital is the optimal decision.

In summary, our results suggest that when the two hospitals are within a competitive range, it is not always such that investing in the hospital that is more desirable at the baseline is a better strategy. Instead, the decision is based on the (a) relative desirability, and (b) geography (relative rurality). Overall, the decision rule favors the rural hospital under the same condition. For example, when a better hospital is rural, optimal investment is to invest in a better/rural hospital. When the better hospital is urban, however, optimal investment is not always the urban hospital. For example, when the better hospital is significantly more urban than the other hospital, it can be optimal to still invest in the worse, rural hospital. If the baseline quality difference between the two hospitals is small, e.g., urban hospital slightly better, investing in the worse, rural hospital becomes more likely to be optimal. However, when the baseline quality difference is large, e.g., when the urban hospital is substantially better, investing in the better, urban hospital becomes more likely to be optimal.

These results provide two important policy implications. First, both DSH and CAH payment policies should incorporate the baseline desirability, in particular, the quality of hospitals into the allocation decision to improve efficiency. Second, even if the rural hospitals (CAH payment recipients) have a slightly lower quality than urban hospitals (DSH payment recipients), it can be more efficient to invest in rural hospitals. This is contrary to the existing belief that investing in rural hospitals mitigates inequalities at the expense of reducing efficiency. We further examine these insights in our simulation analyses.

2.4.2 Hospital Closures

In the previous section, we examined how investment decisions affect social welfare under the assumption that both hospitals are open. A more extreme form of intervention is to let one of the hospitals close, either through regulation or not providing support (e.g., bailout money) when a hospital faces closure. There have been 135 closures in the U.S. between 2015

and 2019, and there is a general increasing trend (Medicare Payment Advisory Commission (US), 2017; Carroll, C., 2019; Song and Saghafian, 2019a). Closing low-quality hospitals, for instance, might be one way to direct patients to higher-quality hospitals. This will improve the overall quality of care patients receive but will lower the access convince for them (on average). To gain insights into hospitals that should be closed, we next examine the case where the decision maker faces a choice between keeping two hospitals open versus keeping only one of them. Recall that the policymaker's ability to change the operating status of each hospital j is reflected in our model via the decision variable y^j . Furthermore, note that in a situation where only one hospital can be kept open, the closure decision depends only on the baseline desirability, but not on the geography. For example, it is always better to keep the hospital that is more desirable at the baseline open and close the other hospital.

To gain some insights, we first define $\bar{\Delta}$ as the differences in expected utilities between only keeping the less desirable hospital open and keeping both hospitals open, and $\underline{\Delta}$ as the difference in expected utilities between only keeping the more desirable hospital open and keeping both hospitals open. That is, we let

$$\underline{\Delta} = \mathbb{E}[\bar{\Pi}_4(x^*)] - \max(\mathbb{E}[\bar{\Pi}_1(B)], \mathbb{E}[\bar{\Pi}_1(0)]), \quad (2.28)$$

and

$$\bar{\Delta} = \mathbb{E}[\bar{\Pi}_4(x^*)] - \min(\mathbb{E}[\bar{\Pi}_1(B)], \mathbb{E}[\bar{\Pi}_1(0)]). \quad (2.29)$$

Theorem 3. $\underline{\Delta} > 0$ and $\bar{\Delta} > 0$.

Theorem 3 indicates that keeping both hospitals open always generates greater social welfare than closing one. This is consistent with the existing studies showing that having more choice (e.g., two instead of one hospital) generates greater societal gains (Gowrisankaran *et al.*, 2018; Capps *et al.*, 2010; McNamara, 1999). Thus, we quantify the additional utility by keeping both hospitals instead of one.

Proposition 4. $\lim_{\bar{q} \rightarrow 0} \underline{\Delta} = 0$,

where $\bar{q}$ is defined in Corollary 2, and

Proposition 5. $\frac{\partial \Delta}{\partial \bar{q}} < 0$ and $\frac{\partial \bar{\Delta}}{\partial \bar{q}} > 0$.

The results above indicate that as the baseline quality advantage of Hospital 1 relative to Hospital 2 diminishes, the additional utility gain of keeping both hospitals becomes small. In addition, when Hospital 1 is substantially better than Hospital 2, keeping only Hospital 1 provides almost the same aggregate utility as keeping both.

While the results presented above indicate that it is always better to keep both hospitals open, there is an issue: in practice, it is unlikely that policymakers can allocate funds in an optimal way. Thus, it is important to study what happens when funds are allocated in an imperfect way (suboptimal). Thus, we next examine the situation when closing one hospital is better than keeping both but misallocating resources. Define $\tilde{\Delta}$ as the difference between the utility generated by investing in the wrong hospital (e.g., investing in Hospital 2 when the investment to Hospital 1 is optimal) and closing the less desirable hospital.

$$\tilde{\Delta} = \min(\mathbb{E}[\tilde{\Pi}_4(B)], \mathbb{E}[\tilde{\Pi}_4(0)]) - \max(\mathbb{E}[\tilde{\Pi}_1(B)], \mathbb{E}[\tilde{\Pi}_1(0)]) \quad (2.30)$$

Proposition 6. *There exists c_0 such that $\tilde{\Delta} < 0$.*

Proposition 6 indicates that there exists a situation where closing the less desirable hospital (and investing in the more desirable hospital) generates greater social welfare than allocating the fund suboptimally and keeping both hospitals. In general, the closure of a hospital has generated greater concerns to the society than the misallocation of existing funds. Our results indicate, however, that under the circumstances that the current allocation of funds is suboptimal and the improvement through reallocation of funds is not feasible, it can be welfare improving to not pursue the policies aimed at preventing hospital closures (e.g., providing direct bailout funds to hospitals on the verge of closures). In the following sections, we show that the current policy and the suggested legislative changes on the level and allocation of fundings for DSH and CAH payment programs are suboptimal. We also show that the allocations that incorporate both quality and geography into account are

welfare improving. However, if these changes are not feasible, it can result in a situation where the closure of less desirable hospitals improves welfare over the status quo.

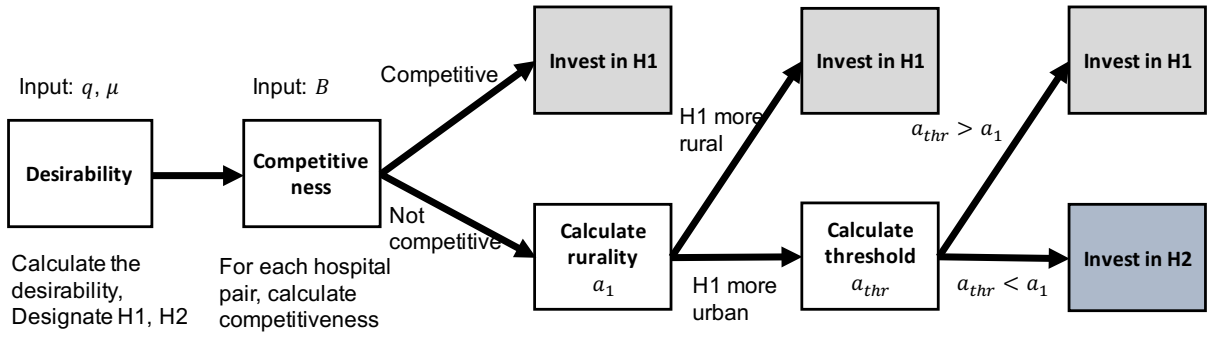
In summary, our results indicate that although hospital closures induced by DSH or CAH payment reduction can result in reduced patient utility, it is possible to mitigate the impact through identifying the hospital whose closures are least likely detrimental to the patients' utility. Given that hospital closure will result in cost savings through reduced DSH or CAH payments, achieving similar levels of social aggregate utility while achieving cost saving can be welfare improving. In addition, in situations where improving welfare through reallocation of funds is not feasible, the closure of less desirable hospitals can improve welfare. In the following section, we examine this idea further through simulation analysis calibrated with our data (see Table 1 for the sources of our data).

2.4.3 Summary and heuristic

Below is the summary of our results in response to the research questions we raised in the beginning.

1. Patient behavior: Patients have a geographical preference such that urban patients tend to choose the urban hospital, and rural patients tend to choose rural hospitals. Because of such preference, patients who choose rural hospitals tend to achieve lower utility than those who choose the urban hospital.
2. Investment: Investing all the funds in one hospital generates greater social welfare than splitting the funds between the two. When the two hospitals are out of a competitive range, it is optimal to invest in the more desirable hospital. When the two hospitals are within a competitive range, the decision is based on the (a) relative desirability and (b) geography at baseline. When the more desirable hospital is rural, optimal investment is to invest in it. When the more desirable hospital is urban, optimal investment based on the threshold.
3. Closure: Keeping both hospitals generates greater social welfare than closing one.

Figure 2.3: *heuristic for Investment Decision between Two Hospitals*



However, when one hospital is substantially more desirable, the additional gain of keeping both becomes very small. It is possible that closing the less desirable hospital (and investing in the more desirable hospital) generates greater social welfare than allocating the fund suboptimally and keeping both hospitals.

4. Policy evaluation and recommendations for DSH/CAH: In the following sections, we examine the impact of DSH or CAH payment reform through simulation analyses that use the real data. Then we apply our heuristic described in Figure 2.3 and other insights obtained from analysis into the simulation to identify ways to improve current policies under debate.

2.5 Simulation Experiments

We conduct simulation experiments to complement our analytical findings. We calibrated our model by making use of nationally representative patient and hospital-level data (Table 2.1), which allow us to get insights by observing the impact of some potential policy interventions. We next describe our data, parameter estimation methods, and simulation assumptions.

2.5.1 Data and Study Population

We utilized the traditional fee-for-service (FFS) Medicare inpatient claims data for the year 2014, the first year the DSH payment has been reduced and the payment rule has been revised, to identify each patient's hospital choices, and Medicare Beneficiary Summary Files (BSF) to identify each beneficiary's demographic, clinical, and geographic information. Information on hospital quality and the patient experience was obtained from the Hospital Compare data,¹¹ the publicly available quality ratings provided by the CMS and the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS). Our study population is the Medicare beneficiaries who paid at least one inpatient visit during 2014 and are enrolled in the FFS Medicare throughout the year.¹² We further excluded the patients and hospitals in the market (defined as the Hospital Referral Region, or HRR (Wennberg, 1996)) that has more than 16 hospitals (28.8% of total markets), because the trade-off between distance and quality becomes less clear when multiple hospitals are concentrated in an area. The hospitals were limited to the short-term acute care hospitals or the CAHs. Hospitals receiving DSH payments were identified as those who received any DSH payment during the study period, according to the Medicare Cost Report. Overall, our final data includes 361,410 patients who visited 1,515 hospitals in 2014.

2.5.2 Variable Definitions

In general, FFS Medicare patients can choose to go to any Medicare-participating hospitals, which are essentially almost all U.S. hospitals. However, in reality, our data shows that the majority (84.6%) of inpatient visits are made by the patients who reside in the same market (HRR) as the hospital, and the majority (81.2%) of the within-market hospital choice is to either the nearest or the second-nearest hospital when there are three or fewer hospitals in a market. Thus, to estimate the parameters of our model, we examined the patient's trade-off

¹¹Hospital Compare data draw detailed information on hospital quality from multiple sources, including hospital submitted electronic health records, surveys, and Medicare claims data.

¹²We restricted the population to those who are aged 66 or older and are entitled due to age (as opposed to disability).

between the nearest and the second nearest hospital. For our simulation experiments, we did not restrict the patient's choice set, allowing them to choose among all hospitals in the market.

Hospital Quality. For hospital quality measures, we used the 30-day risk-adjusted acute myocardial infarction (AMI) survival (%) defined as $100 \times (1 - \text{30-day risk-adjusted AMI mortality rate})$, which is a widely used proxy of hospitals' overall quality (Bradley *et al.*, 2006). In addition, we examined overall patient experience ratings, defined as the percent of patients who rated good for their inpatient care experience. The mortality rate was obtained from Hospital Compare Data, widely used measure of hospital care quality, and overall patient experience ratings were obtained from HCAHPS data (Emmert and Schlesinger, 2017; Gutacker *et al.*, 2016). Because patients with higher risk or specific conditions are likely to incorporate the diseases-specific quality measures, we also used disease-specific 30-day readmissions rates and disease-specific 30-day risk-adjusted survival rates among patients with specific conditions in our robustness checks (see Section 6).

Distance. To estimate the distribution of distance from patients to hospitals, we need observations of each patient's travel distance from their departure location (e.g., home or work) to the destination (hospital). However, we only observe the 5-digit zip codes of each beneficiary's home address and 5-digit zip codes of the hospital in our data. Thus, we estimated the distance parameters as follows. For each market, the mean distance from each hospital to all patients is calculated using the zip code centroid based distance. We also estimated the value of $\Sigma^{(j,k)}$ defined in (13) for each hospital pair in the market using the zip code centroid based distance, from which we obtained the value of a_1 to the pair. Table 2.2 provides a summary of our quality and distance measures. The DSH payment recipients tend to have similar 30-day AMI risk-adjusted survival and worse patient experience than the CAHs and have a smaller mean distance than the CAHs. The majority of DSH payment recipients are urban, whereas the majority of CAHs are rural. Consistent with our definition, DSH payment recipients have $a_1 > 0.5$ and CAH payment recipients $a_1 < 0.5$, where the mean a_1 of all hospitals in the sample is 0.5.

Table 2.2: *Summary of Hospital Characteristics*

	All	DSH	CAH
Quality, risk-adjusted AMI survival (%)	85.82	85.80	85.41
Quality, patient experience (% good)	71.07	68.65	74.57
Average distance (miles)	33.65	26.14	59.43
a_1	0.50	0.54	0.41
Urban (%)	55.96	70.2	19.9
Rural (%)	44.04	29.8	80.1

2.5.3 Parameter Estimation

There are two key parameters in our model: (a) the distance quality trade-off constant (α), and (b) the effectiveness of investment on quality (k). We discuss our estimation approaches for these two parameters below.

Quality-Distance Trade-Off (α). We directly estimated the quality-distance trade-off parameter α from our data using the maximum log likelihood (MLE) approach. Formally, in our data, we observe Z_i^* , a binary dependent variable indicating each patient's hospital choice. For each inpatient visit, define $Z_i^* = 1$ if $U_i^1 > U_i^2$ (i.e., the patient chooses Hospital 1) and $Z_i^* = 0$ if $U_i^1 < U_i^2$ (i.e., chooses Hospital 2). We have the model parameters $(\alpha, \mu, \Sigma) \in \Lambda$ and $Z_i \sim f(z_i|\Lambda)$ and $Z_i, i = 1, \dots, n$, are i.i.d. The parameters μ and Σ are estimated from our data, and thus we only need to estimate α . For estimating α , we assume that the investment level is the amount distributed in the previous year. For simulation, we defined x as the differences in the funding that the hospitals receive compared to the previous year.

Proposition 7. *For patient i , let $c_i = (q_i^1 - q_i^2) - \alpha(\mu_i^1 - \mu_i^2)$. The log likelihood equation $l(\alpha|z)$, i.e., the probability of observing a given data is*

$$l(\alpha|z) = \sum_{i=1}^n [z_i \Phi(c_i/\alpha\sigma_i) + (1 - z_i)(1 - \Phi(c_i/\alpha\sigma_i))]. \quad (2.31)$$

The maximum likelihood estimate of α is given by

$$\hat{\alpha} = \operatorname{argmax}_{\alpha} \sum_{i=1}^n [z_i \Phi(c_i/\alpha\sigma_i) + (1 - z_i)(1 - \Phi(c_i/\alpha\sigma_i))]$$

Table 2.3: Parameters Used in Our Model

Parameter		Value	Source
q	Quality, risk-adjusted AMI survival	Varies by hospital	Hospital Compare
	Quality, patient experience	Varies by hospital	HCAHPS
α	Distance-quality trade-off (AMI survival)	3.28 % survival/log mile	Our data using MLE
	Distance-quality trade-off (patient experience)	13.89 % good/log mile	Our data using MLE
k	Investment effectiveness on quality (AMI survival)	1.17% per \$100 per capita	Baicker, K. and Staiger, D. O. (2005)
B	Total budget for DSH and CAH program (2014)	5.1 billion USD	Medicare Cost Report, CMS

The MLE $\hat{\alpha}$ is such that $\frac{\partial l(\alpha|z)}{\partial \alpha}|_{\alpha=\hat{\alpha}} = 0$. Given the functional form, we cannot analytically solve for the $\hat{\alpha}$, so we numerically solved using the quasi-Newton method (Broyden, 1970). Because of large N , our parameter estimate $\hat{\alpha}$ approximates the true value α with a tight bound.

Table 2.3 shows the estimated α for various outcomes. Overall, our results indicate that patients are willing to travel one additional mile on average to receive treatment from a hospital that has about 0.2% higher risk-adjusted survival rate and 0.6% higher patient experience rating. The estimates are similar to those reported in other studies (Romley and Goldman, 2011; Saghafian, S. and Hopp, W. J., 2020; Chandra and Staiger, 2007), although some differences in magnitude might be due to the differences in the study population and the definitions of quality in our setting. To further validate our estimate of α , we also performed the following. We first estimated α using a random 80% subset of the population as training data. We then used our model along with the estimated value of α to predict the hospital choice of the remaining 20% out-of-sample test data population. This revealed that our MLE estimation approach well approximates the observed patient choice.

Investment Effectiveness (k). Unlike α , we cannot directly estimate k from our data, so we adapted the estimates from existing literature (Baicker, K. and Staiger, D. O., 2005; Jha *et al.*, 2010; Li *et al.*, 2007) and conducted various sensitivity analyses to test our assumptions. Several studies provide a direct link between DSH payment and an improvement in hospitals' outcomes. Baicker, K. and Staiger, D. O. (2005) shows that the DSH program improves health outcomes such as neonatal health and for AMI patients, and Jha *et al.* (2010) shows that

hospitals with a high DSH index were associated with greater improvements in performance for AMI and pneumonia. Li *et al.* (2007) provides evidence that when a hospital converted into a CAH, there was an improvement in patient safety indicators. We made use of estimates from Baicker, K. and Staiger, D. O. (2005), which provides evidence that \$100 per patient increase in DSH payment results in a 1.17% reduction in the risk-adjusted percent of Medicare patients who died within 90 days of having a heart attack. To translate their findings into our setting, we used $k = 1.17\%$ per hospital per 100 DSH dollars, which is equivalent to 0.96% increase in risk-adjusted survival percentage per 1 million DSH spending. Similarly, quality improvement was one of the main goals of the CAH program, and evidence shows that hospitals have used many of the CAH funding to quality improvement and quality assurance activities (Casey *et al.*, 2015). Thus, in our main analysis, we assumed that CAH would have an equivalent effectiveness rate as the DSH investment.

Budget (B). We identified the total amount of Medicare DSH or CAH payment each hospital received in 2014 in order to calculate the allocation scenarios in the policy simulation. Medicare DSH payment was obtained directly from the Cost Report. The CAH payment was calculated, based on previous studies, as the difference between the revenue generated under cost-based reimbursement in the study period and the simulated revenue that would have been generated under DRG based reimbursement (Carroll, C., 2019; Gowrisankaran *et al.*, 2018). The total budget for our simulation is 5.1 billion USD, a sum of 3.4 billion USD for the DSH program in 2014, and 1.7 billion USD for the CAH program in 2014. Note that, however, our simulation does not consider the possibility of moving funds from one program to the other.

2.6 Policy Simulation

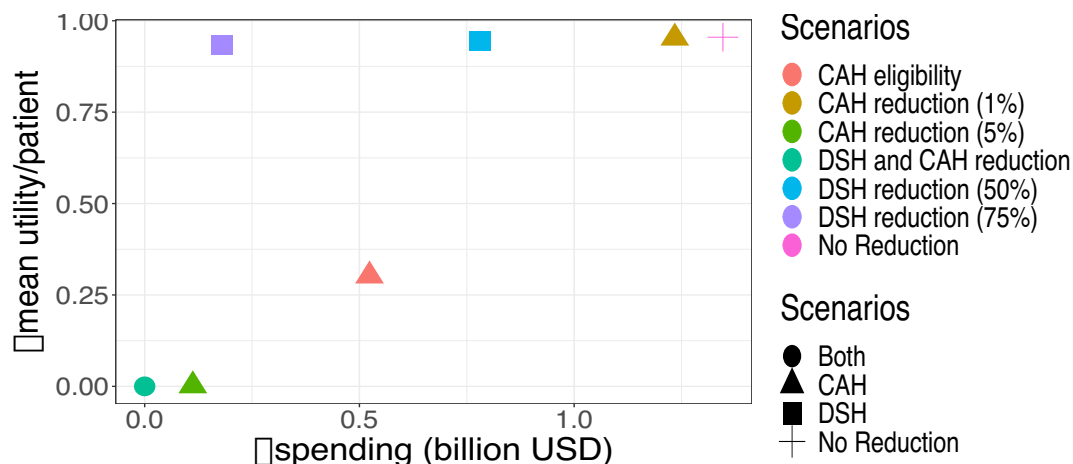
2.6.1 Policy Scenarios

We simulate allocating the budget for DSH or CAH payment among the 1,515 hospitals in our sample according to the following policy scenarios.

- **ACA DSH Reduction (Status Quo):** We examined the current status of funding, where DSH fundings are reduced according to the ACA to 25% of the amount they previously have received.
- **No Reduction:** The total and per hospital spending for DSH does not change as intended by the ACA. Thus, each hospital's DSH payments are calculated as the amount the hospitals received in 2013, prior to the DSH reduction. Under no reduction scenario, the CAH program remains the same as the observation year (2014). Because the reduction has originally scheduled as gradual changes between fiscal years 2014 and 2019, we also examined the intermediate DSH reduction where hospitals receive 50% of the amount they previously have received.
- **CAH Reduction:** We examined the case where the DSH payments are not reduced as planned in the ACA, but CAH payment recipients receive reduced payment from the current 101% of the allowable cost to 100%, as has been proposed. We additionally examine the case where the reduction is by 5%, i.e., 96% of the allowable cost.
- **ACA DSH Reduction and CAH Reduction:** We examined the scenario where both fundings are reduced, as specified above.
- **CAH Eligibility:** The CAH payment recipients who were exempt from the distance requirement lose their CAH status. We identified these hospitals as those acquired the CAH status between 1997-2005 and are within 35 miles from the nearest hospital (total 840 CAHs). Based on a study that the average increase in simulated revenues as a result of qualifying for CAH status is 14%, we calculated the revenue as a result of losing their CAH status (Carroll, C., 2019).

Figure 2.4 shows the changes in spending and the changes in expected mean utility associated with each scenario. The two CAH-related strategies—CAH reduction (5%) and CAH eligibility reduction—are strongly dominated by other strategies. The incremental cost-effectiveness ratio (ICER)—defined by the difference in cost between two scenarios, divided

Figure 2.4: Cost-Effectiveness Plane for Baseline Policy Scenarios



by the difference in their effect—for DSH reduction (75%) is 5.18 (utility per patient per 1 billion USD). Translating the number into a more realistic interpretation, this is equivalent to about 810 fewer AMI death for an average-sized hospital per year. The ICER for DSH reduction (50%) is 0.0184, and the ICER for no reduction is 0.0180 increase in utility per patient per 1 billion USD. The strategy of CAH reduction (1%) has a very similar ICER as the strategy of no reduction but is marginally weakly dominated. Therefore, only the strategies of reducing DSH payments are on the cost-effectiveness frontier.

Our analysis suggests that the original strategy of reducing both fundings is unlikely to be welfare-improving. Instead, the strategies of reducing DSH but not reducing the CAH payment are cost-effective. In particular, DSH reduction, as originally planned, is the most cost-effective. This is consistent with our analytic results in that allocating a greater amount of funds to rural hospitals can outperform the allocation to urban hospitals. Secondly, our result shows that between reducing CAH eligibility and reducing CAH funding (1%), reducing eligibility is the dominant strategy. This is also consistent with our observation from the analysis in Section 4 that allocating funds to a fewer number of hospitals is more efficient than reducing the level of funding.

2.6.2 Strategies to Redistribute Funds

Since our results in the previous sections indicated that the original strategy of reducing both fundings is unlikely to be welfare-improving, we next study how to improve the DSH or CAH programs by better redistribution the available funds.

Strategies. Next, we examined various strategies below that could improve the current DSH or CAH policies. The baseline comparison point is the planned DSH and/or CAH payment reduction originally intended, described in the earlier section. Within each policy change, we examine targeting hospitals as a mechanism of redistributing funds based on various criteria. For quality-based allocation, we re-allocated the budget proportionally to baseline quality, with a greater amount to high-quality hospitals. For size-based allocation, we allocated the budget proportionally to the total number of beds and separately based on annual patient volume in 2014. Finally, we examine equally allocating the funds to all hospitals participating in each program.

Our Heuristic. We also examine the heuristic-based allocation rules we proposed in Figure 2.3. Table 2.4 summarizes the distribution of parameters for each of the hospital pairs in our data. Based on the a_1 and a_{inv} for each hospital pair, we identified whether Hospital 1 or Hospital 2 is recommended for investment. Overall, the more desirable hospitals are recommended for investment 61.7% of the time, and their less desirable pairs 38.3% of the time. Overall, 980 out of total 1,515 hospitals were recommended for investment.

Figure 2.5 shows the changes in social welfare for each of the allocation strategies compared to the baseline comparison point. All allocations are budget neutral. Among characteristics based allocations, equal distribution of funds results in the worst outcome for all three programs, followed by bed-size based allocation. Size based distribution performed better, and quality based distribution resulted in the greatest improvement in utility. This can be due to that quality based allocation distributes greater resources to the hospitals that more patients are likely to choose. The results suggest that contrary to existing belief, allocating more resources to hospitals with the greater size is not an efficient policy. The fact that quality-targeted investment is the next best strategy indicates that the current payment

Table 2.4: Summary of Parameters Used in heuristic for Hospital Investment

	1st Qu.	Median	Mean	3rd Qu.		1st Qu.	Median	Mean	3rd Qu.
$p(B)$	0.890	0.950	0.928	0.990	w_{inv}	0.677	1.429	1.435	2.244
$p(0)$	0.220	0.340	0.370	0.500	y_{inv}	-0.687	-0.455	-0.461	-0.223
$c(B)$	3.236	4.013	4.130	4.899	z_{inv}	0.219	0.584	0.664	1.048
$c(0)$	-2.420	-1.644	-1.526	-0.757	a_1	0.462	0.560	0.557	0.652
$\lambda(c(B)/\alpha)$	1.396	1.650	1.698	1.948	a_{inv}	-0.897	1.144	-0.992	2.500
$\lambda(c(0)/\alpha)$	0.241	0.437	0.489	0.680					
$\lambda(-c(B)/\alpha)$	0.039	0.148	0.206	0.317					
$\lambda(-c(0)/\alpha)$	0.807	1.049	1.069	1.313					

Note. Summary of each parameter calculated for a pair of hospitals. All hospital pairs that are in the same market are included. Hospital 1 was designated as the one with greater desirability than Hospital 2.

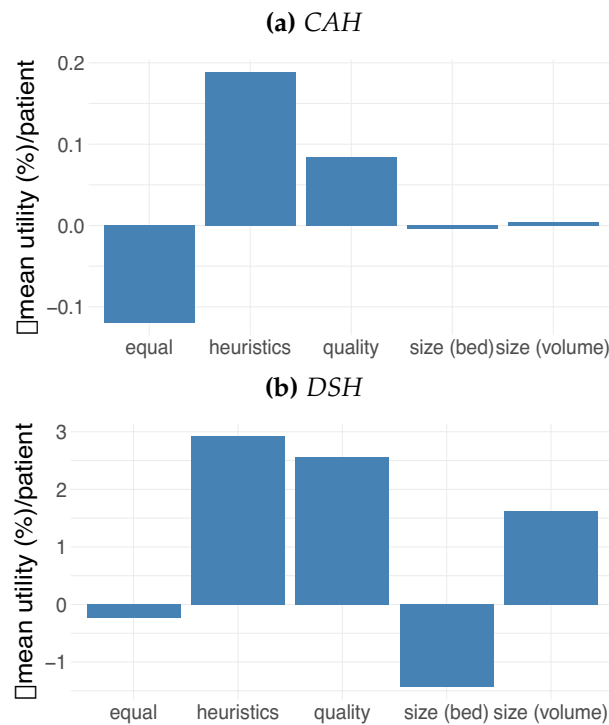
reform toward payment in hospitals based on value such as the value-based purchasing (VBP) programs can be more welfare improving than others. Between CAH and DSH, CAH programs are more likely to experience an improvement as a result of volume or quality based allocation than DSH programs. Finally, our heuristic outperforms other allocation strategies. Consistent with other strategies, allocation to the CAH program had more improvement in social welfare compared to the DSH program.

2.6.3 Hospital Closure Scenarios

Our analytic results suggest that the optimal result under keeping both hospitals generates greater social welfare than the optimal results under closing one hospital. However, in case one hospital is substantially more desirable, our results show that the additional welfare gain of keeping both hospitals becomes very small. Moreover, closing the less desirable hospital (and investing in the more desirable hospital) can generate greater social welfare than allocating the fund suboptimally and keeping both hospitals. We compare the current trend of closures as a result of policy changes with optimal closure scenarios based on our results.

To perform our analysis, we first identify the hospitals that are most likely to close under DSH payment reduction based on published studies (Cole *et al.*, 2014). These are the hospitals that are: (a) in high-DSH states and (b) with high reliance on DSH payments. We

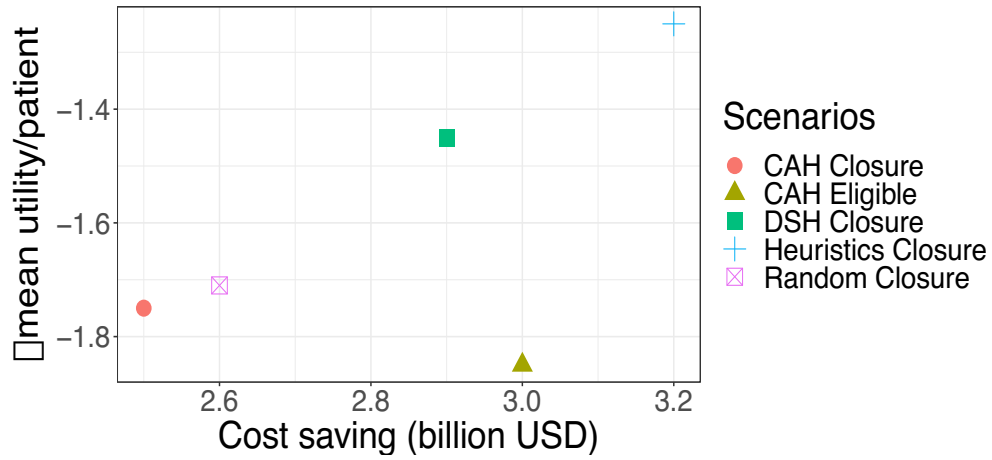
Figure 2.5: *Changes in Social Welfare (Mean per Patient Utility) under Various Investment Scenarios*



Note. Each bar indicates the size of percent changes in expected utility per patient. All strategies are budget neutral. Status Quos are the equal budget level and allocation as year 2014.

then measure the high reliance on DSH payments as a proportion of the Medicaid DSH payments out of total revenue from the Cost Report. Based on these criteria, we ranked the hospitals that are most likely to close due to the reduction in DSH payment. We identify hospitals likely to close as those most likely to be affected by the CAH eligibility changes, i.e., those acquired the CAH status between 1997-2005 and are within 35 miles from the nearest hospital. We also find the hospital whose closure will be the least detrimental on the expected utility based on our heuristic. Similar to the investment decision, we measure the differences in desirability for each hospital pair in the market. This allows us to rank hospitals based on their likelihood to achieve a similar level of utility with or without closure of the worse hospital. Given that approximately 20% of existing hospitals are at high risk of closure, we are able to identify 20% of hospitals in our sample (n=303) most likely to close

Figure 2.6: Cost-Effectiveness Plane for Hospital Closures (N=303)



as a result of CAH, DSH, or the heuristic. For each closure, we also estimate cost saving as the total DSH payment made in the year 2014 for DSH payment recipients, or 14% of the total payment made to CAH payment recipients in the same year.

Figure 2.6 shows the impact of each closure scenario. Although DSH policy-led closures are less detrimental to society than random closures, CAH policy change-led closures result in a greater reduction in utility. The heuristic-based determination of closure results in 0.26 greater utility per patient than the DSH-led closure. This is equivalent to about a 0.26% increase in risk-adjusted AMI survival rate for each patient who receives treatment. Overall, closure of hospitals recommended by the heuristic results in cost-saving because of a reduction in DSH and CAH payments. The closure as a result of a reduction in CAH eligibility shows a greater amount of cost-saving than the closure as a result of fund reduction, which is likely due to that these hospitals that lose eligibility have greater sizes and receive greater CAH payments (Neuhausen *et al.*, 2014). Our results suggest that identifying the type of hospitals that are least likely to be detrimental to society can substantially improve social welfare. Strategies that consider relative quality and geography among hospitals in the market can achieve better results than the current payment reduction plans.

2.7 Conclusions and Policy Recommendations

We studied the type of hospitals that should be supported by funds. Our framework allows targeting suitable hospitals by taking into account quality, distance, and geography under budget constraints. Contrary to existing belief, we find that investing all resources into a few specific hospitals yield greater social welfare than current programs that aim at supporting a wide range of hospitals but each with a lower amount of support. In addition, we observe that allocating resources to urban instead of rural hospitals does not necessarily result in better outcomes. In particular, allocating greater resources to more rural hospitals, even if their quality is lower, can improve social welfare by reducing the bypassing cost (see Theorem 2). Furthermore, our results indicate that quality based allocation rules can increase social welfare compared to the currently suggested legislative fund reduction policies. Finally, we find that although hospital closures generally reduce social welfare, some closures can result in nearly optimal results, if closure decisions are based on both quality and geography.

Based on our findings, we make the following policy recommendations.

1. Revising the eligibility criteria for hospital payment programs.
 - Current discussion on the eligibility criteria for CAH payment focuses on the hospitals that have obtained the status through exemptions for distance requirements. These hospitals are less likely to be in rural areas and have higher quality than other CAH payment recipients that do meet the distance requirement (Neuhausen *et al.*, 2014). Based on our findings, we recommend revising the CAH eligibility to ensure that only the hospitals that originally met the intention of CAH receive payments (instead of reducing the CAH payment to all hospitals).
 - The original rationale of DSH payment reduction was that hospitals would experience substantially less proportion of uncompensated care as a result of Medicaid expansion, and therefore, will need less support from the DSH program. However, some states ended up not expanding Medicaid, which resulted in the hospitals

in the non-expansion state experiencing the same level of uncompensated care. Based on our analysis, we find that reducing DSH payment by supporting the hospitals in the non-expansion state, instead of reducing DSH payments to all recipients, will result in greater societal gains.

2. Implementing quality-based investment strategies can improve social welfare. However, quality-based payment is not a gold-standard.

- We find that, in general, investing in higher quality hospitals is more efficient than investment decisions based on other criteria such as equity or size. Currently, eligibility and the size of payments are mostly determined by the proportion of low-income patients, but not determined by quality. Our result supports the current ongoing effort to move toward a value-based payment system. However, we observe that only accounting for quality will result in suboptimal outcomes, and geography should also be incorporated.
- Our results also recommend prioritizing investment in hospitals in more rural areas, even if their quality is lower than the urban hospitals. This means that in addition to the current reimbursement method of the CAH program are based on the cost, incorporating the geography and supporting hospitals in more rural areas can also improve social welfare.

There are limitations from the key assumptions we make in our study. First, we do not incorporate the variations in waiting time in our patients' utility function, but such effect may affect patient choice as well as hospital quality (see, e.g., Dong *et al.* (2018)). Second, we assume that patients have perfect information. In reality, patients' knowledge of hospital quality may have been affected by their own experience, the reputation that is not captured through risk-adjusted survival (see, e.g., Saghafian, S. and Hopp, W. J. (2020)). Third, we do not incorporate additional heterogeneity in investment effectiveness, such as the volume-outcome effect (see, e.g., Chandra and Staiger (2007); Saghafian, S. and Hopp, W. J. (2020) that examine the scenarios when the providers' quality depends on his/her

patient volume). Fourth, we do not assume that hospitals are strategic. Evidence shows that hospitals can change their operations in response to the market structure, and such spillover effects can have implications for the remaining hospitals (Song and Saghafian, 2019a). Future research is needed to extend our results.

Chapter 3

The Impact of Vertical Integration on Physician Behavior and Healthcare Delivery: Evidence from Gastroenterology Practices¹

3.1 Introduction

The U.S. health care system is undergoing a period of substantial consolidation between physicians and hospitals, with hospitals purchasing many physician practices or directly employing physicians (Kocher and Sahni, 2011; Burns *et al.*, 2014; Alpert *et al.*, 2017; Scott *et al.*, 2017; Baker *et al.*, 2018). The number of physicians who have “vertically integrated” (e.g., consolidated) with hospitals has doubled in the past decade, and the trend is expected to continue (Kocher and Sahni, 2011; Neprash *et al.*, 2015; Scott *et al.*, 2017; Nikpay *et al.*, 2018). In theory, there are potential benefits of vertical integration such as greater efficiency through achieving economies of scale and better quality through achieving care coordination

¹Co-authored with Soroush Sagharian, Joseph Newhouse, Mary Beth Landrum, and John Hsu

and information sharing (Kocher and Sahni, 2011; Burns *et al.*, 2014; Carlin *et al.*, 2015; Baker *et al.*, 2018; Baicker, K. and Levy, H., 2013; Lammers, 2013). However, the trend of vertical integration has also generated concerns among policymakers because of its purported anticompetitive effect, which could increase price and spending and lower quality (Neprash *et al.*, 2015; Capps, C. *et al.*, 2018; Baker, L. C. *et al.*, 2014).

A growing number of studies have examined the impact of integration on care delivery, but the evidence quantifying these effects on quality is mixed (Post *et al.*, 2018; Scott *et al.*, 2017; Carlin *et al.*, 2015; Lammers, 2013). For example, the majority of peer-reviewed studies present none or ambiguous findings, and the studies that employ more rigorous methods either examine limited geographic areas (Carlin *et al.*, 2015) or focused on the integration between post-acute care centers and hospitals (Konetzka *et al.*, 2018). Although the variations in the use of data sets or methods among existing studies could have contributed to the lack of consensus on the findings, these studies also have limitations in that they focus on measuring the direct relationship between the integration and quality outcomes. Yet, a key question remains unanswered is whether vertical integration promotes a fundamental change in the way physicians operate and how such changes affect quality, efficiency, and spending. Understanding the mechanism through such changes in outcomes occur can help better interpret the implications of vertical integration on the healthcare delivery system.

In this study, we generate insights into this question by examining how hospital-physician integration affects various dimensions of physician behavior. We focus on the hospitals' purchases of physician outpatient practices, where despite a clear change in the financial ownership of the organization, other factors that may affect the care delivery (e.g., physicians, patients, and geographic location) could remain relatively fixed. This enables to tease out the behavioral responses of the physicians from other contemporaneous changes. We find that vertical integration is associated with a reduction in some key dimensions of quality (e.g., 30-day complication rates) and efficiency, and the effect is explained by the changes in physicians' behavior in reducing some recommended care processes (e.g., anesthesia with deep sedation). We further provide evidence that such changes are due

to the resource constraints of integrated providers that limit the provision of value-adding care steps after integration. Finally, we show that in addition to the negative impact on the quality of care and efficiency, integration is associated with a significant increase in per-patient spending. Overall, our results suggest that the changes in financial ownership, without appropriate changes in the incentive structure to motivate the physicians' care processes in a positive direction, can have negative impacts on the healthcare delivery system through both an increase in adverse patient outcomes and spending.

3.1.1 Policy Context and Settings

Our study setting is the fee-for-service (FFS) Medicare, national health insurance that covers the majority of the elderly U.S. population.² Under the FFS payment model, services are not bundled, and providers are paid for each service at the administratively set prices. Importantly, Medicare reimburses the same procedures that occurred in different settings differently: in general, procedures occurred at hospital outpatient departments (HOPDs) are paid higher rates than those at physician offices or ambulatory surgery centers (ASCs) (e.g., Medicare reimburses \$1,090 for the colonoscopies that occurred at HOPDs, whereas only \$413 for those at physician offices.) Yet, FFS Medicare patients can receive many of the outpatient procedures at different settings, and there is limited evidence that patients select themselves into different settings to justify such price differentials. Thus, these policies have been criticized for accelerating the trend of integration and contributing to the growth in Medicare outpatient spending by motivating hospitals to acquire physician-owned practice and convert the practice (usually in the same location) into an HOPD (Office of Inspector General, 2014; Medicare Payment Advisory Commission (US), 2017; Forlines, 2017; United States Government Accountability Office, 2015). Although there has been a legislative attempt to eliminate the fee differentials for non-grandfathered practices in January 2017 (Bipartisan Budget Act of 2015 (Section 603)), the more recent legislation has revised the

²In particular, the FFS provider payment policy relevant to our study question is Medicare's payment for outpatient services (outpatient prospective payment system, or OPPS).

Table 3.1: *Overview of Our Dataset*

Name	Unit	Years
FFS Medicare Claims (Inpatient, Outpatient, Office)	Patient visit	2008-2015
Medicare Beneficiary Summary Files	Patient	2008-2015
CMS Physician and Other Supplier Data	Physician	2012-2015
CMS Hospital Compare	Hospital	2015
Medicare Provider of Service	Hospital	2008-2015
Area Resource Files	County	2012
CMS State/County/Plan Enrollment Data	County	2008-2015

site-neutral payment policy to expand exemptions (21st Century Cures Act). Clear policy recommendations have been difficult, mainly because of the lack of evidence on the impact of vertical integration on other dimensions of care delivery beyond expenditure (e.g., quality or efficiency).

3.1.2 Challenges and Framework

There are several challenges in providing evidence on vertical integration at the national level. First, there is limited national-level data that can be used to provide generalizable evidence. To address this challenge, we (a) make use of a data set that covers 20% of all fee-for-service (FFS) Medicare patients between 2008-2015, which includes 3.6 million observations of patient visits provided by 5,488 physicians, and (b) combine it with multiple other data sources that allow us to examine a wide range of measures, including quality, efficiency, expenditures, and patient experience (Table 3.1).

Second, because the majority of acquisitions occur at a small scale, vertical integration is not easily identifiable from the survey data. We take advantage of the Medicare payment rules and providers' billing patterns to infer the financial relationship between physicians and hospitals. In this way, we are able to identify the integration status of individual physicians during each year in our study period (2008–2015).

Third, measuring the changes in outcomes before and after integration in a meaningful way is challenging. To address this challenge, we focus on a homogenous clinical area: Gastroenterology (GI). GI is one of the specialty areas that has experienced a rapid in-

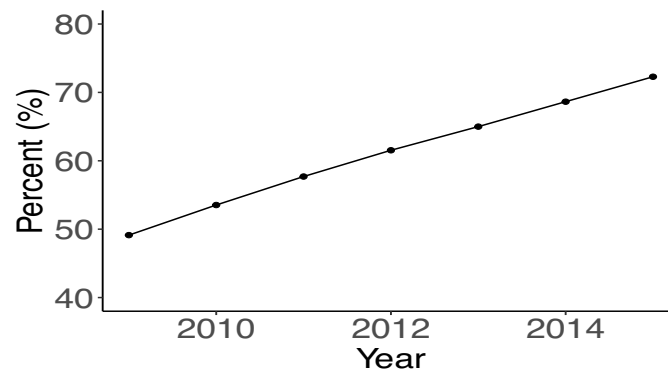


Figure 3.1: *Trend of GI Physician Integration (2009-2015)*

Note. Source: our analysis of Medicare FFS claims.

crease in vertical integration. Figure 3.1 shows the trend of vertical integration among gastroenterologists, where integration has increased more than 20% between 2009 and 2015. Focusing on GI has clear advantages because colonoscopy—a primary type of endoscopy for colorectal cancer (CRC) screening and diagnosis—has a set of well-validated process quality measures that are sensitive to the physicians’ skills and are linked to important long-term patient outcomes such as interval cancer.

Fourth, there are various identification challenges that one needs to overcome to establish a causal effect of vertical integration. For example, physicians’ decision to integrate is not exogenous; the physicians who are more likely to benefit from integration will do so, thus leaving the effect subject to selection bias. Specifically, the decision can depend on physicians’ ability, strategy, and technology, which in turn can be correlated with their care delivery patterns. Thus, in order to show that vertical integration causes changes in care delivery, one needs to address the inherent differences between the physicians who decide to integrate with hospitals versus those who do not. We take advantage of our panel dataset and make use of a Difference-In-Differences (DID) fixed effects model that controls for unobserved differences between the comparison groups. To address the fact that the integration trend might be driven by other market factors such as technology, market demand, insurance structure, and socioeconomic factors, we link our panel data to other data sources and

adjust for a set of relevant covariates. To gain further confidence, we also conduct multiple robustness checks. Finally, we employ a few econometric methods (e.g., mediation analysis) to identify the drivers of the changes we observe, which, in turn, allow us to provide policy recommendations.

3.1.3 Main Findings and Contributions

We find that vertical integration negatively affects the quality of care delivery. After physicians integrate, they reduce recommended care processes such as deep sedation (about 3.9 fewer patients receiving deep sedation per 100 patients).³ Furthermore, the patients of integrated physicians experience a significant increase in major post-colonoscopy complications such as bleeding (5.7 per 1,000 colonoscopies) and minor complications such as GI or cardiac symptoms (4.7 per 1,000 colonoscopies). There is no clear impact of integration on operational aspects such as waiting time or physician efficiency. These effects occur even after adjusting for changes in patient composition and market characteristics. Thus, contrary to existing belief, our results indicate that vertical integration not only does not result in improvements in quality or operational efficiency but indeed worsens some aspects of them. Through mediation analyses, we further show that the reduction in the use of deep sedation is the main mechanism through which the increase in adverse outcomes occurs. Finally, we find that after integration, hospitals (a) no longer allocate expensive anesthesiologists to relatively unprofitable procedures such as outpatient colonoscopy, and (b) spend about 132.7 more dollars per colonoscopy procedure.

Taken together, our results indicate that, despite an increase in spending, vertical integration does not result in improvements in quality or operational efficiency. Rather, the shift in the incentive structures of the organization as a result of integration can generate unintended negative consequences. Moreover, our cost-effectiveness analyses reveal that increasing financial incentives can alter physicians' behavior; paying up to twice as much as

³Compared to other types of sedation, deep sedation requires more resources and coordination efforts since only anesthesiologists can administer it, whereas other types of sedation can be administered by nurses. This describes why the impact on deep sedation is higher than other sedation types.

the current price for deep sedation can be cost-effective. Overall, these findings show that future policies that promote integration should carefully align the incentives of integrated practices to prevent unintended consequences while promoting coordination.

3.2 Conceptual Framework

3.2.1 Background in GI Practices

Colonoscopy is a primary screening modality in GI practices for Colorectal Cancer (CRC) screening and diagnosis. During a colonoscopy, colonoscope—a long, flexible tube—is inserted into the rectum to examine the presence of polyps. Polyps can potentially develop into cancer and can be removed (“polypectomy”) by the physician during colonoscopy. Thus, the key to a good outcome is how well the physician removes the polyps or adenomas (a benign tumor that may be precursor lesions to CRC). However, polyp removal is not always straightforward, and because the physician controls the entire process, the outcome is highly dependent on the quality of the preparation (e.g., bowel preparation, types of sedation) and the skill of the physician. There are common post-colonoscopy complications such as bleeding, infection, and in rare cases, perforation of the bowel. Currently, there is a wide unexplained variation in physician quality and colonoscopy outcomes (Corley *et al.*, 2014).

3.2.2 Theories of Vertical Integration

Vertical integration in economics refers to the common ownership of two or more stages of production (or distribution) that are usually separate. In a vertical chain of healthcare, patients usually go through the chain of primary care physicians, specialists, hospitals, and rehabilitation facilities. In this chain, specialists and hospitals are typically considered to be the upstream and the downstream entities, respectively. Economic theories suggest that there are roughly two types of motivation for vertical integration: efficiency-based and strategy-based (Post *et al.*, 2018; Baker *et al.*, 2018). Efficiency based theories propose that providers integrate primarily to eliminate inefficiencies in production. Strategy based

theories propose that providers integrate to increase market power and/or employ the anticompetitive tactic to create barriers to entry (Gaynor, 2014).

The efficiency-based theory claims that if physicians and hospitals are under the same system, it can be easier to achieve smooth communication, reduction in duplicate services and waste, and goal setting/standardization of practices (Kocher and Sahni, 2011; Burns *et al.*, 2014; Baker *et al.*, 2018; Baicker, K. and Levy, H., 2013; Lammers, 2013). This will result in clinical integration, which management literature defines as the coordination of patient care services across the various functions, activities, and operating units of a delivery system (Gillies *et al.*, 1993). It is believed that clinical integration is the gold standard for improving care quality. Yet, there is limited direct evidence that vertical integration will achieve clinical integration (Singer *et al.*, 2018). On the other hand, the strategy-based theories suggest that integration for strategic purposes (e.g., vertical foreclosure) will have a less direct impact on the outcomes because the integration is to increase the market power and buy referrals. Such moves will not necessarily motivate organizations to achieve any clinical integration, or even worsen care if increased market power results in lower motivation to compete on quality. These theories predict the effect of integration in different directions. Because both positive and negative effects from the two theories are not inherently mutually exclusive, the pivotal question centers on the magnitude of changes.

The changes in financial ownership of the organization can result in the changes in financial structures both at the organizational and the physician level, which can, in turn, also affect the care delivery patterns. First, as we discussed earlier, the hospital-owned organization receives greater payment per most of the same outpatient services from Medicare and other private insurance. At the organizational level, a positive income shock will increase the marginal utility of income, which will decrease the quantity (McGuire, 2000). At the same time, if the quantity of care is not reimbursed at the margin, providers may increase their quantity. Second, when physician practices get acquired, physicians no longer face the direct financial risk, and compensation is largely based on salary. These changes can result in the changes in the service mix, as well as the overall level of service

quantity. Thus, integrated physicians may have less incentive for health care utilization, resulting in a reduction in operational efficiency (e.g., lower service volume or greater waiting time) after vertical integration.

3.3 Literature

Our study is related to the stream of literature that examines the operational efficiency and the quality of service organizations. In particular, it is relevant to studies that examine the role of public policy (e.g., payment policy) and the provider market structures (e.g., mergers, exit, competition) in the operation of service organizations (Chen and Savva, 2018; Song and Saghafian, 2019b). Within the context of vertical integration, most studies focus on their anticompetitive effect, i.e., how it affects spending and price (Neprash *et al.*, 2015; Baker, L. C. *et al.*, 2014). Our study focuses on how integration affects the organizations' operational behaviors and quality.

Within the operations management literature that examines worker behavior in service organizations, our study is related to the empirical studies that examine how organizational settings affect the operational efficiency and quality of services (Tan, T. F. and Netessine, S., 2019; Wang, J. and Zhou, Y., 2018; Meng *et al.*, 2018; Staats *et al.*, 2017). Studies have examined how specific characteristics of the organization affect worker behavior, such as the structural layout of the facility (Meng *et al.*, 2018) or a quality monitoring program (Staats *et al.*, 2017). Our study examines how the ownership of the organization affects the behavior of workers, and identify the specific characteristics that change as a result of the changes in ownership that affect the behavior.

Within the supply chain management literature, our study is related to the stream of literature on how vertical integration can improve efficiency through reducing the double marginalization problems, i.e., when upstream and downstream firms with monopolistic power fail to identify optimal cost (thus less profit) due to lack of information sharing. Double marginalization problem has been studied extensively in the operations management literature, mostly through supply chain models (Heese, 2007; Li *et al.*, 2013). Fewer studies,

however, have empirically evaluated how the behavioral changes within the integrating entities may influence the overall effect of vertical integration. Our study is thus relevant to an empirical examination of how the changes in the supply chain affect production quality. For example, some studies have examined how the particular characteristics of vertical integration (e.g., distance or resource sharing) affect the efficiency or the quality of production (Bray *et al.*, 2019; Ramdas and Randall, 2008).

Lastly, our study contributes to medical literature that explores the determinants of medical care quality (Song *et al.*, 2010). Specifically, there is a large unknown variation in the quality of CRC screening and diagnosis (Warren *et al.*, 2009; Rabeneck *et al.*, 2008). Our study thus addresses the question of how the health system can reduce the variations in physician practices and improve provider quality, thus guiding the clinical and public health practitioners.

3.4 Data and Study Setting

3.4.1 Data

Table 3.1 provides an overview of all data sources we have used. Our main data source is a 20% random sample of traditional FFS Medicare claims (Parts A and B) for inpatient, outpatient, and office visits between 2005-2015. The FFS Medicare claims provide detailed information on the healthcare services use of each patient visit, such as the procedures received through the Healthcare Common Procedure Coding System (HCPCS) codes, diagnosis through International Classification of Diseases (ICD-9) codes, and spending. We obtained each patient's sociodemographic information such as the age, sex, and 9-digit ZIP code from the Medicare Beneficiary Summary Files (BSF). We obtained each physician's information from the Physician and Other Supplier Public Use File from the Centers for Medicare & Medicaid Services (CMS), which provides information on the characteristics, utilization, and payment information on services and procedures provided to FFS Medicare beneficiaries by physicians (Centers for Medicare & Medicaid Services, 2016a). We also

aggregated the physician information directly from the FFS claims based on the physician-level National Provider Identifier (NPI) or unique physician identification number (UPIN) identifiers as well as the tax identification number (TIN) of individual physicians and/or practices. We incorporated the area level health care utilization, supply, and sociodemographic information from the Bureau of Health Professions' Area Resource Files (ARF), and the county-level penetration of Health Maintenance Organization (HMO) plan from the CMS State/County/Plan Enrollment Data.

3.4.2 Measuring Vertical Integration

There have been two streams of approaches that existing studies have taken to measure vertical integration: survey-based and claims-based. Many studies have used the survey data such as the American Hospital Association (AHA) Annual Survey or SK&A data of physician survey, which include questions on the hospitals or physicians' relationship with the other (Madison, 2004; Cuellar and Gertler, 2006; Scott *et al.*, 2017; Baker, L. C. *et al.*, 2014; Wagner, 2016; Capps, C. *et al.*, 2018; Koch *et al.*, 2017). Although survey data can provide a direct source of information on integration, they may miss small integration, be subject to misclassification, or do not capture physician-level changes. Other studies used claims-based approaches, which infers the status of physicians' or practices' integration status from the billing patterns (Neprash *et al.*, 2015; Konetzka *et al.*, 2018; Desai and McWilliams, 2018; Capps, C. *et al.*, 2018). These are based on the rationale that (a) providers have a strong financial incentive to report the services that occurred at the hospital-owned practices due to the payment differential between the HOPDs and physician offices, and (b) only the practices that are 100% owned by a hospital can bill at a higher HOPD rate (Neprash *et al.*, 2015; Konetzka *et al.*, 2018; Desai and McWilliams, 2018; Capps, C. *et al.*, 2018).

Using the claims-based approach, we measured integration as the following. For each physician j in year t , we first calculated the degree of integration as:

$$INTEG_{jt} = \frac{HOPD_{jt}}{HOPD_{jt} + OFFICE_{jt} + ASC_{jt}}, \quad (3.1)$$

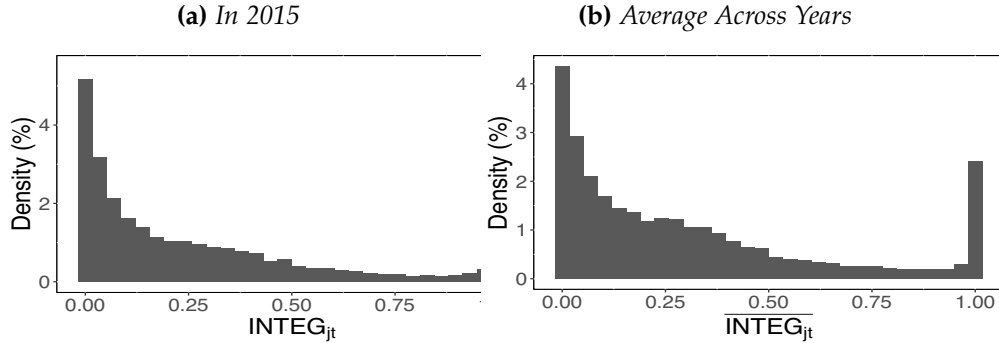
where $HOPD_{jt}$, $OFFICE_{jt}$, and ASC_{jt} represent the total number of unique HOPD, office, and ASC-based colonoscopy claims, respectively.⁴ Thus, $INTEG_{jt}$ takes a numeric value between 0 and 1 (Figures 3.3). We also note that a large proportion of physicians (i.e., about 49.1%) have $0 < INTEG_{jt} < 1$, which means they are neither fully independent (i.e., $INTEG_{jt} = 0$) nor fully integrated (i.e., $INTEG_{jt} = 1$). From further investigation, we confirm this is because most GI physicians practice at multiple practices, often split among hospital-owned and physician-owned ones.

We next defined a binary integration variable of integration I_INTEG_{jt} for each year for each physician given a threshold value thr as $= 0$ if $INTEG_{jt} < thr$ and $= 1$ otherwise. Previous studies used the threshold values of 50% or 75% to dichotomize the physicians into integrated versus independent groups (Neprash *et al.*, 2015; Desai and McWilliams, 2018). However, we note that there is a noticeable difference in the integration patterns among the gastroenterologist physicians vs. other specialists examined in these studies. Relatively fewer cardiologists or PCPs in these studies were “partially integrated” (e.g., $0 < INTEG_{jt} < 1$), whereas about half of the GI physicians who are partially integrated were the driving force of vertical integration. Thus, we used a lower threshold for defining integration in our study to capture the GI physicians who are partially integrated. Assuming that a physician who submits any claims under HOPD is vertically integrated with at least one organization, we defined the threshold for defining the binary integration variable as 10% in our main analysis. However, we later test our assumption on this threshold by varying them in our robustness check (Section 9).

Based on I_INTEG_{jt} , we defined the control group as the physicians who did not change their integration status throughout the observation period. These could be the physicians who either remain independent or who remain integrated ($I_INTEG_{jt} = 0$ for all t or $I_INTEG_{jt} = 1$ for all t). Our control group is the physicians who change their integration status (there exists at least one year such that $I_INTEG_{jt} = 0$ and $I_INTEG_{jt} = 1$). The

⁴Unique claims are defined as the claims with the same beneficiary ID, service date, and the provider NPI (UPIN).

Figure 3.3: Distribution of $INTEG_{jt}$



Note. For each physician j , $\overline{INTEG}_{jt} = \sum_t INTEG_{jt} / \sum_t T_t$ for years t for which the physician submitted any claims.

majority of physicians who integrated remained integrated after the year of first integration. However, there existed a small portion of physicians ($< 9.5\%$) who “disintegrated” after integrated. In our main analyses, we included these physicians. In the Robustness Check, we have examined the effect of removing them.

3.4.3 Study Population and Comparison Groups

Our patient population is the FFS Medicare beneficiaries who received colonoscopies at any outpatient care settings during our study period, are aged between 65 and 85 at the time of the procedure and are entitled to Medicare due to age.⁵ We also focused on the patients who have received colonoscopy from GI physicians.⁶ Although the vast majority of the physicians who perform colonoscopies are GI physicians, a small number of other specialists such as colorectal surgeons and primary care physicians also perform them. Because the other specialists are likely to have different baseline skills, training, as well as

⁵The USTSPF guideline recommends against screening above age 85 (Force *et al.*, 2008). Thus, we removed those above age 85 when they received colonoscopy from our main analysis, as are likely to be clinically different from the rest. We also limited the analysis to those aged 65 or above, since the Medicare beneficiaries under 65 are often sicker than a typical Medicare population (Cubanski *et al.*, 2016). To ensure the observation of post-procedure adverse outcomes, we further restricted our analysis to those who have continuous enrollment in the FFS Medicare Parts A and B of one year before and 30 days after the colonoscopy date.

⁶Gastroenterologists, or GI physicians, were identified by the specialty code on claims (gastroenterology=10). We removed the colonoscopies performed by physicians with specialties such as “Colorectal surgery (=28)”, “Internal medicine (=11)”, or “family practice (=08).”

patient characteristics (Baxter *et al.*, 2012), we removed them from our analysis.

3.5 Variable Definitions and Descriptions

3.5.1 Outcome Variables

We divide our outcome variables into four categories: process-related quality, outcome-related quality, operational efficiency, and spending.⁷ Below, we describe each of these separately.

Process-Related Quality. We selected the process-related quality measures based on the endorsed quality indicators by professional societies (Rex *et al.*, 2006). Among the suggested quality indicators, we selected those obtainable from the claims data that are (a) widely accepted, and (b) have the potential to be affected by the change in practice settings. The definitions are in Table 3.2, and approaches more specific information are provided in the Appendix. We also address some of the limitations of using claims data for measuring quality in our robustness checks (Section 9). Studies show that GI physicians with higher polypectomy rates tend to have better patient outcomes such as lower interval cancer (Warren *et al.*, 2009; Kaminski *et al.*, 2010). Thus, we measured the rate of polypectomy of physicians as a proxy for their process-related quality, after adjusting for various patient risk factors. We identified polypectomy from claims by the concurrent pathology bills (Warren *et al.*, 2009). Incomplete colonoscopies can also be used as another proxy for measuring process quality, since they can result in missed lesions, a contributor to the interval cancer (Cooper *et al.*, 2012). We obtained incomplete colonoscopies directly from the HCPCS modifier codes 53, 73, or 74 on colonoscopy claims.⁸

Another key process measure of colonoscopy quality that we examined is the method

⁷In our robustness checks (Section 8), we also consider patient experience as yet another outcome variable to test the validity and the spread of the results we observe using our main outcome variables.

⁸Modifier 53 indicates a discontinued procedure of physician services. Modifier 73 indicates a discontinued HOPD/ASC procedure prior to the administration of anesthesia. Modifier 74 indicates a discontinued HOPD/ASC procedure after the administration of anesthesia. Both modifiers 73 and 74 apply to facility charges.

Table 3.2: Quality Measures and Definitions

Category	Variable	Definition
Process-Related Quality	Polypectomy	Removal of at least one polyp during a colonoscopy.
	Incomplete colonoscopy	A colonoscopy that does not evaluate the colon past the distal third of the colon.
	Deep sedation	Use of propofol sedation during colonoscopy.
Outcome-Related Quality	Perforation	Incidence of a hole in the wall of part of the gastrointestinal tract.
	Gastrointestinal bleeding ¹	All forms of bleeding in the gastrointestinal tract.
	Infection	Incidence of bacterial infections after colonoscopy.
	Cardiac symptoms	Incidence of arrhythmia, congestive heart failure, cardiac or respiratory arrest, syncope, hypotension, or shock.
	Minor GI symptoms	Incidence of paralytic ileus, nausea, vomiting, dehydration, abdominal pain, diverticulitis, and enterocolitis.
	Interval cancer	Incidence of CRC 6 to 36 months after a negative colonoscopy.
Operational Efficiency	GI physician efficiency (throughput)	Total number of procedures per GI physician per year.
	Anesthesiologist efficiency (throughput)	Total number of procedures per anesthesiologist per year.
	Time to complete colonoscopy	Time interval between incomplete colonoscopy to next complete colonoscopy.
	Time to treatment	Time interval between confirmatory colonoscopy to initiation of cancer treatment
Spending	Spending per procedure	Total spending occurred during the service event.
	Spending per episode	Total spending occurred during the episode, i.e., the day of, and the 30 days after the index service event.

Note. All process and outcome variables were measured at the patient level. All variables were measured primarily from the FFS Medicare inpatient, outpatient, and carrier claims. SEER-Medicare data were used to supplement for the interval cancer. ¹Major bleeding events include intracranial hemorrhage, hemoperitoneum, and inpatient or emergency department stays for gastrointestinal, hematuria, or not otherwise specified hemorrhage. Minor bleeding events included epistaxis, hemoptysis, vaginal hemorrhage, hemarthrosis and any outpatient claim for hematuria, gastrointestinal, and not otherwise specified hemorrhage.

of sedation. Previously, the primary sedation method for screening colonoscopies had been through midazolam and an opioid. More recently, propofol sedation for outpatient colonoscopies has increasingly become popular (Khiani *et al.*, 2012). Evidence shows that propofol sedation can provide fast onset of action, short duration of action, amnestic effects, faster recovery and discharge times and increased patient satisfaction (Chen and Rex, 2004). However, there is a wide variation in the adoption of propofol for outpatient endoscopy. Thus, we measured the use of deep sedation (i.e., propofol as a method of anesthesia) as another indicator of procedure quality. We used the presence of anesthesiologist or nurse anesthetist to identify anesthesiology involvement (Cooper *et al.*, 2012; Khiani *et al.*, 2012).⁹

⁹Because of the FDA regulation, another provider (i.e., an anesthesiologist or nurse anesthetist) must be present during the endoscopic procedure if propofol sedation is used during a colonoscopy. We followed the existing studies that relied on the presence of the CPT-4 code 00810, anesthesia assistance with endoscopic procedure distal to the duodenum, occurring on the same date as the colonoscopy of interest.

Outcome-Related Quality. We examined the three most common major complications—perforation, bleeding, and infection—that can result in serious health conditions, including death (Rex *et al.*, 2006; Rabeneck *et al.*, 2008). We also measured minor complications, defined as the presence of minor GI or cardiac symptoms discussed in the medical literature (Warren *et al.*, 2009). For all complications, we identified these events using ICD-9 codes that are present either on or within seven days after colonoscopy. Finally, we measured the downstream health outcome, interval cancer, which is a type of CRC that occurs despite receiving a screening colonoscopy. We identified the interval cancer based on existing claims-based approaches (Quantin *et al.*, 2012). The colonoscopy that was received six to 36 months before the diagnosis of cancer was defined as the one attributed to cancer. Conditional on genetic and clinical variations, interval CRC may occur due to an inadequate polypectomy or missed lesions. Thus, a higher interval cancer rate is a signal of poor physician quality after adjusting for the patient risk factors (Kaminski *et al.*, 2010). The details of measurement and validation are in the Appendix.

Operational Efficiency. We used physicians' throughput and patients' waiting times as measures of operational efficiency. We calculated throughput as the total colonoscopy performed per physician per year, the total number of services provided per physician per year, the total number of unique procedures given per physician per year, and the total number of unique patients treated per physician per year. Next, we measured two waiting times: time from incomplete colonoscopy to the next follow up, and time from positive colonoscopy to the initiation of surgery. To address the physician attribution problem where a patient receives initial and follow up colonoscopy from different physicians, we limited the analysis to the patients who received the two procedures from the same organization.

Spending. Although Medicare prices are administratively set, we examined the changes in spending per procedure to examine whether there is a change in the procedure mix that affects spending. For example, colonoscopy reimbursement rate varies by the type of specific procedures to remove polyps. Spending was defined as the total amount paid to the provider per procedure (e.g., a unique date-physician-provider triplet), obtained from the

beneficiary's claims.

3.5.2 Independent Variables

We divide the independent variables used in our models into three categories: patient characteristics, physician characteristics, and market characteristics. Below, we describe each of these separately.

Patient Characteristics. In our models, we controlled for patient characteristics including demographic information such as age, gender, race/ethnicity, the reason for Medicare entitlement (i.e., whether or not a beneficiary is entitled to Medicare due to end-stage renal disease, or ESRD), and Medicare-Medicaid dual eligibility status (or "Duals"), a proxy used for low-income status. We accounted for the heterogeneity in patients' overall health risk by calculating each patient's Charlson Comorbidity Index (Elixhauser Comorbidity Index) (Elixhauser *et al.*, 1998).¹⁰ We also calculated the indicators for chronic conditions from the Chronic Conditions Data Warehouse algorithm (Chronic Condition Data Warehouse, 2014), which uses diagnosis and procedure codes from the previous year to determine the 27 chronic conditions of the patient.

Physician Characteristics. We measured the physician characteristics such as his/her geographic region of practice, gender, the total number of affiliated practices measured as tax id number (TIN), and the indication of whether s/he is affiliated with a multispecialty clinic. For example, we counted the number of physician specialties affiliated with each organization and indicated the organization as a multispecialty practice if there were specialists other than gastroenterology or anesthesiology. Finally, we identified the ASC status of each physician's practice based on whether the practice submitted any ASC-based claims.

Market Characteristics. We controlled for the market concentration for a variety of reasons including the fact that horizontal and vertical integration in the market can be correlated.

¹⁰The Elixhauser Comorbidity Index includes 30 diagnoses that can potentially increase the probability of adverse outcomes. We calculated the index directly from the patient's inpatient and outpatient claims history in the previous year and used the total number of chronic conditions in our main model (Elixhauser *et al.*, 1998).

Table 3.3: Definition of Independent Variables and Data Sources

Variable	Description	Data source
Patient characteristics		
Age	Numeric, 64-86.	Medicare BSF
Gender	Binary, male or female.	Medicare BSF
Race	Factor, White, Black, Hispanic, Asian, or others.	Medicare BSF
Medicare entitlement	Binary, ESRD or not.	Medicare BSF
Medicaid eligibility	Binary, dual or non-dual.	Medicare BSF
Comorbidity	Numeric, from 0 (least severe) to 21 (most severe).	Medicare inpatient, outpatient claims
Chronic conditions	Numeric, from 0 to 27.	Medicare BSF
Location	Binary, rural or urban.	Medicare Cost Report, POS
Physician characteristics		
Number of affiliations	Numeric, greater than 0	Medicare inpatient, outpatient claims
Multispecialty	Binary, 0 or 1	Medicare inpatient, outpatient claims
ASC affiliation	Binary, 0 or 1	Medicare inpatient, outpatient claims
Market characteristics		
Managed care penetration	Numeric, from 0 (no penetration) to 1 (full penetration).	State/County/Plan Enrollment Data
Herfindahl-Hirschman Index	Numeric, greater than 0	Medicare inpatient claims, Medicare Cost Report
Provider supply	Numeric, number of GI physicians per 10,000 persons.	AHRF

To measure market concentration, we computed Herfindahl-Hirschman Indices (HHIs) for hospitals and specialist physicians for each market (HRR/HSA). HHIs were calculated by summing the squared market shares of firms.¹¹ We used population, number of Medicare beneficiaries, number of GI physicians, and median household income from ARF. Finally, we calculated and included variables related to the insurance market structure (e.g., HMO penetration rates).

3.5.3 Descriptive Statistics

Table 3.4 summarizes the cross-sectional patient characteristics and the outcome variables by their physicians' integration status I_INTEG_{jt} in a given year. The cross-sectional measure of outcome rates was consistent with the estimates from existing literature (Rex *et al.*, 2006). Overall, patients who receive colonoscopy from integrated physicians are more likely to be Black race, more likely to be Duals, have higher comorbidity index, and more likely to

¹¹According to the Department of Justice (DOJ) and Federal Trade Commission (FTC), the Horizontal Merger Guidelines, markets with HHIs below 1,500 are considered as unconcentrated, those with HHIs of 1,500–2,500 as moderately concentrated, and those with HHIs above 2,500 as highly concentrated (Shapiro, 2010).

Table 3.4: Patient Characteristics and Outcome Variables by Integration Status

	Independent	Integrated
Observations (N)	1,443,329	2,119,634
Patients (N)	1,123,285	1,628,126
Demographic		
Age (mean)	71.38	70.73
Gender, Male (%)	44.29	44.03
Race, White (%)	82.88	82.24
Race, Black (%)	9.32	11.10
Race, Asian (%)	2.06	1.38
Race, Hispanic (%)	2.05	1.76
Duals (%)	14.51	17.90
Rural (%)	12.47	18.85
Comorbidity (mean)	1.53	1.60
Chronic conditions (%)	94.11	93.96
Process-Related Quality		
Polypectomy (%)	59.77	59.44
Incomplete colonoscopy (%)	1.84	1.88
Deep sedation (%)	56.58	39.65
Outcome-Related Quality		
Perforation (%)	0.12	0.16
Bleeding (%)	29.67	32.39
Infection (%)	0.21	0.31
Minor complications	13.27	15.62
Interval CRC (%)	0.23	0.25
Operational Efficiency		
Total colonoscopies per year (N/physician/year)	181.09	156.18
Total services per year (N/physician/year)	1,457.58	947.14
Total procedure types per year (N/physician/year)	39.20	40.21
Total patients seen year (N/physician/year)	462.35	416.61
Median time to followup (days)	24	28
Median time to surgery (days)	32.79	33.83
Spending		
Total spending per colonoscopy (USD)	264.74	552.23
Physician spending per colonoscopy (USD)	207.50	191.82
Facility spending per colonoscopy (USD)	410.94	585.07

Note. Patients' integration status is assigned based on the physician they received colonoscopy from. Independent physicians have $I_INTEG_{jt} = 0$, and integrated physicians have $I_INTEG_{jt} = 1$. All characteristics differed at the significance level 0.001.

reside in rural areas. The patients of integrated physicians are more likely to have higher unadjusted adverse outcomes such as perforation, bleeding, infection, minor complications, and interval CRC. Notably, integrated physicians provide deep sedation substantially less, about 70% of the use of the independent physicians. The fixed differences between the integrated and independent physicians and patients are adjusted in our DID analysis.

3.6 Main Empirical Strategy

3.6.1 Overview

Our main empirical strategy is based on a DID analysis with the physician, area, and year fixed effects. Under certain assumptions that we describe in the next section, the coefficient for the treatment variable in our model can provide a causal interpretation on how vertical integration affects care delivery. The unit of analysis in our model is a colonoscopy, and the treatment status for each colonoscopy was based on the integration status of the physician who performed the procedure. We also accounted for individual, physician, and market characteristics described above. We allowed multiple colonoscopies provided on the same patient to have different treatment status if the patient received multiple colonoscopies from physicians. The majority (73.0%) of patients, however, received only one colonoscopy during the study period. To estimate the effect of integration by performing DID analysis, we made use of the following model:

$$Y_{ijt} = \alpha POST_{jt} + \beta \mathbf{X}_{ijt} + \gamma \mathbf{Z}_{jt} + PHYSICIAN_j + MARKET_i + YEAR_t + \epsilon_{ijt}, \quad (3.2)$$

where Y represents the outcome variables such as process-related quality, outcome-related quality, spending, and operational efficiency, $POST$ is a binary variable that indicates that the observation is made post-integration for the treated group, $PHYSICIAN$ is the physician fixed effect, $MARKET$ is the market fixed effect, and $YEAR$ is the year fixed effect. $\mathbf{X}$ is the vector of patient characteristics, $\mathbf{Z}$ is the vector of market characteristics, and ϵ is an error term. Indices i , j , and t represent a patient, physician, and year, respectively. Bold notation is used to represent vectors.

3.6.2 Main Assumptions

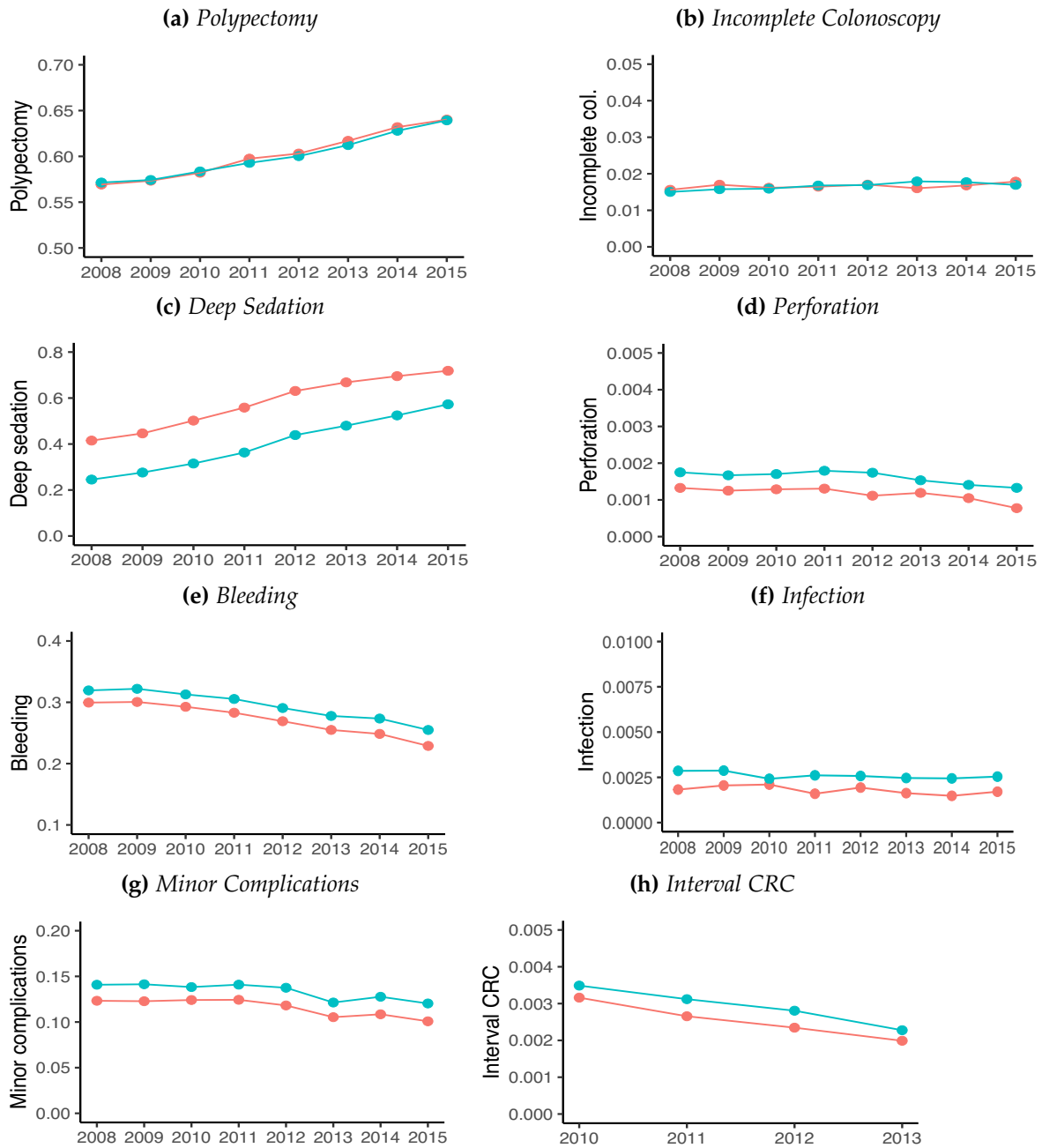
The main assumptions of our identification strategy that are needed to provide a causal interpretation are: (a) parallel trends between the treatment and the control group, and (b) strict exogeneity. The parallel trend assumption imposes that, in the absence of treatment,

the difference between the treatment and the control group is constant over time. There is no statistical test that can directly test this assumption, so we first examine the pre-integration trends for our outcomes. Our visual inspection in Figure 3.4 confirms that there is a similar trend in outcomes before the physicians integrate. However, we note that the outcomes that are relatively rare, such as incomplete colonoscopy, perforation, infection, and interval cancer, the parallel trend is less stable, likely because there is a fewer number of observations. Yet, the outcomes that show significant changes based on our DID analysis, such as deep sedation, bleeding, and minor complications, show more stable parallel trends. We also statistically test the differences in trend between the two groups by interacting each of the pre-integration years with our treatment variable. None of the interaction terms are significant, suggesting there is no significant difference in time trends between the comparison groups before the integration.

The strict exogeneity condition assumes that the regressors are uncorrelated with the error terms. Such an assumption can be violated if, for example, the errors are correlated with unobserved and time-varying characteristics. Our rich set of covariates for patient and area characteristics, as well as the fixed effects at multiple levels, can address the heterogeneity between the comparison groups and year-specific shocks. However, there are still three sets of threats in our setting that can yield violations of the strict exogeneity assumption: patient selection to physicians, physician selection to integration status, and the changes in physician behavior post-integration. Below, we discuss each of them separately.

Patient Selection to Physicians. Patients may select into different physicians in a way that is unobservable to us. For example, when the physicians join a hospital or a large healthcare organization, the patients' perception of the quality of the service may change such that poorer or sicker patients select into integrated physicians. This can make it appear that integration worsens quality. Indeed, we have observed some baseline differences in patient characteristics in our data (Table 3.4). However, the concern for such selection is mitigated, because the selection has to be both time-varying and unobservable, which is not to be the case. This concern is further mitigated since we adjust for CRC specific risk, overall

Figure 3.4: Trends of Process-Related and Outcome-Related Quality



Note. Col. indicates colonoscopy. Red lines indicate the control group. Green lines indicate the treatment group before integration. Because the physicians in our treatment group integrate at different years, we do not show the post-integration trends.

comorbidity, and chronic conditions, and sociodemographic status. Also, relative to the choice of primary care physicians (PCPs), patients are less likely to choose their specialists

directly but are referred by their PCPs (Barnett *et al.*, 2012). Nevertheless, we later examine the changes in observed patient risk composition before and after the physicians integrate to further address this concern.

Physician Selection to Integration Status. We also note that the physicians' decision to integrate is unlikely to be exogenous, but can be a result of their strategic behavior. For example, those who decide to integrate may prefer to collaborate with others more, which may also be correlated with their quality. It is also possible that the physicians' decisions to integrate are based on the pre-integration characteristics of their patient group. For example, physicians with a greater proportion of low-income or high-risk patients may decide to integrate to alleviate financial struggles. Finally, the market characteristics such as the degree of horizontal integration, the degree of managed care penetration, or the input costs may be correlated with both physicians' propensity to integrate and underlying patient health (Gaynor, M. *et al.*, 2013). These threats again only apply to an extent the differences are unobservable and time-varying, which will be mitigated by various controls we include using our multidimensional panel data. Nevertheless, we further examine the physicians selection effect in multiple ways, including examining the effect of market conditions and making use of an inverse probability weighting (IPW) method to adjust for the physicians' propensity for integration.

Physician Behavior Post-Integration. Integrated physicians may change their behavior in a way that confounds the integration effect. For example, physicians affiliated with an integrated organization are more likely to increase the coding intensity for reporting complications, which may affect the outcome-related quality measures we examine without affecting the true underlying quality. In particular, if physicians code for complications after colonoscopy more actively after they integrate (e.g., due to increased monitoring efforts), it may appear that the quality has worsened post-integration. However, we believe that this concern is mitigated for several reasons. First, major complications are less likely to be subject to variations in this coding behavior than minor complications, but our results show a stronger effect for major complications—bleeding. Second, we account for the fact

that the patients can visit any of the inpatient, outpatient, or office settings for subsequent adverse outcomes, and are not limited to the same practice that they originally visited. This adverse outcome is, however, attributed to the original practice where they received a colonoscopy, not the practice they visit with the complications. Third, our examination of patient composition shows that the proportion of high-risk patients is fairly consistent, which weakens the argument that providers increase their coding intensity (Figure 3.7). Nevertheless, to gain further confidence, in our robustness checks, we test these assumptions in various ways.

3.7 Results and Discussions

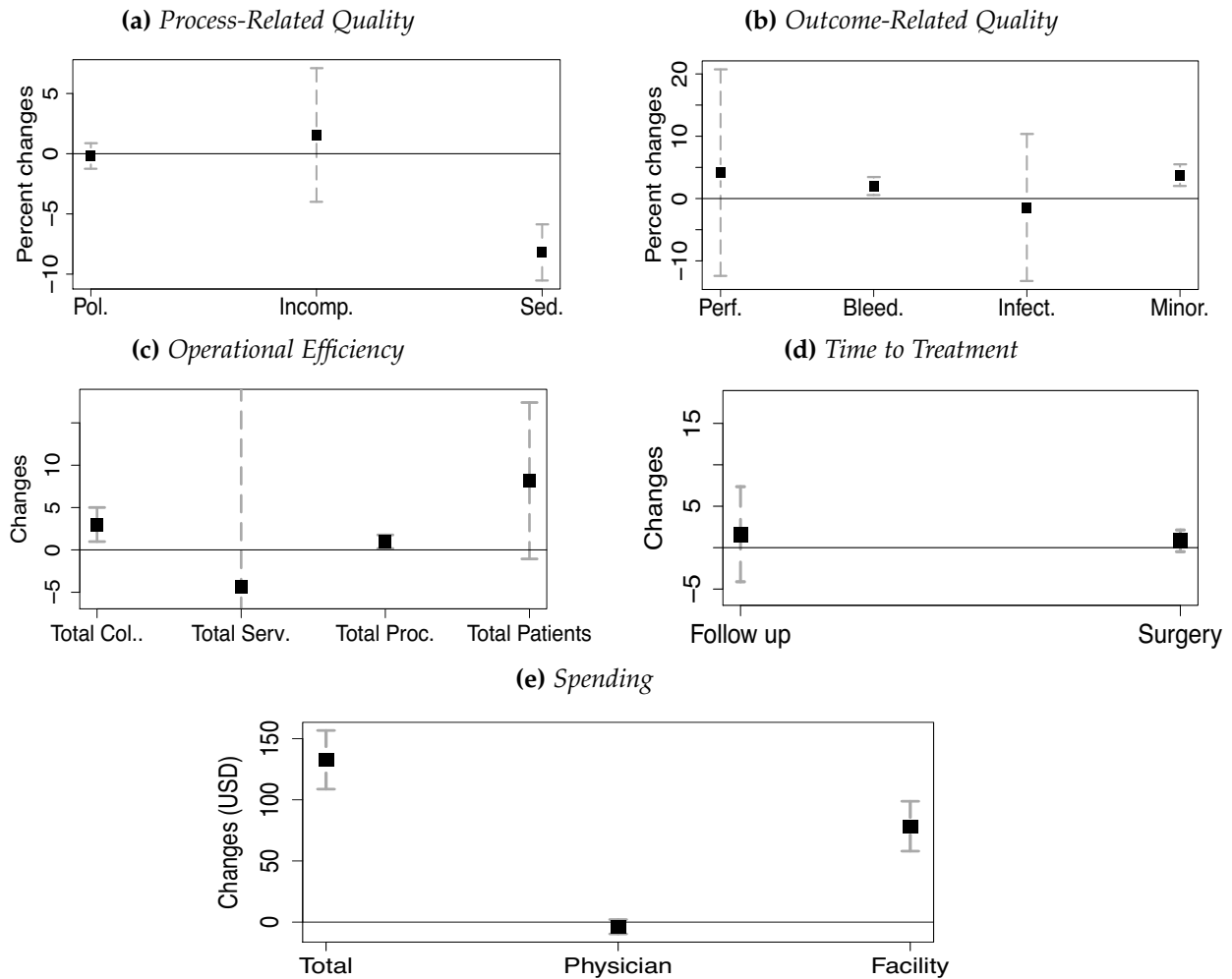
3.7.1 Average Effect

Process-Related Quality. Figures 3.5, Table 3.5, and Table 3.6 show the estimates of our DID coefficients. Among the three process-related measures (polypectomy, incomplete colonoscopy, and deep sedation), polypectomy and incomplete colonoscopy do not change after integration. However, the physicians who integrate reduce the use of deep sedation by 8.2%, equivalent to about 3.9 fewer patients receiving deep sedation per 100 patients receiving colonoscopies.

Outcome-Related Quality. Among the outcome-related quality (perforation, bleeding, infection, minor complications, and interval cancer), we observe that patients experience a significant increase in bleeding and minor complications after colonoscopy after their physicians integrate. Because both are relatively common complications (e.g., the average 30-day incidence of bleeding and minor complications are 28.6% and 12.6%, respectively), such increases translate into about 5.7 and 4.7 additional bleeding and minor complications out of 1,000 colonoscopies.

Operational Efficiency. Overall, when GI physicians integrate, there are no significant changes in the waiting time, either to a follow-up colonoscopy or surgery after a positive colonoscopy (Figure 3.5). Similarly, there is a minimal change in other throughput measures.

Figure 3.5: *Difference-in-Differences Estimates: Average Effect of Integration*



Note. Pol. indicates polypectomy. Incomp. indicates incomplete colonoscopy. Sed. indicates deep sedation. Perf. indicates perforation. Bleed. indicates bleeding. Infect. indicates infection. Minor. indicates minor complications. Col. indicates colonoscopy. Serv. indicates services. Proc. indicates procedures. Each dot indicates the size of the DID coefficient. Each dot indicates the size of the DID coefficient. Grey lines depict the 95% confidence intervals around the coefficient of the DID variable. Standard errors are robust and clustered at the physician and the year levels.

There is no measurable change in the total number of services provided per physician per year, and the total number of patients treated per physician per year. Although physicians experience a significant increase in the total number of colonoscopies and the total number of unique procedures, the effects are very small, i.e., 3.0 and 0.96 increase respectively. Compared to the average of 165.6 and 39.6, these changes are relatively small.

Spending. Our results show that for each colonoscopy visit, a physician's integration is

Table 3.5: Regression Result: Process-Related and Outcome-Related Quality

	<i>Dependent variable:</i>							
	Pol. (1)	Incomp. (2)	Deep Sed. (3)	Perf. (4)	Bleed. (5)	Infect. (6)	Minor Comp. (7)	Int. CRC (8)
POST	−0.0011 (0.0016)	0.00026 (0.00047)	−0.039*** (0.0056)	0.00006 (0.00012)	0.0057** (0.0016)	−0.000033 (0.00013)	0.0047** (0.0011)	−0.00018 (0.00027)
Obs.	2,582,769	2,582,769	2,582,769	2,582,769	2,582,769	2,582,769	2,582,769	1,631,431
R ²	0.096	0.031	0.528	0.013	0.104	0.019	0.083	0.020
Adj. R ²	0.091	0.026	0.525	0.008	0.099	0.013	0.078	0.012
Res. Std. Err.	0.467	0.126	0.344	0.038	0.429	0.047	0.319	0.056

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Pol. indicates polypectomy; Incomp. indicates incomplete colonoscopy. Dee Sed. indicates deep sedation; Perf. indicates perforation; Bleed. indicates bleeding. Infect. indicates infection; Minor Comp. indicates minor complication; Int. CRC indicates interval CRC; Obs. indicates observations; Adj. R² indicates adjusted R²; Res. Std. Err. indicates residual standard error.

Table 3.6: Regression Result: Operational Efficiency and Spending

	<i>Dependent variable:</i>						
	Total Col. (1)	Total Serv. (2)	Total Proc. (3)	Total Pat. (4)	Total Spend. (1)	Phy. Spend. (2)	Fac. Spend. (3)
POST	2.993* (0.939)	−4.387 (18.965)	0.962** (0.160)	8.167 (3.950)	132.716*** (12.238)	−3.790*** (0.506)	78.412*** (10.404)
Obs.	73,793	28,008	38,626	37,555	2,582,824	2,582,824	1,090,038
R ²	0.869	0.925	0.915	0.944	0.369	0.257	0.352
Adj. R ²	0.843	0.876	0.881	0.921	0.365	0.253	0.344
Res. Std. Err.	52.547	664.106	5.899	79.736	290.830	60.239	292.282

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Col. indicates colonoscopy; Serv. indicates services; Proc. indicates procedures; Pat. indicates patients; Spend. indicates spending; Phy. indicates physician; Fac. indicates facility; Obs. indicates observations; Adj. R² indicates adjusted R²; Res. Std. Err. indicates residual standard error.

associated with \$132.7 increase in total Medicare spending for colonoscopy (Figure 3.5). The effect is driven by an increase in the facility fee of about \$78.4 after a physician integrated.

In summary, our main analysis shows that integration affects various dimensions of healthcare. There were insignificant or minimal changes in operational efficiency measures as well as some procedure-related quality measures after physicians integrated. However, there was a significant reduction in one aspect of the procedure— the use of deep sedation.

In addition, patients of integrated physicians experienced a substantial increase in some outcome-related quality measures such as bleeding and minor complication rates. In the following sections, we further examine the mechanisms that drive our results.

3.7.2 Drivers of Main Results

Drivers of Increase in Adverse Outcomes. The increase in adverse outcomes can be driven by at least three factors: changes in the process-related quality, changes in the operational efficiency, and other unmeasured changes accompanying integration. First, insights from operations management literature suggest that changes in operational processes in the service system, such as the service speed and/or the customers' waiting time, can affect the quality of service (Chan *et al.*, 2016; KC and Terwiesch, 2009). Although we only observe a small amount of increase in some measures of operational efficiency, we test if such an increase is a driver of adverse outcomes. Second, medical literature indicates several benefits of deep sedation during colonoscopies such as the fast onset of action, short duration of action, amnestic effects, and faster recovery and discharge times, which in turn can improve quality (Chen and Rex, 2004). Although there is no evidence of a direct link between deep sedation rates and reduction in bleeding or minor complication rates, the reduction in deep sedation among integrated providers may have resulted in increased 30-day complications by reducing some aspects of quality. Finally, the organizational changes as a result of integration unmeasured by our data may have driven physicians to reduce quality. To test these three possibilities, we made use of mediation analysis (see, e.g., Baron and Kenny 1986): we first included deep sedation or physician's efficiency as an independent variable and re-examined the effect of integration on adverse outcomes (Baron and Kenny, 1986). Specifically, we use the following model:

$$Y_{ijt} = \alpha POST_{jt} + \delta SEDATION_{ijt} + \beta \mathbf{X}_{ijt} + \gamma \mathbf{Z}_{jt} + PHYSICIAN_j + MARKET_i + YEAR_t + \epsilon_{ijt}, \quad (3.3)$$

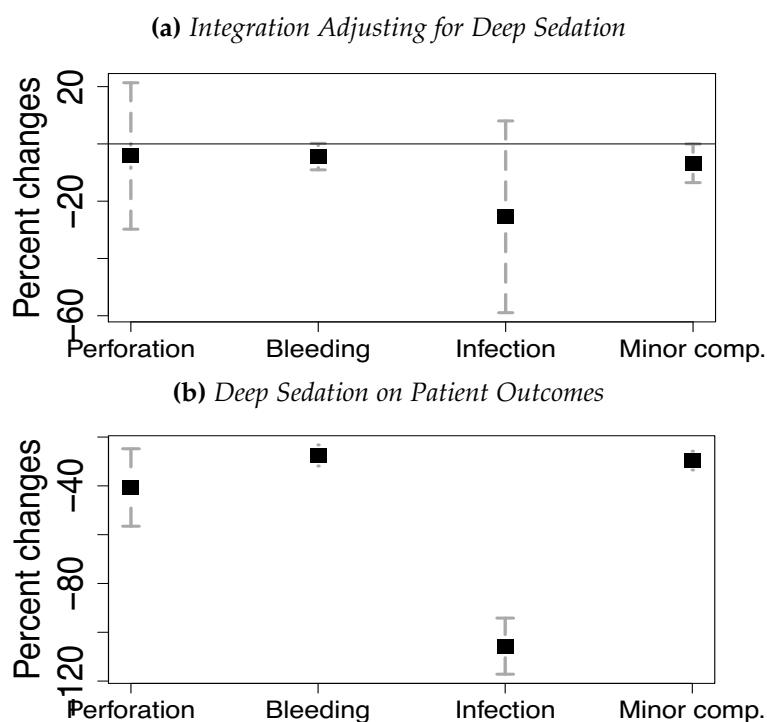
where $SEDATION_{ijt}$ is the binary variable for whether deep sedation was accompanied,

and all variables are as previously defined. For examining the role of efficiency, we substituted $SEDATION_{ijt}$ with $EFFICIENCY_{jt}$, which is the variable of the total number of colonoscopies provided per year for physician j in year t .

Our results presented in Figure 3.6 show that the effect of integration on adverse outcomes is either significantly reduced in magnitude or is no longer statistically significant after adjusting for deep sedation during the procedure, whereas the effect is consistent after adjusting for physicians' efficiency. At the same time, the coefficient for deep sedation is significant and negative, indicating that providing deep sedation is associated with a reduction in adverse outcomes. Lastly, after including deep sedation, the coefficient for integration is no longer significant or positive. We did not observe such changes when we included operational efficiency instead of deep sedation. This suggests that deep sedation is more likely a mediator of adverse outcomes than the physician's operational efficiency: the reduction in the use of deep sedation among integrated physicians explains an increase in some of the adverse health outcomes.

Drivers of Reduction in Deep Sedation. Our results suggest that integration reduces the use of recommended services such as deep sedation, which in turn increase adverse outcomes among patients. We take a closer look at the potential reasons the providers change the use of deep sedation. Compared to other types of sedation, deep sedation requires more resources and coordination efforts because only anesthesiologists can administer it, whereas other types of sedation can be administrated by nurses. Thus, the provision of deep sedation can be more sensitive to resource constraints, i.e., the anesthesiologist availability, that an organization faces. There are two potential ways that integrated GI physicians face anesthesiologist supply constraints: (1) Constraints in the external margin (e.g., fewer total number of available anesthesiologists) and (2) changes in the internal margin (e.g., fewer provision of deep sedation for colonoscopy per anesthesiologist). For the first case, an example would be when a physician gets hired to an integrated practice that owns a fewer number of anesthesiologists. For the second case, an example would be when an acquired practice allocates anesthesiologists to procedures other than colonoscopies.

Figure 3.6: Difference-in-Differences Estimates: Average Effect



Note. Minor comp. indicates minor complications. All effects are scaled as changes in percentages. Each dot indicates the size of the DID coefficient. Grey lines depict the 95% confidence intervals around the coefficient of the DID variable. Standard errors are robust and clustered at the physician and the year levels.

To examine the first pathway, we tested whether a GI physician experiences a reduction in the availability of anesthesiologists. For each GI physician, we measured the total number of the available anesthesiologist as the number of a unique anesthesiologist who has ever worked with a GI physician in a given year.¹² A typical GI physician worked with 6.6 anesthesiologists per year. We next used a DID model similar to our main model and used the total number of anesthesiologists a GI physician worked with per year as an outcome variable to identify whether the change in integration status of a GI physician is associated with a change in the number of available anesthesiologists. We do not find such effect (DID coefficient: -0.23 , SD: 0.15 , p-value: 0.18). To examine the second pathway, we measured

¹²A GI physician and an anesthesiologist were considered to have worked together if they treat the same patient on the same date.

the average number of deep sedation for colonoscopy performed by an integrated vs. independent anesthesiologist. We find that a typical independent anesthesiologist provides deep sedation for 140.2 colonoscopies, whereas a typical integrated anesthesiologist deep sedation for 112.3 colonoscopies.

These provide suggestive evidence that the reduction in anesthesia use is driven more by the changes in the internal, as opposed to the external margins. Indeed, anecdotal evidence from physicians who work at HOPDs suggests that administrative processes for accessing anesthesiologists for colonoscopies are challenging because they usually have to compete with other procedures for their availability. One potential explanation for such a challenge is the current payment structure for colonoscopies and deep sedation. The average payment for deep sedation per GI procedure via anesthesiologist involvement in HOPDs for FFS Medicare patients is low, only \$157.3 in 2012. The reimbursement rate is similar to physician offices and hospitals. A practice that focuses on providing large volume colonoscopy may still make profits, despite a high fixed cost of hiring anesthesiologists and low per procedure revenue, by economies of scale (e.g., attracting more patients, making each procedure shorter, and also providing large volume). In fact, most GI offices or ASCs have a dedicated anesthesiologist. However, for integrated practices potentially jointly managed with other specialties or practices, if the supply of anesthesiologists is fixed, if there are other competing practices with greater revenue (e.g., orthopedics or pain management), the opportunity cost of providing deep sedation through anesthesiologist for the practice may be low and may make more sense to allocate anesthesiologist to other procedures. This could potentially have affected the organization such that it allocates anesthesiologists to the services that generate greater revenue. Our results can only provide speculative evidence in this regard. However, it is worth noticing that the adverse consequences that occur as a result of changes in deep sedation use are substantial, and if these misaligned incentives result in low colonoscopy quality, fixing it can potentially have a substantial impact on improving the quality of colonoscopies. We later discuss the amount of reasonable incentive if it can change integrated physicians' behavior to provide more deep sedation.

Table 3.7: *Difference-in-Differences Estimates: Heterogeneous Effect of Integration*

Outcomes	Coefficients	Intensity	ASC Affiliation	Reasons	Competition
Deep sedation	Integration	-0.036*** (0.0054)	-0.045*** (0.0063)	-0.039*** (0.0063)	-0.046*** (0.0054)
	Integration x ASC		0.022** (0.0061)		
	Integration x hired			0.0034 (0.011)	
Bleeding	Integration x HHI				0.0046* (0.0018)
	Integration	0.0067** (0.0016)	0.0019 (0.0025)	0.0042* (0.0016)	0.0058† (0.0027)
	Integration x ASC		-0.0001 (0.0026)		
	Integration x hired			0.016* (0.0059)	
Minor complications	Integration x HHI				0.0001 (0.0001)
	Integration	0.0049** (0.0012)	0.0034† (0.0018)	0.0038* (0.0014)	0.0054* (0.0016)
	Integration x ASC		-0.0016 (0.0019)		
	Integration x hired			0.0088† (0.0046)	
	Integration x HHI				0.0001 (0.0001)

Note. † $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. HHI was scaled out of 1,000. Standard errors are robust and clustered at the physician and the year levels.

3.7.3 Heterogeneous Effect

We next examined whether the effects differ by the nature of integration, such as the degree of integration, affiliation with ASC, whether the physician integrated as a result of employment vs. acquisition of practice, and market competition environment.

Integration Intensity. We observe that the majority of GI physicians practice at multiple practices. Thus, there is a large number of partially integrated physicians who split their time between independent and integrated organizations. Although these physicians are categorized as integrated in our main analyses, they are exposed to integrated practice setting less degree than the physicians who exclusively work at one (or multiple) integrated practices. To differentiate these two cases, we used an ordered factor variable of integration, where physicians of partial integration (e.g., only a fraction of their affiliated practices integrate) have $INTEG_{jt} = 1$, and those of full integration (e.g., all of their practices integrate) have $INTEG_{jt} = 2$. Table 3.7 shows the result of our analysis: the intensity of integration is positively correlated with the integration effect. Those who were integrated with greater intensity experience greater reduction in deep sedation and increases in bleeding and minor complications.

ASC Affiliation. Studies report meaningful differences in the patient outcomes for the procedures performed at ASCs versus physician offices, where ASCs tend to be larger also yield better quality outcomes for certain surgical procedures (Strope *et al.*, 2009). Thus, whether the integrating physicians are ASC-based or office-based can affect the integration effect. To examine this, we calculated the proportion of ASC based claims out of all outpatient claims and found that about 56.0% of the physicians had submitted more than 25% of their claims under ASC. We defined these physicians as ASC-based and included a binary indicator of ASC-based as an interaction term with the DID coefficient. Table 3.7 shows that ASC-based physicians experience less reduction in deep sedation use compared to office-based physicians. They also had less increase in minor complications, but these effects were not significant. Thus, there is a potential, although weak, evidence that ASC-based physicians experience less negative consequences of integration.

Reasons for Integration. We also note that there are at least two ways physicians can integrate: acquisition of their practices, and direct hire to a different, already integrated, practice. We defined the physicians who are integrated through integration if the majority (≥ 50) of the other physicians affiliated with the same organization (based on the organization NPI in the facility claims) also integrate in the same year. We defined the physicians as integrated through hire if the majority of their fellow physicians are already integrated. We included a binary indicator of whether the physician was integrated through a hire (vs. acquisition) as an interaction term with the DID coefficient. Table 3.7 shows that the physicians who integrate through acquisition experience a greater adverse effect than those who integrate through the hire.

Competition. Integration can contribute to a reduction in the competition in the market. Given the fixed Medicare price, this might incentivize the practices to reduce their efforts on improving quality and/or efficiency. Thus, we examined whether the competition in the market plays a role by including an interaction term between the competition (measured as hospital HHI) and *POST* in our main model. Interestingly, patients in low competition areas (e.g., high HHI) experienced less reduction in anesthesia than patients in high competition

areas (e.g., low HHI). However, variation in the degree of competition did not play a role in other outcomes.

3.8 Robustness Checks

To test the spread and validity of our results, we performed various robustness checks. As we describe next, these include examining the impact of integration on an additional outcome measure (patient experience), testing changes in patient risk composition, investigating physicians' coding and gaming behavior, changing the measure of integration as well as some other variable definitions, and performing a variety of additional statistical tests such as using Inverse Probability Weighting (IPW) to adjust for potential baseline differences in physicians. Overall, our results give us confidence that our results are fairly robust and are not sensitive to our assumptions and model specifications.

3.8.1 Effect on Patient Experience

In addition to the outcome variables discussed earlier, we examined the effect of integration on patient experience to test whether integration affects other dimensions of care. To measure patient experience, we used the Outpatient Consumer Assessment of Healthcare Providers and Systems (OAS CAHPS) survey data for HOPDs. The OAS CAHPS survey asks patients who made outpatient visits about various aspects of their experience, including cleanliness, communication, and overall recommendation (Centers for Medicare & Medicaid Services, 2016b). Because these surveys were initiated in 2018, we were not able to perform a DID analysis similar to what we have done for our main analyses. Thus, we conduct a matched regression of physicians who do and do not integrate and compare their outcomes. Among the HOPD-affiliated physicians, we performed matching based on their patients' age, demographics, and clinical risk, and physician's own characteristics such as gender and geographic regions. We ran a linear regression on the matched sample, stratified by physicians in partial and full integration, for the four patient experience outcomes. Overall, we find that integration is associated with a significant increase in all aspects of

patient experience except for the cleanliness for both types of integrations. The association between vertical integration and an indication of better patient experience may suggest that despite a reduction in the actual care processes, there can be a potential positive impact of vertical integration—patients may perceive integration as having improved quality. Yet, the results need to be interpreted with caution, as we were not able to adjust for unobserved heterogeneity among physicians.

3.8.2 Changes in Patient Risk Composition

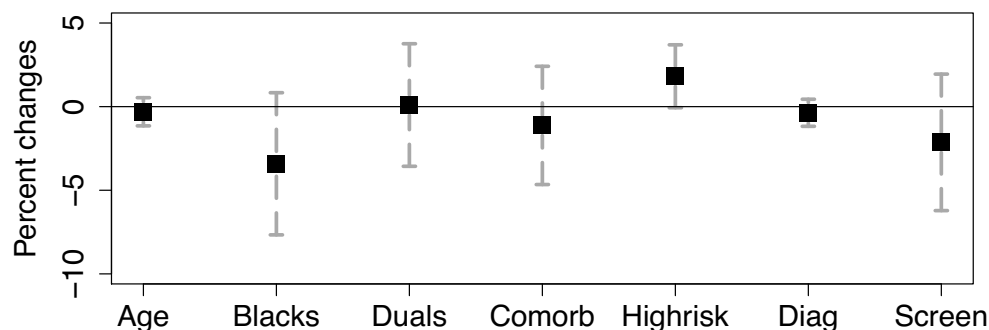
One important assumption of our identification strategy is that the changes in the quality of GI physicians who alter their integration are not due to the changes in patient characteristics. Given that distance is one of the primary factors for patients' physician choice, and given that the physicians' physical locations do not change after integration, such concerns on patient selection before and after integration can be mitigated. Yet, it is possible that the changes in the ownership status of a physician can result in attracting a different set of patients as a result of changes in insurance, perceived quality, or actual quality.

Patient Composition. We first examined whether there are any changes in observed patient composition when GI physicians integrate. For each physician, we measured the proportion of patients with certain characteristics (demographic, clinical) per year. We then tested whether the patient composition changed following integration using a physician-level DID model:

$$Y_{jt} = \alpha \text{INTEG}_{jt} + \beta \mathbf{X}_{jt} + \gamma \mathbf{Z}_{jt} + \text{PHYSICIAN}_j + \text{YEAR}_t + \epsilon_{jt}, \quad (3.4)$$

where Y is a measure of patient composition (e.g., out of all patients of physician j , number of patients with specific characteristics) for physician-year. Other variables are defined as the same as our main model. Figure 3.7 shows that integration was not associated with any changes in demographic or clinical composition, including colonoscopy risks. If the unobserved risk also changed in the same direction as the observed risk after integration, the result suggests that the changes in patient composition would not have contributed to

Figure 3.7: Average Effect of Integration on Patient Composition, Scaled into Percent Changes



Note. All effects are scaled as changes in percentages. Each dot indicates the size of the DID coefficient. Grey lines depict the 95% confidence intervals around the coefficient of the DID variable. Standard errors are robust and clustered at the physician and the year levels.

Table 3.8: Robustness Checks Results

	(2) Patient risk		(3) Coding behavior	(4) Measure integration		(5) IPW		
	Stratify, high-risk	Stratify, low-risk	Primary condition only	Threshold, 5%	Threshold, 15%	Threshold, 20%	Threshold, 25%	
Deep sedation	-0.035*** (0.0057)	-0.038*** (0.0056)	-0.037*** (0.008)	-0.043*** (0.0084)	-0.041*** (0.0069)	-0.040*** (0.0068)	-0.036*** (0.0065)	-0.037*** (0.0080)
Bleeding	0.0050* (0.0018)	0.0061* (0.0019)	0.007** (0.002)	0.00030 (0.0021)	0.0055** (0.0014)	0.0038* (0.0015)	0.0043† (0.0020)	0.0081** (0.0020)
Minor complications	0.0052* (0.0016)	0.0051** (0.0013)	0.005** (0.001)	0.0054** (0.0011)	0.0069** (0.0018)	0.0052** (0.0014)	0.0036 (0.0021)	0.0052** (0.0010)
Spending	127.8*** (11.9)	132.4*** (11.9)	59.1*** (4.7)	97.8*** (11.5)	152.5*** (13.0)	157.5*** (13.8)	146.3*** (14.0)	133.84*** (11.47)

Note. Comp. indicates complications. † indicates marginally significant at p -value < 0.10 . We only present the results for the outcomes that had any significant changes in the main analyses. Standard errors in parentheses are robust and clustered at the physician and the year levels.

the increases in adverse outcomes.

Stratification. We further stratify the sample by patient risk. We estimated the effect among the patients with high vs. low CRC risk since high-risk colonoscopies have a greater chance of adverse outcomes. Table 3.8 shows that the effect was consistent among both groups for all outcomes. The results further suggest that the chance that an unobserved heterogeneity might be driving our results is minimal.

3.8.3 Coding and Gaming Behavior

Another important threat to our identification strategy is that physicians may change the way they code in a way that does not reflect the true changes in quality, depending on the administrative infrastructure of the newly integrated system. To test this assumption, we formally examined the following. First, we examined the total number of related conditions that physicians submit. We then tested the primary conditions to compare the intensity. Finally, we made use of the primary condition only to re-examine the integration effect. Our results in Table 3.8 suggest that there is no evidence that the physicians are increasing the coding intensity after integration. Further, we note that there is no good reason for physicians to change their coding behavior for the purpose of gaming because the measures we are most worried about (e.g., polyp detection) are not used for payment or direct quality measures tied to incentives during our observation period. Our data also shows that the variation in polypectomy rates across different physicians was very low, and the rates do not change after integration, suggesting that the effect is negligible.

3.8.4 Measuring Integration

In our main analysis, we made use of specific threshold values (e.g., 10%) to define ingratiation. To test the sensitivity of our findings to these values, we made use of other thresholds: 5%, 15%, 20%, and 25%. Table 3.8 shows that varying the threshold from 5% to 20% for integration does not affect our main findings. However, setting the threshold at greater than 25% resulted in weakening the significance of integration effect on bleeding and minor complications. We argue that because of the nature of GI physicians' practice patterns where the majority of physicians practice at multiple practices, limiting the physicians to a small subset of physicians who have higher integration threshold may miss the average effect.

3.8.5 Other Statistical Tests

Bonferroni Correction. Some of our outcome measures are compound outcomes that are potentially correlated with each other. For example, perforation can result in increased

bleeding. Thus, there is a risk of having positive results by chance when we conduct multiple comparisons on different outcomes. We correct for the potential correlation between different outcomes using Bonferroni test simply by dividing the p-value of 0.05 into five, the total number of major and minor complications that are potentially correlated.¹³ The results of the adjusted outcomes are still significant or marginally significant for all our quality outcomes.

Inverse Probability Weighting (IPW). We adjusted for the baseline differences in physicians, especially the pre-integration characteristics, using IPW. In IPW, each observation is weighted to balance the representation of subgroups within the full data set. The weight is the reciprocal of the predicted probability of the observation being in the group that was observed for each observation. We used a logistic model with pre-integration characteristics of patient, physician, and area to estimate the propensity score that each physician will integrate. Using the propensity score, we weighed the entire study sample by inverse probability of treatment weights defined as $z/e + (1 - z)/(1 - e)$, where z denotes treatment status and e denotes the estimated propensity score. We truncated the propensity at the 99th percentile to address the unstable weights or for providers with a very low probability of receiving the treatment. To assess whether the propensity score model has been adequately specified, we examined the distribution of baseline covariates between the treatment groups in the balanced sample by checking the summary statistics and t-statistics. Using the weighted sample, we estimated the average treatment effect from our main model (Table 3.8).

3.9 Conclusion and Policy Recommendation

We examine the impact of vertical integration on the quality of care (in terms of both process and patient outcomes), operational efficiency, and spending. Table 3.9 summarizes our

¹³We also applied several approaches to adjust for p-values, including Hochberg, Hommel, Holm, and Benjamini Hochberg and Yekutieli procedures. These approaches vary in the conservativeness of how many p-values they adjust. The Bonferroni test is the most conservative. (Benjamini and Hochberg, 1995)

Table 3.9: *Summary of the Impact of Vertical Integration*

	Outcomes	DID effect direction		Outcomes	DID effect direction
Process-related quality	Polypectomy	—	Operational efficiency	Time to follow-up	—
	Incomplete colonoscopy	—		Time to surgery	—
	Deep sedation	↓		Any services	—
Outcome-related quality	Perforation	—		Colonoscopies	↑
	Bleeding	↑		Procedure types	↑
	Infection	—		Patients seen	—
	Minor complication	↑	Spending	Provider	—
	Interval CRC	—		Facility	↑
				Total	↑

findings, where favorable changes are indicated in upward arrows, and unfavorable changes are presented in downward arrows. Overall, we find that unfavorable changes caused by vertical integration are primarily driven by a reduction in some value-added care processes (e.g., deep sedation), which leads to an increase in some adverse patient outcomes. We further show that this reduction in the value-added care process is, in turn, driven by the inherent financial incentive structure and subsequent resource constraints. Although there is a sign of an increase in some aspects of the operational efficiency, the effect is small. We also observe that spending increases post-integration, which is mainly due to the greater price that Medicare pays for integrated practices. Put together, our results suggest that vertical integration has some negative consequences on the various dimension of healthcare delivery, and hence, requires careful consideration by policymakers.

Our study has a number of limitations. First, it focuses on a specific specialty (gastroenterology) and population (FFS Medicare). The findings may differ for other specialty practices that have different quality measures, provider roles, and characteristics of the disease. The younger population or the Medicare managed care population may also have different responses than the FFS population we studied. Second, there are various limitations from the nature of the data we used as well as our empirical strategy of DID. Although we discuss why the concerns for biases from both data and methods are mitigated (Section 6) and conduct various robustness checks (Section 8), it does not eliminate all of the threats to internal validity. For example, for measuring quality outcomes, there are

various challenges with identifying variations in coding and billing patterns that one needs to consider. Similarly, measuring integration from our data is imperfect and subject to error. There are multiple forms of integration, and using claims data to infer them is inherently challenging. Having a rich dataset that can identify various nuances of integrated entities apart in the future would be helpful.

To assist policymakers, we estimate the reasonable amount of an incentive if it could alter physicians' behavior by performing a simple cost-effectiveness analysis. Suppose we can prevent the reduction of deep sedation use after integration through increasing the incentives for providing deep sedation. If such incentives can maintain the same level of deep sedation use after physicians' integration, this would result in 8.2 % increase in deep sedation * 2,226,850 colonoscopy patients/year = 182,602 additional colonoscopies with deep sedations among FFS Medicare population. Thus, it results in 182,602 colonoscopies * (30.02 – 2.2)% reduced chance of bleeding = 51,129 fewer bleedings among FFS Medicare population who go through colonoscopy each year. Based on the estimated utility loss of 0.0384 quality-adjusted life years (QALY) for bleeding after colonoscopy from literature (Goede *et al.*, 2017), this change will gain 51,129 colonoscopies * 0.0384 QALY gains = 1,963 QALY gained in a given year. Given that spending up to \$22,289 for a unit of QALY is considered cost-effective in developed countries (Bertram *et al.*, 2016), our QALY translates into monetary value of 1,963 QALY*\$22,289 /QALY= \$43,753,307, which means \$43,753,307/182,602 colonoscopies = \$239.6 additional payment for each colonoscopy with deep-sedation. In other words, paying up to \$239.6 additional incentives per colonoscopy to prevent the reduction of deep sedation use among integrating physicians can still be cost-effective. This estimate critically depends on the assumption that providing financial incentives can alter physicians' or practices' behavior in a way that more anesthesiologists would be available to provide deep sedation for colonoscopies. Given that assumption, the magnitude of incentive is over twice the current payment rate for deep sedation, which we view could be a reasonable size for inducing physicians' behavioral changes.

Our results shed light on the recent policy discussion around healthcare delivery and

financing reform. One of the rationales for these reforms is to address the fragmented healthcare system by encouraging coordination among care providers. Yet, there have been concerns that these policies may increase vertical integration without clear evidence on positive impact. For example, Medicare Access and CHIP Reauthorization Act (MACRA) of 2015, which has revised the physician payment methods, creates a potential pathway for physicians to earn substantial bonuses for participating in alternative payment models favoring large organizations. Other provider payment reforms such as bundled payment programs or the Federal 340B drug discount program all provide direct or indirect incentives for consolidations among providers in different production segment. Our results indicate that incentives that can accelerate vertical integration can negatively impact some dimensions of performance in the healthcare sector, and hence, need to be cautiously evaluated by policymakers. At the same time, because such negative impacts are driven by resource constraints, changing the financial incentive for the providers to address the constraints can potentially mitigate the unintended consequences of integration.

References

- ADIDA, E., MAMANI, H. and NASSIRI, S. (2016). Bundled payment vs. fee-for-service: Impact of payment scheme on performance. *Management Science*, **63** (5), 1606–1624.
- ALPERT, A., HSI, H. and JACOBSON, M. (2017). Evaluating the role of payment policy in driving vertical integration in the oncology market. *Health Affairs*, **36** (4), 680–688.
- ASWANI, A., SHEN, Z.-J. M. and SIDDIQ, A. (2019). Data-driven incentive design in the medicare shared savings program. *Operations Research*.
- ATA, B., KILLALY, B. L., OLSEN, T. L. and PARKER, R. P. (2013). On hospice operations under medicare reimbursement policies. *Management Science*, **59** (5), 1027–1044.
- AVDIC, D. (2016). Improving efficiency or impairing access? health care consolidation and quality of care: Evidence from emergency hospital closures in sweden. *Journal of health economics*, **48**, 44–60.
- BAICKER, K., CHERNEW, M. E. and ROBBINS, J. A. (2013). The spillover effects of Medicare managed care: Medicare Advantage and hospital utilization. *Journal of Health Economics*, **32** (6), 1289–1300.
- BAICKER, K. and LEVY, H. (2013). Coordination versus competition in health care reform. *The New England journal of medicine*, **369** (9), 789.
- BAICKER, K. and STAIGER, D. O. (2005). Fiscal shenanigans, targeted federal health care funds, and patient mortality. *The quarterly journal of economics*, **120** (1), 345–386.
- BAKER, L. C., BUNDORF, M. K., DEVLIN, A. M. and KESSLER, D. P. (2018). Hospital ownership of physicians: hospital versus physician perspectives. *Medical Care Research and Review*, **75** (1), 88–99.
- BAKER, L. C., BUNDORF, M. K. and KESSLER, D. P. (2014). Vertical integration: hospital ownership of physician practices is associated with higher prices and spending. *Health Affairs*, **33** (5), 756–763.
- BARNES, B. G. and HARP, N. L. (2018). The us medicare disproportionate share hospital program and capacity planning. *Journal of Accounting and Public Policy*, **37** (4), 335–351.
- BARNETT, M. L., KEATING, N. L., CHRISTAKIS, N. A., O'MALLEY, A. J. and LANDON, B. E. (2012). Reasons for choice of referral physician among primary care and specialist physicians. *Journal of general internal medicine*, **27** (5), 506–512.

- BARON, R. M. and KENNY, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, **51** (6), 1173.
- BARTEL, A. P., CHAN, C. W. and KIM, S.-H. H. (2014). Should hospitals keep their patients longer? the role of inpatient care in reducing post-discharge mortality. *National Bureau of Economic Research*.
- BASTANI, H., BAYATI, M., BRAVERMAN, M., GUMMADI, R. and JOHARI, R. (2016). Analysis of medicare pay-for-performance contracts. *Available at SSRN* 2839143.
- BATES, L. J., MUKHERJEE, K. and SANTERRE, R. E. (2006). Market structure and technical efficiency in the hospital services industry: A DEA approach. *Medical Care Research and Review*, **63** (4), 499–524.
- BATT, R. J. and TERWIESCH, C. (2012). Doctors under load: An empirical study of state-dependent service times in emergency care. *Working Paper*.
- BAXTER, N. N., WARREN, J. L., BARRETT, M. J., STUKEL, T. A. and DORIA-ROSE, V. P. (2012). Association between colonoscopy and colorectal cancer mortality in a us cohort according to site of cancer and colonoscopist specialty. *Journal of Clinical Oncology*, **30** (21), 2664.
- BAZZOLI, G. J., BREWSTER, L. R., LIU, G. and KUO, S. (2003). Does U.S. hospital capacity need to be expanded? *Health Affairs*, **22** (6), 40–54.
- BAZZOLI, G. J., FAREED, N. and WATERS, T. M. (2014). Hospital financial performance in the recent recession and implications for institutions that remain financially weak. *Health Affairs*, **33** (5), 739–745.
- BENBASSAT, J. and TARAGIN, M. (2000). Hospital readmissions as a measure of quality of health care: advantages and limitations. *Archives of Internal Medicine*, **160** (8), 1074–1081.
- BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, **57** (1), 289–300.
- BERTRAM, M. Y., LAUER, J. A., DE JONCHEERE, K., EDEJER, T., HUTUBESSY, R., KIENY, M.-P. and HILL, S. R. (2016). Cost–effectiveness thresholds: pros and cons. *Bulletin of the World Health Organization*, **94** (12), 925.
- BERTSIMAS, D., FARIAS, V. F. and TRICHAKIS, N. (2013). Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research*, **61** (1), 73–87.
- BIRKMEYER, J. D., SIEWERS, A. E., FINLAYSON, E. V., STUKEL, T. A., LUCAS, F. L., BATISTA, I., WELCH, H. G. and WENNBURG, D. E. (2002). Hospital volume and surgical mortality in the United States. *New England Journal of Medicine*, **346** (15), 1128–1137.
- BLAVIN, F. (2016). Association between the 2014 Medicaid expansion and us hospital finances. *JAMA*, **316** (14), 1475–1483.

- BRADLEY, E. H., HERRIN, J., ELBEL, B., McNAMARA, R. L., MAGID, D. J., NALLAMOTHU, B. K., WANG, Y., NORMAND, S.-L. T., SPERTUS, J. A. and KRUMHOLZ, H. M. (2006). Hospital quality for acute myocardial infarction: correlation among process measures and relationship with short-term mortality. *Jama*, **296** (1), 72–78.
- BRANAS, C. C., MACKENZIE, E. J. and REVELLE, C. S. (2000). A trauma resource allocation model for ambulances and hospitals. *Health Services Research*, **35** (2), 489.
- BRAY, R. L., SERPA, J. C. and COLAK, A. (2019). Supply chain proximity and product quality. *Management Science*, **65** (9), 4079–4099.
- BROYDEN, C. G. (1970). The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA Journal of Applied Mathematics*, **6** (1), 76–90.
- BUCHMUELLER, T. C., JACOBSON, M. and WOLD, C. (2006). How far to the hospital?: The effect of hospital closures on access to care. *Journal of Health Economics*, **25** (4), 740–761.
- BUENO, H., ROSS, J. S., WANG, Y., CHEN, J., VIDÁN, M. T., NORMAND, S.-L. T., CURTIS, J. P., DRYE, E. E., LICHTMAN, J. H., KEENAN, P. S. *et al.* (2010). Trends in length of stay and short-term outcomes among Medicare patients hospitalized for heart failure, 1993–2006. *JAMA*, **303** (21), 2141–2147.
- BURNS, L. R., GOLDSMITH, J. C. and SEN, A. (2014). Horizontal and vertical integration of physicians: a tale of two tails. In *Annual review of health care management: Revisiting the evolution of health systems organization*, Emerald Group Publishing Limited, pp. 39–117.
- CACHON, G. P. and ZHANG, F. (2007). Obtaining fast service in a queueing system via performance-based allocation of demand. *Management Science*, **53** (3), 408–420.
- CAPPS, C., DRANOVE, D. and LINDROOTH, R. C. (2010). Hospital closure and economic efficiency. *Journal of Health Economics*, **29** (1), 87–109.
- CAPPS, C., DRANOVE, D. and ODY, C. (2018). The effect of hospital acquisitions of physician practices on prices and spending. *Journal of health economics*, **59**, 139–152.
- CARLIN, C. S., DOWD, B. and FELDMAN, R. (2015). Changes in quality of health care delivery after vertical integration. *Health services research*, **50** (4), 1043–1068.
- CARROLL, C. (2019). Impeding access or promoting efficiency? effects of rural hospital closure on the cost and quality of care. *Working Paper*.
- CASEY, M. M., MOSCOVICE, I., HOLMES, G. M., PINK, G. H. and HUNG, P. (2015). Minimum-distance requirements could harm high-performing critical-access hospitals and rural communities. *Health Affairs*, **34** (4), 627–635.
- CENTER FOR MEDICAID AND STATE OPERATIONS/SURVEY AND CERTIFICATION GROUP (2007). Critical access hospitals (cahs): Distance from other providers and relocation of cahs with a necessary provider designation.

- CENTERS FOR MEDICARE & MEDICAID SERVICES (2016a). Medicare provider utilization and payment data: physician and other supplier. *Available at* (Accessed November 17, 2016) <https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-provider-charge-data/physician-and-other-supplier.html> View in Article.
- CENTERS FOR MEDICARE & MEDICAID SERVICES (2016b). Outpatient and ambulatory surgery cahps (oas cahps).
- CENTERS FOR MEDICARE & MEDICAID SERVICES (2018). National health expenditure data tables.
- CHAN, C. W., FARIAS, V. F. and ESCOBAR, G. J. (2016). The impact of delays on service times in the intensive care unit. *Management Science*, **63** (7), 2049–2072.
- CHANDRA, A. and STAIGER, D. O. (2007). Productivity spillovers in health care: evidence from the treatment of heart attacks. *Journal of Political Economy*, **115** (1), 103–140.
- CHEN, C. and SAVVA, N. (2018). Unintended consequences of hospital regulation: The case of the hospital readmissions reduction program. *Available at SSRN 3236983*.
- CHEN, S. and REX, D. (2004). Registered nurse-administered propofol sedation for endoscopy. *Alimentary pharmacology & therapeutics*, **19** (2), 147–155.
- CHERNEW, M., SCANLON, D. and HAYWARD, R. (1998). Insurance type and choice of hospital for coronary artery bypass graft surgery. *Health Services Research*, **33** (3 Pt 1), 447.
- CHRONIC CONDITION DATA WAREHOUSE (2014). Chronic condition data warehouse.
- COLE, E. S., WALKER, D., MORA, A. and DIANA, M. L. (2014). Identifying hospitals that may be at most financial risk from medicaid disproportionate-share hospital payment cuts. *Health Affairs*, **33** (11), 2025–2033.
- COOPER, G. S., XU, F., BARNHOLTZ SLOAN, J. S., SCHLUCHTER, M. D. and KOROUKIAN, S. M. (2012). Prevalence and predictors of interval colorectal cancers in medicare beneficiaries. *Cancer*, **118** (12), 3044–3052.
- CORLEY, D. A., JENSEN, C. D., MARKS, A. R., ZHAO, W. K., LEE, J. K., DOUBENI, C. A., ZAUBER, A. G., DE BOER, J., FIREMAN, B. H., SCHOTTINGER, J. E. *et al.* (2014). Adenoma detection rate and risk of colorectal cancer and death. *New England Journal of Medicine*, **370** (14), 1298–1306.
- CUBANSKI, J., NEUMAN, T. and DAMICO, A. (2016). Medicare’s role for people under age 65 with disabilities. *Kaiser Family Foundation Issue Brief, August*.
- CUELLAR, A. E. and GERTLER, P. J. (2006). Strategic integration of hospitals and physicians. *Journal of health economics*, **25** (1), 1–28.
- DE CHAISEMARTIN, C. and D’HAUTFÈUILLE, X. (2017). Fuzzy differences-in-differences. *The Review of Economic Studies*, **85** (2), 999–1028.

- DEBO, L. G., TOKTAY, L. B. and VAN WASSENHOVE, L. N. (2008). Queuing for expert services. *Management Science*, **54** (8), 1497–1512.
- DEO, S., IRAVANI, S., JIANG, T., SMILOWITZ, K. and SAMUELSON, S. (2013). Improving health outcomes through better capacity allocation in a community-based chronic care model. *Operations Research*, **61** (6), 1277–1294.
- DESAI, S. and MCWILLIAMS, J. M. (2018). Consequences of the 340b drug pricing program. *New England Journal of Medicine*, **378** (6), 539–548.
- DONG, J., YOM-TOV, E. and YOM-TOV, G. B. (2018). The impact of delay announcements on hospital network coordination and waiting times. *Management Science*, **65** (5), 1969–1994.
- DUFLO, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, **91** (4), 795–813.
- EARNSHAW, S. R., RICHTER, A., SORENSEN, S. W., HOERGER, T. J., HICKS, K. A., ENGELGAU, M., THOMPSON, T., VENKAT NARAYAN, K., WILLIAMSON, D. F., GREGG, E. *et al.* (2002). Optimal allocation of resources across four interventions for type 2 diabetes. *Medical Decision Making*, **22** (1_suppl), 80–91.
- ELIXHAUSER, A., STEINER, C., HARRIS, D. R. and COFFEY, R. M. (1998). Comorbidity measures for use with administrative data. *Medical Care*, **36** (1), 8–27.
- EMMERT, M. and SCHLESINGER, M. (2017). Hospital quality reporting in the united states: does report card design and incorporation of patient narrative comments affect hospital choice? *Health services research*, **52** (3), 933–958.
- FISCELLA, K., FRANKS, P., GOLD, M. R. and CLANCY, C. M. (2000). Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. *JAMA*, **283** (19), 2579–2584.
- FISHMAN, L. E. and BENTLEY, J. D. (1997). The evolution of support for safety-net hospitals. *Health Affairs*, **16** (4), 30–47.
- FONAROW, G. C., SMITH, E. E., SAVER, J. L., REEVES, M. J., HERNANDEZ, A. F., PETERSON, E. D., SACCO, R. L. and SCHWAMM, L. H. (2011). Improving door-to-needle times in acute ischemic stroke: the design and rationale for the American Heart Association/American Stroke Association’s Target: stroke initiative. *Stroke*, **42** (10), 2983–2989.
- FORCE, U. P. S. T. *et al.* (2008). Screening for colorectal cancer: Us preventive services task force recommendation statement. *Annals of internal medicine*, **149** (9), 627.
- FORLINES, G. L. (2017). Trends and drivers of physician-hospital integration: The role of medicare reimbursement. *Working Paper*.
- FREEDMAN, S., NIKPAY, S., CARROLL, A. and SIMON, K. (2017). Changes in inpatient payer-mix and hospitalizations following medicaid expansion: Evidence from all-capture hospital discharge data. *PloS one*, **12** (9), e0183616.

- FREEMAN, M., SAVVA, N. and SCHOLTES, S. (2018). Economies of scale and scope in hospitals: An empirical study of volume spillovers. *London Business School Working Paper*.
- FRIEDMAN, A. B., OWEN, D. D. and PEREZ, V. E. (2016). Trends in hospital ED closures nationwide and across Medicaid expansion, 2006-2013. *The American Journal of Emergency Medicine*, **34** (7), 1262–1264.
- GAYNOR, M. (2014). Competition policy in health care markets: Navigating the enforcement and policy maze. *Health Affairs*, **33** (6), 1088–1093.
- GAYNOR, M., MORENO-SERRA, R. and PROPPER, C. (2013). Death by market power: reform, competition, and patient outcomes in the national health service. *American Economic Journal: Economic Policy*, **5** (4), 134–66.
- GILLIES, R. R., SHORTELL, S. M., ANDERSON, D. A., MITCHELL, J. B. and MORGAN, K. L. (1993). Conceptualizing and measuring integration: findings from the health systems integration study. *Journal of Healthcare Management*, **38** (4), 467.
- GOEDE, S. L., RABENECK, L., VAN BALLEGOOIJEN, M., ZAUBER, A. G., PASZAT, L. F., HOCH, J. S., YONG, J. H., KROEP, S., TINMOUTH, J. and LANSDORP-VOGELAAR, I. (2017). Harms, benefits and costs of fecal immunochemical testing versus guaiac fecal occult blood testing for colorectal cancer screening. *PloS one*, **12** (3).
- GOLDSTEIN, E., FARQUHAR, M., CROFTON, C., DARBY, C. and GARFINKEL, S. (2005). Measuring hospital care from the patients' perspective: An overview of the CAHPS hospital survey development process. *Health Services Research*, **40** (6p2), 1977–1995.
- GOWRISANKARAN, G., LUCARELLI, C., SCHMIDT-DENGLER, P. and TOWN, R. (2018). Can amputation save the hospital? the impact of the medicare rural flexibility program on demand and welfare. *Journal of health economics*, **58**, 110–122.
- GRAVES, J. A. (2012). Medicaid expansion opt-outs and uncompensated care. *New England Journal of Medicine*, **367** (25), 2365–2367.
- GREEN, L. V. and NGUYEN, V. (2001). Strategies for cutting hospital beds: the impact on patient service. *Health Services Research*, **36** (2), 421.
- GUJRAL, K. and BASU, A. (2019). Impact of rural and urban hospital closures on inpatient mortality. *NBER Working Paper*.
- GUPTA, D. and MEHROTRA, M. (2015). Bundled payments for healthcare services: Proposer selection and information sharing. *Operations research*, **63** (4), 772–788.
- GUTACKER, N., SICILIANI, L., MOSCELLI, G. and GRAVELLE, H. (2016). Choice of hospital: which type of quality matters? *Journal of health economics*, **50**, 230–246.
- HAAS, S., GAWANDE, A. and REYNOLDS, M. E. (2018). The risks to patient safety from health system expansions. *JAMA*, **319** (17), 1765–1766.
- HART, L. G., LARSON, E. H. and LISHNER, D. M. (2005). Rural definitions for health policy and research. *American Journal of Public Health*, **95** (7), 1149–1155.

- HECKMAN, J. J., HONORE, B. E. *et al.* (1990). The empirical content of the roy model. *Econometrica*, **58** (5), 1121–1149.
- HEESE, H. S. (2007). Inventory record inaccuracy, double marginalization, and rfid adoption. *Production and Operations Management*, **16** (5), 542–553.
- HODGSON, A., ROBACK, P., HARTMAN, A., KELLY, E. and LI, Y. (2015). The financial impact of hospital closures on surrounding hospitals. *Journal of Hospital Administration*, **4** (3), 25.
- HOPP, W. J., IRAVANI, S. M. and YUEN, G. Y. (2007). Operations systems with discretionary task completion. *Management Science*, **53** (1), 61–77.
- HSIA, R. Y., KANZARIA, H. K., SREBOTNJAK, T., MASELLI, J., MCCULLOCH, C. and AUERBACH, A. D. (2012). Is emergency department closure resulting in increased distance to the nearest emergency department associated with increased inpatient mortality? *Annals of Emergency Medicine*, **60** (6), 707–715.
- HUDSON, S., HULL, P. and LIEBERSOHN, J. (2017). Interpreting instrumented difference-in-differences. *Metrics Note*.
- JHA, A. K., ORAV, E. J. and EPSTEIN, A. M. (2010). The effect of financial incentives on hospitals that serve poor patients. *Annals of internal medicine*, **153** (5), 299–306.
- JOSKOW, P. L. (1980). The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital. *The Bell Journal of Economics*, pp. 421–447.
- JOUINI, O., DALLERY, Y. and NAIT-ABDALLAH, R. (2008). Analysis of the impact of team-based organizations in call center management. *Management Science*, **54** (2), 400–414.
- JOYNT, K. E., CHATTERJEE, P., ORAV, E. J. and JHA, A. K. (2015). Hospital closures had no measurable impact on local hospitalization rates or mortality rates, 2003–11. *Health Affairs*, **34** (5), 765–772.
- JOYNT, K. E., ORAV, E. J. and JHA, A. K. (2011). Thirty-day readmission rates for medicare beneficiaries by race and site of care. *JAMA*, **305** (7), 675–681.
- KAMINSKI, M. F., REGULA, J., KRASZEWSKA, E., POLKOWSKI, M., WOJCIECHOWSKA, U., DIDKOWSKA, J., ZWIERKO, M., RUPINSKI, M., NOWACKI, M. P. and BUTRUK, E. (2010). Quality indicators for colonoscopy and the risk of interval cancer. *New England Journal of Medicine*, **362** (19), 1795–1803.
- KAUFMAN, B. G., THOMAS, S. R., RANDOLPH, R. K., PERRY, J. R., THOMPSON, K. W., HOLMES, G. M. and PINK, G. H. (2016). The rising rate of rural hospital closures. *The Journal of Rural Health*, **32** (1), 35–43.
- KC, D. S. and TERWIESCH, C. (2009). Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Science*, **55** (9), 1486–1498.
- KC, D. S. and TERWIESCH, C. (2012). An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing & Service Operations Management*, **14** (1), 50–65.

- KEELER, T. E. and YING, J. S. (1996). Hospital costs and excess bed capacity: A statistical analysis. *The Review of Economics and Statistics*, pp. 470–481.
- KHADEMI, A., SAURE, D. R., SCHAEFER, A. J., BRAITHWAITE, R. S. and ROBERTS, M. S. (2015). The price of nonabandonment: Hiv in resource-limited settings. *Manufacturing & Service Operations Management*, **17** (4), 554–570.
- KHIANI, V. S., SOULOS, P., GANCAYCO, J. and GROSS, C. P. (2012). Anesthesiologist involvement in screening colonoscopy: temporal trends and cost implications in the medicare population. *Clinical Gastroenterology and Hepatology*, **10** (1), 58–64.
- KOCH, T. G., WENDLING, B. W. and WILSON, N. E. (2017). How vertical integration affects the quantity and cost of care for medicare beneficiaries. *Journal of health economics*, **52**, 19–32.
- KOCHER, R. and SAHNI, N. R. (2011). Hospitals’ race to employ physicians—the logic behind a money-losing proposition. *New England Journal of Medicine*, **364** (19), 1790–1793.
- KONETZKA, R. T., STUART, E. A. and WERNER, R. M. (2018). The effect of integration of hospitals and post-acute care providers on medicare payment and patient outcomes. *Journal of health economics*, **61**, 244–258.
- LAMMERS, E. (2013). The effect of hospital–physician integration on health information technology adoption. *Health economics*, **22** (10), 1215–1229.
- LEVINSON, D. R. and GENERAL, I. (2013). Most critical access hospitals would not meet the location requirements if required to re-enroll in medicare.
- LI, P., SCHNEIDER, J. E. and WARD, M. M. (2007). Effect of critical access hospital conversion on patient safety. *Health services research*, **42** (6p1), 2089–2108.
- LI, X., LI, Y. and CAI, X. (2013). Double marginalization and coordination in the supply chain with uncertain supply. *European Journal of Operational Research*, **226** (2), 228–236.
- LINDROOTH, R. C., PERRAILLON, M. C., HARDY, R. Y. and TUNG, G. J. (2018). Understanding the relationship between Medicaid expansions and hospital closures. *Health Affairs*, **37** (1), 111–120.
- LINDROOTH, R. C., SASSO, A. L. and BAZZOLI, G. J. (2003). The effect of urban hospital closure on markets. *Journal of health economics*, **22** (5), 691–712.
- LIU, C., SREBOTNJAK, T. and HSIA, R. Y. (2014). California emergency department closures are associated with increased inpatient mortality at nearby hospitals. *Health Affairs*, **33** (8), 1323–1329.
- LIU, H., WAXMAN, D. A., MAIN, R. and MATTKE, S. (2012). Utilization of anesthesia services during outpatient endoscopies and colonoscopies and associated spending in 2003–2009. *Jama*, **307** (11), 1178–1184.
- LU, L. X. and LU, S. F. (2018). Distance, quality, or relationship? interhospital transfer of heart attack patients. *Production and Operations Management*, **27** (12), 2251–2269.

- MADISON, K. (2004). Hospital–physician affiliations and patient treatments, expenditures, and outcomes. *Health services research*, **39** (2), 257–278.
- MANARY, M. P., BOULDING, W., STAELIN, R. and GLICKMAN, S. W. (2013). The patient experience and health outcomes. *New England Journal of Medicine*, **368** (3), 201–203.
- MCGUIRE, T. G. (2000). Physician agency. In *Handbook of Health Economics*, vol. 1, Elsevier, pp. 461–536.
- McNAMARA, P. E. (1999). Welfare effects of rural hospital closures: a nested logit analysis of the demand for rural hospital services. *American Journal of Agricultural Economics*, **81** (3), 686–691.
- MEDICARE PAYMENT ADVISORY COMMISSION (US) (2017). Report to the congress, medicare payment policy. *Medicare Payment Advisory Commission*.
- MENG, L., BATT, R. and TERWIESCH, C. (2018). The impact of facility layout on service worker behavior: An empirical study of nurses in the emergency department. *Available at SSRN* 3143883.
- MERETOJA, A., STRBIAN, D., MUSTANOJA, S., TATLISUMAK, T., LINDSBERG, P. J. and KASTE, M. (2012). Reducing in-hospital delay to 20 minutes in stroke thrombolysis. *Neurology*, **79** (4), 306–313.
- MILLER, R. H. and LUFT, H. S. (1997). Does managed care lead to better or worse quality of care? *Health Affairs*, **16** (5), 7–25.
- MILLS, A. F., ARGON, N. T. and ZIYA, S. (2018). Dynamic distribution of patients to medical facilities in the aftermath of a disaster. *Operations Research*, **66** (3), 716–732.
- MOSCONE, F., TOSETTI, E. and VITTADINI, G. (2012). Social interaction in patients’ hospital choice: evidence from Italy. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **175** (2), 453–472.
- NAVIGANT (2019). Analysis shows one-in-five U.S. rural hospitals at high risk of closing unless financial situation improves.
- NEPRASH, H. T., CHERNEW, M. E., HICKS, A. L., GIBSON, T. and McWILLIAMS, J. M. (2015). Association of financial integration between physicians and hospitals with commercial health care prices. *JAMA internal medicine*, **175** (12), 1932–1939.
- NEUHAUSEN, K., DAVIS, A. C., NEEDLEMAN, J., BROOK, R. H., ZINGMOND, D. and ROBY, D. H. (2014). Disproportionate-share hospital payment reductions may threaten the financial stability of safety-net hospitals. *Health Affairs*, **33** (6), 988–996.
- NGUYEN, N. X. and SHEINGOLD, S. H. (2011). Indirect medical education and disproportionate share adjustments to medicare inpatient payment rates. *Medicare & medicaid research review*, **1** (4).
- NIKPAY, S. S., RICHARDS, M. R. and PENSON, D. (2018). Hospital-physician consolidation accelerated in the past decade in cardiology, oncology. *Health Affairs*, **37** (7), 1123–1127.

- OFFICE OF INSPECTOR GENERAL (2014). Medicare and beneficiaries could save billions if cms reduces hospital outpatient department payment rates for ambulatory surgical center-approved procedures to ambulatory surgical center payment rates. *Department of Health and Human Services Audit Report*, (A-05-12-00020).
- OLIVA, R. and STERMAN, J. D. (2001). Cutting corners and working overtime: Quality erosion in the service industry. *Management Science*, **47** (7), 894–914.
- PICKENS, G., KARACA, Z., CUTLER, E., DWORSKY, M., EIBNER, C., MOORE, B., GIBSON, T., IYER, S. and WONG, H. S. (2018). Changes in hospital inpatient utilization following health care reform. *Health Services Research*, **53** (4), 2446–2469.
- PINES, J. M., ZOCCHI, M., MOGHTADERI, A., BLACK, B., FARMER, S. A., HUFSTETLER, G., KLAUER, K. and PILGRIM, R. (2016). Medicaid expansion in 2014 did not increase emergency department use but did change insurance payer mix. *Health Affairs*, **35** (8), 1480–1486.
- POST, B., BUCHMUELLER, T. and RYAN, A. M. (2018). Vertical integration of hospitals and physicians: Economic theory and empirical evidence on spending and quality. *Medical Care Research and Review*, **75** (4), 399–433.
- QUANTIN, C., BENZENINE, E., HÄGI, M., AUVERLOT, B., ABRAHAMOWICZ, M., COTTENET, J., FOURNIER, E., BINQUET, C., COMPAIN, D., MONNET, E. *et al.* (2012). Estimation of national colorectal-cancer incidence using claims databases. *Journal of cancer epidemiology*, **2012**.
- RABENECK, L., PASZAT, L. F., HILSDEN, R. J., SASKIN, R., LEDDIN, D., GRUNFELD, E., WAI, E., GOLDWASSER, M., SUTRADHAR, R. and STUKEL, T. A. (2008). Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. *Gastroenterology*, **135** (6), 1899–1906.
- RAMANARAYANAN, S. (2008). Does practice make perfect: An empirical analysis of learning-by-doing in cardiac surgery. *SSRN Working Paper*.
- RAMDAS, K. and RANDALL, T. (2008). Does component sharing help or hurt reliability? an empirical study in the automotive industry. *Management Science*, **54** (5), 922–938.
- RAMSEY, S. D., MANDELSON, M. T., ETZIONI, R., HARRISON, R., SMITH, R. and TAPLIN, S. (2004). Can administrative data identify incident cases of colorectal cancer? a comparison of two health plans. *Health Services and Outcomes Research Methodology*, **5** (1), 27–37.
- RAVAL, D. and ROSENBAUM, T. (2018). *Why is Distance Important for Hospital Choice? Separating Home Bias from Transport Costs*. Tech. rep., Working Paper.
- REX, D. K., PETRINI, J. L., BARON, T. H., CHAK, A., COHEN, J., DEAL, S. E., HOFFMAN, B., JACOBSON, B. C., MERGENER, K., PETERSEN, B. T. *et al.* (2006). Quality indicators for colonoscopy. *The American journal of gastroenterology*, **101** (4), 873.
- ROMLEY, J. A. and GOLDMAN, D. P. (2011). How costly is hospital quality? a revealed-preference approach. *The Journal of Industrial Economics*, **59** (4), 578–608.

- ROSKO, M. D. and MUTTER, R. L. (2010). Inefficiency differences between critical access hospitals and prospectively paid rural hospitals. *Journal of Health Politics, Policy and Law*, **35** (1), 95–126.
- ROSKO, M. D. and MUTTER, R. L. (2014). The association of hospital cost-inefficiency with certificate-of-need regulation. *Medical Care Research and Review*, **71** (3), 280–298.
- ROY, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, **3** (2), 135–146.
- SAGHAFIAN, S. and HOPP, W. J. (2019). The role of quality transparency in healthcare: Challenges and potential solutions. *National Academy of Medicine*.
- SAGHAFIAN, S. and HOPP, W. J. (2020). Can public reporting cure healthcare? The role of quality transparency in improving patient–provider alignment. *Operations Research (forthcoming)*.
- SCOTT, K. W., ORAV, E. J., CUTLER, D. M. and JHA, A. K. (2017). Changes in hospital–physician affiliations in us hospitals and their effect on quality of care. *Annals of internal medicine*, **166** (1), 1–8.
- SHAPIRO, C. (2010). The 2010 horizontal merger guidelines: From hedgehog to fox in forty years. *Antitrust LJ*, **77**, 49.
- SHUNKO, M., NIEDERHOFF, J. and ROSOKHA, Y. (2017). Humans are not machines: The behavioral impact of queueing design on service time. *Management Science*, **64** (1), 453–473.
- SINGER, S. J., KERRISSEY, M., FRIEDBERG, M. and PHILLIPS, R. (2018). A comprehensive theory of integration. *Medical Care Research and Review*, p. 1077558718767000.
- SONG, L. and SAGHAFIAN, S. (2019a). Do hospital closures improve the efficiency and quality of other hospitals? *Available at SSRN 3318609*.
- SONG, L. D. and SAGHAFIAN, S. (2019b). The spillover effects of hospital closures on the efficiency and quality of other hospitals.
- SONG, Y., SKINNER, J., BYNUM, J., SUTHERLAND, J., WENNBERG, J. E. and FISHER, E. S. (2010). Regional variations in diagnostic practices. *New England Journal of Medicine*, **363** (1), 45–53.
- STAATS, B. R., DAI, H., HOFMANN, D. and MILKMAN, K. L. (2017). Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science*, **63** (5), 1563–1585.
- STOCK, J. H. and YOGO, M. (2002). Testing for weak instruments in linear IV regression. *NBER Working Paper*.
- STROPE, S. A., DAIGNAULT, S., HOLLINGSWORTH, J. M., ZE, Z., WEI, J. T. and HOLLENBECK, B. K. (2009). Physician ownership of ambulatory surgery centers and practice patterns for urological surgery: evidence from the state of florida. *Medical care*, **47** (4), 403.

- SU, X. and ZENIOS, S. A. (2005). Patient choice in kidney allocation: A sequential stochastic assignment model. *Operations research*, **53** (3), 443–455.
- TAHERI, P. A., BUTZ, D. A. and GREENFIELD, L. J. (2000). Length of stay has minimal impact on the cost of hospital admission. *Journal of the American College of Surgeons*, **191** (2), 123–130.
- TAI, W.-T. C., PORELL, F. W. and ADAMS, E. K. (2004). Hospital choice of rural medicare beneficiaries: patient, hospital attributes, and the patient–physician relationship. *Health Services Research*, **39** (6p1), 1903–1922.
- TAN, T. F. and NETESSINE, S. (2014). When does the devil make work? An empirical study of the impact of workload on worker productivity. *Management Science*, **60** (6), 1574–1593.
- TAN, T. F. and NETESSINE, S. (2019). When you work with a superman, will you also fly? an empirical study of the impact of coworkers on performance. *Management Science*, **65** (8), 3495–3517.
- THE CENTERS FOR MEDICARE & MEDICAID SERVICES (2017). Hospital value-based purchasing.
- THOMPSON, D. A., YARNOLD, P. R., WILLIAMS, D. R. and ADAMS, S. L. (1996). Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. *Annals of Emergency Medicine*, **28** (6), 657–665.
- TOURANGEAU, A. E., DORAN, D. M., HALL, L. M., O'BRIEN PALLAS, L., PRINGLE, D., TU, J. V. and CRANLEY, L. A. (2007). Impact of hospital nursing care on 30-day mortality for acute medical patients. *Journal of Advanced Nursing*, **57** (1), 32–44.
- UNITED STATES GOVERNMENT ACCOUNTABILITY OFFICE (2015). Increasing hospital-physician consolidation highlights need for payment reform. *Report to Congressional Requesters*, (GAO-16-189).
- WAGNER, A. (2016). Effect of physician-hospital financial integration on health outcomes and spending.
- WANG, J. and ZHOU, Y. (2018). Impact of queue configuration on service time: Evidence from a supermarket. *Management Science*, **64** (7), 3055–3075.
- WARREN, J. L., KLABUNDE, C. N., MARIOTTO, A. B., MEEKINS, A., TOPOR, M., BROWN, M. L. and RANSOHOFF, D. F. (2009). Adverse events after outpatient colonoscopy in the medicare population. *Annals of internal medicine*, **150** (12), 849–857.
- WEBER, E. (2014). Measuring welfare from ambulatory surgery centers: a spatial analysis of demand for healthcare facilities. *The Journal of Industrial Economics*, **62** (4), 591–631.
- WEISSMAN, J. S., ROTHSCHILD, J. M., BENDAVID, E., SPRIVULIS, P., COOK, E. F., EVANS, R. S., KAGANOVA, Y., BENDER, M., DAVID-KASDAN, J., HAUG, P. *et al.* (2007). Hospital workload and adverse events. *Medical Care*, **45** (5), 448–455.
- WENNBERG, J. E. (1996). *The Dartmouth Atlas of health care in the United States*. American Hospital Association.

- WERTHEIM, P. and LYNN, M. L. (1993). Development of a prediction model for hospital closure using financial accounting data. *Decision Sciences*, **24** (3), 529–546.
- ZHANG, D. J., GURVICH, I., VAN MIEGHEM, J. A., PARK, E., YOUNG, R. S. and WILLIAMS, M. V. (2016). Hospital readmissions reduction program: An economic and operational analysis. *Management Science*, **62** (11), 3351–3371.

Appendix A

Appendix to Chapter 1

A.1 Additional Methods

A.1.1 Independent Variables

Transfer patients: Transfer patients were identified as those with the source of inpatient admission code “transfer from a different facility,” “transfer from ER,” or “transfer from the same facility.” Admissions for rehabilitation were identified from the presence of ICD-9 codes indicating care involving the use of rehabilitation procedures: V570, V571, V5721, V573, V5781, V5789, V579, and 462.

Categories of diagnosis for inpatient visit: We divided the admissions into 15 clinical categories based on the primary diagnosis codes. These categories include: infections and parasitic diseases, neoplasms, endocrine, nutritional and metabolic diseases, and immunity disorders, diseases of the blood and blood-forming organs, mental disorders, diseases of the nervous system and sense organs, diseases of the circulatory system, diseases of the respiratory system, diseases of the digestive system, diseases of the genitourinary system, diseases of the skin and subcutaneous tissue, diseases of the musculoskeletal system and connective tissue, congenital anomalies, symptoms, signs, and ill-defined conditions, and injury and poisoning.

A.1.2 Cost Estimation for Two Policy Levers

For selective bailout policy, we calculated the bailout cost as the DSH savings that incur as a result of the closures by calculating the average DSH payment made to each hospital type in 2015. For selective elimination of speed-up policy, we estimated the cost of such interventions based on the cost of increased length of stay from previous studies (Bartel *et al.*, 2014; Taheri *et al.*, 2000). Specifically, the annual per hospital cost of eliminating the speed-up behavior was calculated as the product of the inflation-adjusted per-patient cost of the stay for an additional day (\$611 per patient in 2015, see Taheri *et al.* (2000)), the magnitude of the speed-up behavior, and the annual patient volume per hospital.

A.2 Tables and Figures

Table A.1: Summary of Hospital and Patient Characteristics, Matched Sample, (a) Hospital Characteristics Only (Top) and (b) Both Hospital and Patient Characteristics (Bottom)

	Hospitals (Treat)	Hospitals (Control)	P-value
<i>Hospital characteristics only</i>			
Total (n)	197	243	
Observation years	4,030	4,071	
Disproportionate Share Hospital (%)	83.2	81.3	0.024
Teaching hospitals (%)	43.4	45.1	0.14
Ownership-Nonprofit (%)	57.9	66.2	< 0.001
Ownership-Private (%)	20.9	21.0	
Ownership-Public (%)	21.1	12.8	
Rural (%)			0.65
Avg Star Rating	2.7	3.0	< 0.001
Avg Discharges	11,422.8	11,238.4	0.44
Avg Beds	235.2	227.3	0.077
Avg Efficiency	44.4	44.8	0.27
Avg Utilization	0.6	0.6	0.0027
Avg Patient Experience (% Good)	66.6	67.9	< 0.001
Avg HMO pen (%)	0.2	0.2	0.63
Avg HHI	1,241.6	1,482.0	< 0.001
Avg Pop (n)	1,355,771.0	1,198,569.0	0.0028
Avg Unemp (%)	9.3	8.8	< 0.001
Avg Poverty (%)	17.4	16.1	< 0.001
Avg Under 65 (%)	84.0	82.8	< 0.001
Avg Hosp (/10,000)	0.3	0.3	0.0022
<i>Hospital and patient characteristics</i>			
Total (n)	196	239	
Observation years	4,013	4,022	
Disproportionate Share Hospital (%)	83.1	81.9	0.15
Teaching hospitals (%)	43.2	41.6	0.16
Ownership-Nonprofit (%)	57.7	61.7	< 0.001
Ownership-Private (%)	21.0	23.4	
Ownership-Public (%)	21.2	14.9	
Rural (%)			0.92
Avg Star Rating	2.7	2.9	< 0.001
Avg Discharges	11,470.8	11,620.1	0.54
Avg Beds	236.1	230.8	0.24
Avg Efficiency	44.5	45.9	0.00044
Avg Utilization	0.6	0.6	< 0.001
Avg Patient Experience (% Good)	66.6	67.0	0.17
Avg HMO pen (%)	0.2	0.2	0.12
Avg HHI	1,243.2	1,451.4	< 0.001
Avg Pop (n)	1,358,006.0	1,390,784.0	0.56
Avg Unemp (%)	9.3	9.4	0.053
Avg Poverty (%)	17.4	17.2	0.06
Avg Under 65 (%)	84.0	83.4	< 0.001
Avg Hosp (/10,000)	0.3	0.3	0.34

Table A.2: *Summary Statistics of Hospital and Patient Variables*

	Mean	SD	Q1	Q3	Cor,vol	Cor,bed	Cor,effc	Cor,util		
Discharges (n)	9,724	10,035.50	2,689	13,540	1					
Beds (n)	196.70	185.74	71	263	0.95	1				
Operational efficiency	45.92	16.09	35.97	56.19	0.42	0.23	1			
Bed utilization (%)	0.54	0.20	0.40	0.69	0.62	0.49	0.80	1		
	Mean	SD	Q1	Q3	Cor,age	Cor,comb	Cor,chron	Cor,serv	Cor,read	Cor,mort
Age	78.98	8.11	72	85	1					
Comorbidity	3.37	2.90	1	5	0.05	1				
Chronic conditions	23.72	8.15	27	27	0.18	0.04	1			
Service duration (days)	4.81	4.12	2	6	0.05	0.19	0.001	1		
Readmission (%)	0.16	0.37	0	0	0.02	0.19	0.01	0.09	1	
Mortality (%)	0.06	0.24	0	0	0.11	0.08	0.02	0.13	0.09	1

Note. Q1 and Q3 indicate the first and the third quartile, respectively. Cor indicates correlation.

Table A.3: Full Difference-In-Differences Results

	<i>Dependent variable:</i>				<i>Dependent variable:</i>		
	Effic. (1)	Util. (2)	Pat.Exp (3)		Ser. Dur. (1)	Readm. (2)	Mort. (3)
age	−0.255 (0.320)	−0.003 (0.002)	−0.008 (0.127)	race2	0.237*** (0.035)	0.003* (0.001)	−0.006*** (0.001)
sex	−2.964 (2.523)	−0.070 (0.038)	0.344 (1.946)	race3	−0.047 (0.045)	−0.011** (0.004)	−0.004* (0.002)
race.x	−0.151 (5.870)	0.027 (0.058)	−4.427 (4.646)	race4	−0.224** (0.076)	−0.014*** (0.002)	−0.007* (0.003)
race.y	0.073 (6.017)	0.031 (0.072)	−5.581 (4.106)	race5	−0.284*** (0.060)	−0.006*** (0.002)	−0.011*** (0.002)
buyin_mo	2.343 (3.265)	0.004 (0.031)	−0.241 (1.933)	race6	0.075 (0.095)	0.001 (0.006)	0.004 (0.005)
type_adm.x	2.345 (15.735)	−0.152 (0.087)	−17.168** (6.131)	sex	0.006 (0.009)	−0.002* (0.001)	−0.012*** (0.001)
type_adm.y	0.126 (15.532)	−0.164* (0.084)	−15.799** (5.859)	age_grp2	0.235*** (0.014)	0.004* (0.001)	0.018*** (0.001)
type_adm	−1.625 (15.871)	−0.214** (0.082)	−14.175* (5.761)	age_grp3	0.265*** (0.031)	0.002 (0.002)	0.054*** (0.001)
prop_hmo	10.095* (4.539)	0.115* (0.053)	−1.629 (2.300)	ind_dual	0.147*** (0.038)	0.013*** (0.001)	0.001 (0.001)
hhi	−0.001 (0.002)	0.00001 (0.00001)	−0.002 (0.001)	type_adm1	1.230*** (0.051)	0.071*** (0.006)	0.003* (0.001)
pop	−0.00000* (0.00000)	−0.00000 (0.00000)	−0.00000** (0.00000)	type_adm2	1.401*** (0.066)	0.041*** (0.005)	−0.004** (0.001)
unemp	2.637 (1.689)	0.038*** (0.011)	(0.000) (0.000)	type_adm3	0.443*** (0.088)	0.041*** (0.006)	−0.032*** (0.002)
poverty	−3.090** (1.048)	−0.027* (0.011)	−2.134** (0.813)	type_adm4	0.918 (0.783)	0.002 (0.044)	−0.017 (0.025)
				type_adm5	1.711*** (0.138)	0.043*** (0.007)	−0.0001 (0.004)
				type_adm9	0.616** (0.222)	0.048*** (0.010)	−0.002 (0.011)
				crec	−0.160* (0.081)	−0.001 (0.006)	−0.005 (0.004)
				esrd_indY	0.406*** (0.042)	0.045*** (0.002)	0.002 (0.002)
				tot_cc	−0.004*** (0.001)	−0.00002 (0.0001)	0.0001* (0.00003)
				tot_comorb	0.245*** (0.005)	0.024*** (0.001)	0.005*** (0.0001)
				pop	0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)
				unemp	0.014 (0.010)	−0.002*** (0.0004)	0.0001 (0.0004)
				poverty	0.002 (0.003)	0.001*** (0.0001)	0.0002 (0.0001)
				under_65	0.003 (0.004)	0.0002 (0.0002)	−0.0004** (0.0002)
				prop_hosp	1,050.194* (510.694)	−58.352 (32.819)	3.226 (18.700)
				pat_rural	0.077* (0.033)	−0.012*** (0.002)	−0.00004 (0.001)
				hhi	−0.0001* (0.00004)	−0.00000 (0.00000)	0.00000** (0.00000)
				prop_hmo	−1.500*** (0.138)	0.027*** (0.006)	0.009* (0.004)
ind_after	1.589* (0.666)	0.011 (0.006)	−0.487 (0.357)	ind_after	−0.115** (0.036)	0.002 (0.002)	0.002** (0.001)
Observations	4,145	4,145	2,969	Observations	1,654,474	1,654,474	1,654,474
R ²	0.824	0.882	0.865	R ²	0.082	0.049	0.026
Adjusted R ²	0.803	0.868	0.844	Adjusted R ²	0.082	0.048	0.026
Residual	7.453	0.075	3.737	Residual	4.128	0.364	0.242
Std. Error	(df = 3712)	(df = 3712)	(df = 2563)	Std. Error	(df = 1653783)	(df = 1653774)	(df = 1653783)

Note: *p<0.05; **p<0.01; ***p<0.001

Note: *p<0.05; **p<0.01; ***p<0.001

Table A.4: Regression Results: Volume and Capacity

	Dependent variable:	
	Vol.	Beds
	(1)	(2)
ind_after_combined	640.111*** (160.564)	4.453 (2.708)
Observations	4,145	4,145
R ²	0.978	0.983
Adjusted R ²	0.975	0.981
Residual Std. Error (df = 3712)	1,786.602	29.137
Note:	*p<0.05; **p<0.01; ***p<0.001	

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.

Table A.5: Difference-in-Differences Estimates: Average Effect of Hospital Closure on Patient Experience

	Quiet	Clean	Discharge information	Recommend	Doctor communicate	Nurse communicate	Quick help	Staff explain	Pain control
Nearby hospital's closure	−0.387 (0.240)	−0.539 (0.282)	−0.558 (0.368)	−0.650 (0.361)	−0.659* (0.298)	−0.332 (0.405)	−0.555 (0.321)	−0.704** (0.242)	−0.622 (0.364)
Adjusted R ²	0.78	0.82	0.82	0.69	0.71	0.89	0.80	0.80	0.88
Number of observations	2,969	2,969	2,969	2,968	2,968	2,969	2,969	2,969	2,969

Note. The analysis conducted on the same matched sample used for the main analysis. 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.

Table A.6: Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure by Market Characteristics

Hypothesis Outcomes	(1) Operational efficiency	(2) Bed utilization	(3) Service duration	(4) Patient experience	(5) Readmission	(6) Mortality
Competition – high	0.608 (0.742)	−0.001 (0.008)	−0.090 (0.046)	−0.706* (0.320)	−0.0040 (0.0035)	0.0030 (0.0017)
Competition – low	1.053 (0.694)	0.004 (0.006)	−0.127* (0.055)	−0.428 (0.387)	0.0021 (0.0051)	0.0028* (0.0010)
Resources – high	−0.213 (0.725)	−0.013* (0.006)	−0.058 (0.042)	−1.021 (0.628)	−0.0047 (0.0037)	0.0008 (0.0015)
Resources – low	1.660* (0.695)	0.014* (0.006)	−0.152 (0.059)**	−0.089 (0.203)	−0.0010 (0.0055)	0.0029** (0.0009)

Note. The DID coefficient was estimated by stratifying the sample into the corresponding market characteristics. 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.

Table A.7: *Difference-in-Differences Estimates: Heterogeneous Effect of Hospital Closure by Hospital Characteristics*

Hypothesis Outcomes	(1) Operational efficiency	(2) Bed utilization	(3) Service duration	(4) Patient experience	(5) Readmission	(6) Mortality
Size – small	0.943 (0.933)	0.001 (0.007)	–0.277*** (0.069)	–0.843 (0.621)	–0.019** (0.006)	0.004 (0.003)
Size – large	1.316*** (0.327)	0.007 (0.004)	–0.175*** (0.049)	–0.084 (0.127)	–0.004 (0.004)	0.003*** (0.001)
Academic – non teaching	–0.240 (0.664)	–0.009 (0.005)	–0.162** (0.058)	–0.559 (0.414)	–0.011** (0.003)	0.003 (0.002)
Academic – teaching	2.494** (0.769)	0.016* (0.007)	–0.191*** (0.056)	–0.370 (0.318)	–0.004 (0.005)	0.003** (0.001)
Location – rural	–0.875 (1.358)	–0.026** (0.010)	–0.237*** (0.054)	–0.720 (0.625)	–0.017*** (0.005)	0.004 (0.003)
Location – urban	1.761** (0.661)	0.013 (0.007)	–0.174*** (0.049)	–0.278 (0.211)	–0.004 (0.004)	0.003*** (0.001)
Quality – low	–0.106 (0.522)	–0.007 (0.007)	–0.052 (0.066)	–0.001 (0.236)	0.004 (0.005)	0.002 (0.001)
Quality – high	2.437*** (0.559)	0.018*** (0.005)	–0.152** (0.053)	–0.435 (0.496)	–0.002 (0.008)	0.0003 (0.002)

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. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

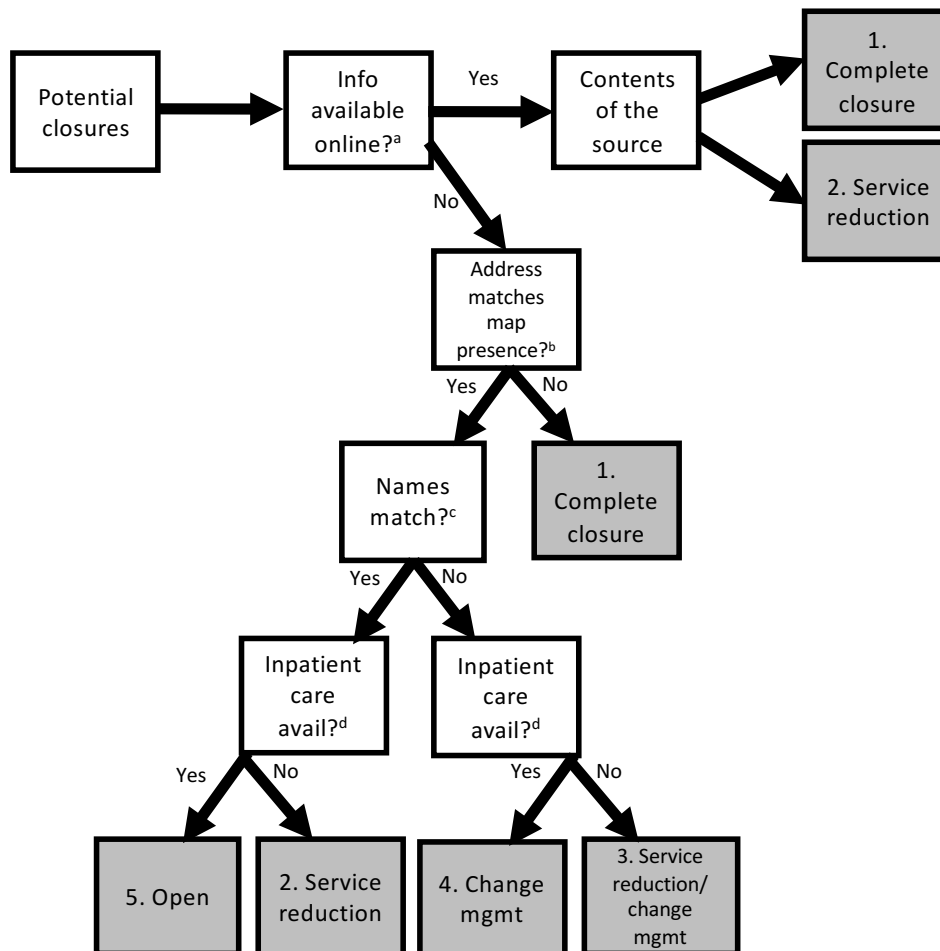
Table A.8: *Difference-in-Differences Estimates: Effect of Medicaid Expansion on Hospital Outcomes*

	Dependent variable:			
	Vol. (1)	Effic. (2)	Util. (3)	Pat.Exp (4)
ind_medicaid_after	199.854 (172.993)	0.911 (0.643)	–0.012 (0.007)	–0.468 (0.462)
Observations	2,448	2,448	2,448	1,711
R ²	0.975	0.8443	0.879	0.869
Adjusted R ²	0.972	0.8237	0.863	0.845
Residual Std. Error	1,847.730 (df = 2162)	7.171 (df = 2162)	0.074 (df = 2162)	3.600 (df = 1451)

Note:

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Note. We estimated the changes in hospital outcomes before and after the Medicaid expansion by comparing the hospitals in and out of the expansion states. We restricted the sample to hospitals that are in our treatment group, i.e., those that did not experience any nearby hospitals' closures.



Notes.

- a. Google searched "Name" + "Closure/merger" for information on closures from local news, Yelp, or Wikipedia.
- b. Examined the physical presence through google map streetview and the date.
- c. Do names of the physical building and the original source match?
- d. Checked the hospital's website to see if inpatient care is available. If the information is not available online, called the hospital.

Figure A.1: *The General Procedure for Identifying Hospital Closures.*

Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
2006 closure	Pre-closure year	Closure year	Post-closure years					Observations removed			
2007 closure	Pre-closure years		Closure year	Post-closure years				Observations removed			
2008 closure	Pre-closure years			Closure year	Post-closure years					Observations removed	
2009 closure	Pre-closure years				Closure year	Post-closure years					Observations removed
2010 closure	Pre-closure years					Closure year	Post-closure years				
2011 closure	Pre-closure years						Closure year	Post-closure years			
2012 closure	Pre-closure years							Closure year	Post-closure years		
2013 closure	Pre-closure years								Closure year	Post-closure years	
2014 closure	Pre-closure years									Closure year	Post-closure year

Figure A.2: Timeline of Hospital Closures

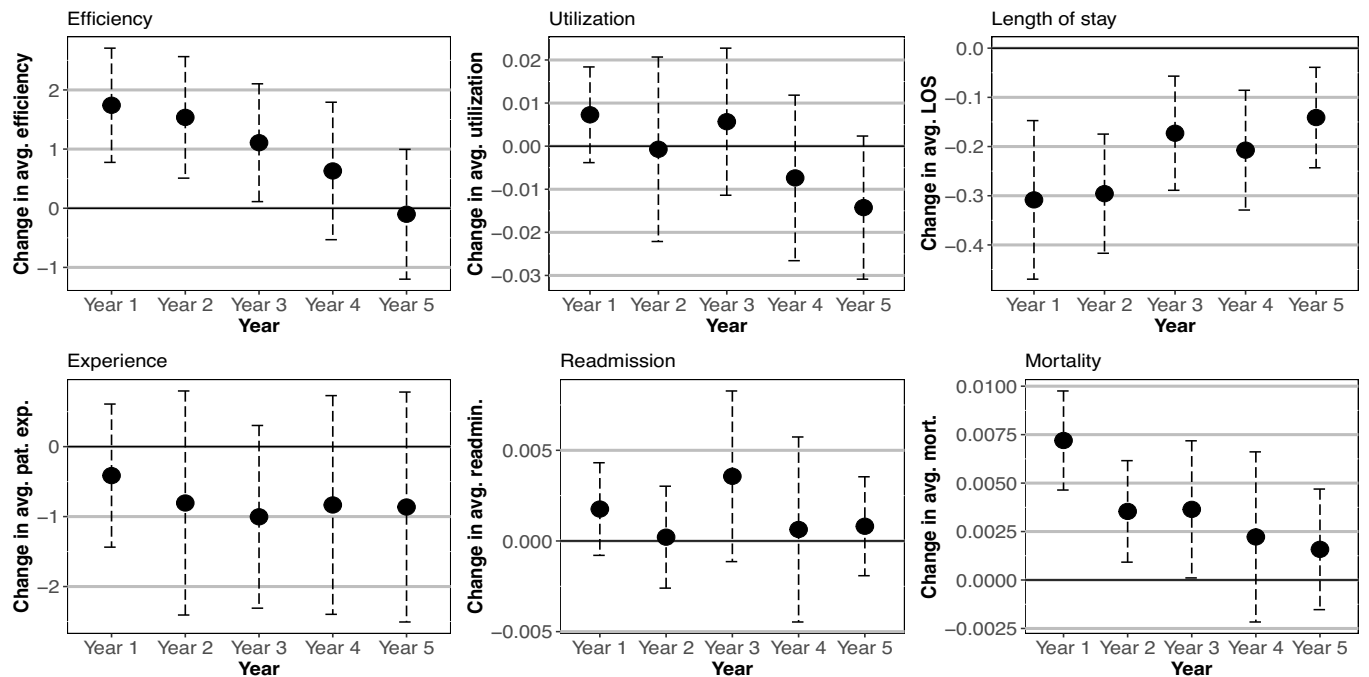


Figure A.3: Difference-in-Differences Estimates: Average Effect of Hospital Closure by Year After Closure

Note. Each dot indicates the size of the DID coefficient for one to five years after closure. 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.

Appendix B

Appendix to Chapter 2

B.1 Proofs

Social welfare

The social welfare, the expected social utility when both hospitals are open is:

$$\begin{aligned}\mathbb{E}[\bar{\Pi}_4] &= p\mathbb{E}[U^1|U^1 > U^2] + (1-p)\mathbb{E}[U^2|U^1 < U^2] \\ &= p\left[q^1 + kx - \alpha\mu^1 - \alpha a_1\mathbb{E}[\tilde{D}|\tilde{D} < c/\alpha]\right] + (1-p)\left[q^2 + k(B-x) - \alpha\mu^2 - \alpha a_2\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha]\right]\end{aligned}\tag{B.1}$$

Note that $\text{Var}(D^1, D^2) = \Sigma = \text{Var}(\bar{D}^1, \bar{D}^2)$. Adapting the framework from Heckman *et al.* (1990), rewrite the variance and covariance $\sigma^2 = \sigma_{11} + \sigma_{22} - 2\sigma_{12}$. $a_1 = \frac{\sigma_{11} - \sigma_{12}}{\sigma^2}$.

$a_2 = a_1 - 1 = \frac{-\sigma_{22} + \sigma_{12}}{\sigma^2}$. $\hat{D} = a_1\bar{D}^2 - a_2\bar{D}^1$. Based on the definitions above, we can rewrite $\bar{D}^1, \bar{D}^2$ $\bar{D}^j = a_j\tilde{D} + \hat{D}$. where $E(\tilde{D}) = 0$ and $E(\hat{D}) = 0$ by construction. U^1 and U^2 can be expressed as $U^j = (kx^j + q^j) - \alpha(\mu^j + \bar{D}^j) + \varepsilon$. $U^1 - U^2 = c - \alpha\tilde{D}$. $p = \Pr[U^1 > U^2] = \Pr[\tilde{D} < c/\alpha] = F_{\tilde{D}}(c/\alpha)$. where $F_{\tilde{D}}$ is the cumulative distribution function of $\tilde{D}$. Thus, expected utility for patients who choose hospital 1 and 2 can also be expressed in terms of c and $\tilde{D}$. $\mathbb{E}[U^1|U^1 > U^2] = (q^1 + kx - \alpha\mu^1) - \alpha a_1\mathbb{E}[\tilde{D}|c/\alpha > \tilde{D}]$. $\mathbb{E}[U^2|U^1 < U^2] = (q^2 + k(B-x) - \alpha\mu^2) - \alpha a_2\mathbb{E}[\tilde{D}|c/\alpha < \tilde{D}]$.

Proof of Proposition 1.

By definition of τ and a_1 ,

$$\tau \geq 0 \Leftrightarrow a_1 \geq \frac{\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] - c/\alpha}{\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] + \mathbb{E}[\tilde{D}|\tilde{D} > -c/\alpha]}$$

By the properties of $\tilde{D}$, $\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] > c/\alpha$ if $c \geq 0$ and $\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] > 0$ if $c < 0$. Thus, $\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] - c/\alpha > 0$. By the properties of $\tilde{D}$, $\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] + \mathbb{E}[\tilde{D}|\tilde{D} > -c/\alpha] > 0$.

Therefore, $0 < \frac{\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] - c/\alpha}{\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] + \mathbb{E}[\tilde{D}|\tilde{D} > -c/\alpha]}$. We also know that $\mathbb{E}[\tilde{D}|\tilde{D} > -c/\alpha] > -c/\alpha$. Thus, $\frac{\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] - c/\alpha}{\mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha] + \mathbb{E}[\tilde{D}|\tilde{D} > -c/\alpha]} < 1$.

Proof of Proposition 2.

Let $\lambda(c/\alpha) = \mathbb{E}[\tilde{D}|\tilde{D} > c/\alpha]$ then $\frac{\partial a_\tau}{\partial c}$ can be written as the following:

$$\frac{\partial a_\tau}{\partial c} = \frac{(\lambda'(c/\alpha) - 1)(\lambda(c/\alpha) + \lambda(-c/\alpha)) - (\lambda(c/\alpha) - c/\alpha)(\lambda'(c/\alpha) - \lambda'(-c/\alpha))}{(\lambda'(c/\alpha) - \lambda'(-c/\alpha))^2} \quad (\text{B.2})$$

We know that $(\lambda'(c/\alpha) - \lambda'(-c/\alpha))^2 > 0$. When $c \geq 0$, from By the properties of $\tilde{D}$, $\lambda'(c/\alpha) - 1 < 0$. $\lambda(c/\alpha) + \lambda(-c/\alpha) > 0$. $\lambda(c/\alpha) - c/\alpha > 0$. $\lambda''(c/\alpha) > 0 \rightarrow \lambda'(c/\alpha) - \lambda'(-c/\alpha) > 0$. When $c < 0$, $-1 < \lambda'(c/\alpha) < 0 \rightarrow \lambda'(c/\alpha) - 1 < 0$. $\lambda(c/\alpha) + \lambda(-c/\alpha) > 0$. $\lambda(c/\alpha) - c/\alpha > 0$. $\lambda'(c/\alpha) - \lambda'(-c/\alpha) > 0$. Therefore, $\frac{\partial \tau_0}{\partial c} < 0$ for all c .

Proof of Theorem 1.

From Proposition 3 and the theorem that the maximum of a convex function on a compact convex set is attained on the boundary.

Proof of Theorem 2.

$$c(B) = (q^1 - q^2) + kB - \alpha(\mu^2 - \mu^2)$$

$$c(0) = (q^1 - q^2) - kB - \alpha(\mu^2 - \mu^2)$$

$$\mathbb{E}[\bar{\Pi}_4(B)] = p(B) \left[q^1 + kB - \alpha\mu^1 - \alpha a_1 \mathbb{E}[\tilde{D}|\tilde{D} < c(B)/\alpha] \right] + (1 - p(B)) \left[q^2 - \alpha\mu^2 - \alpha a_2 \mathbb{E}[\tilde{D}|\tilde{D} > c(B)/\alpha] \right]$$

$$\mathbb{E}[\bar{\Pi}_4(0)] = p(0) \left[q^1 - \alpha\mu^1 - \alpha a_1 \mathbb{E}[\tilde{D}|\tilde{D} < c(0)/\alpha] \right] + (1 - p(0)) \left[q^2 + kB - \alpha\mu^2 - \alpha a_2 \mathbb{E}[\tilde{D}|\tilde{D} > c(0)/\alpha] \right]$$

$$\mathbb{E}[\bar{\Pi}_4(B)] > \mathbb{E}[\bar{\Pi}_4(0)] \leftrightarrow$$

$$p(B) \left[q^1 + kB - \alpha\mu^1 - \alpha a_1 \mathbb{E}[\tilde{D}|\tilde{D} < c(B)/\alpha] \right] - p(0) \left[q^1 - \alpha\mu^1 - \alpha a_1 \mathbb{E}[\tilde{D}|\tilde{D} < c(0)/\alpha] \right]$$

$$> (1 - p(0)) \left[q^2 + kB - \alpha\mu^2 - \alpha a_2 \mathbb{E}[\tilde{D}|\tilde{D} > c(0)/\alpha] \right] - (1 - p(B)) \left[q^2 - \alpha\mu^2 - \alpha a_2 \mathbb{E}[\tilde{D}|\tilde{D} > c(B)/\alpha] \right]$$

Appendix C

Appendix to Chapter 3

C.1 Additional Methods

C.1.1 Measuring Quality

Polypectomy rate: We measured the indication of polypectomy during colonoscopy as a proxy for quality. Adenoma detection is a well-established predictor of interval cancer and is one of the most important colonoscopy quality measures, but measuring adenoma detection requires pathology data that is not available at national level. We instead used a polypectomy, i.e., whether any polyp was removed during a screening colonoscopy, which correlates well with adenoma detection and interval cancer. Polypectomy was identified from claims using HCPCS codes and the existence of pathology bills consistent with previous studies.

Incomplete colonoscopy: Incomplete colonoscopies can result in missed lesions, which is considered a key contributor to the interval cancer. We identified incomplete colonoscopies using the CPT modifier code 53, 73, or 74 based on previous studies. CPT modifier 53 indicates the surgical codes or medical diagnostic codes when the procedure is discontinued because of extenuating circumstances. CPT modifier 73 indicates discontinued outpatient/hospital ambulatory surgical center (ASC) procedure prior to the administration of anesthesia. CPT modifier 74 indicates a surgical or diagnostic procedure requiring anes-

thetia was terminated after the induction of anesthesia or after the procedure was started (e.g., incision made, intubation started, scope inserted) due to extenuating circumstances or circumstances that threatened the well-being of the patient.

Deep sedation use: Because the FDA limits that propofol can be administered by individuals trained in the administration of general anesthesia, another provider (i.e., an anesthesiologist or nurse anesthetist) is usually present during the endoscopic procedure if propofol sedation is used during colonoscopy. We used an approach by an existing study (Khiani *et al.*, 2012; Liu *et al.*, 2012) to identify anesthesiology involvement with colonoscopy. All of the studies relied on the presence of the CPT-4 code 00810, anesthesia assistance with endoscopic procedure distal to the duodenum, occurring on the same date as the colonoscopy of interest. Although this approach does not specifically identify the use of propofol, it is presumed that most anesthesia- assisted procedures would include this agent. Khiani *et al.* (2012) additionally counted the presence of the specialty code for anesthesiologist or certified registered nurse anesthetist (CRNA) on the date of colonoscopy, and Liu *et al.* (2012) included another HCPCS code ("00740" for anesthesia service associated with upper gastrointestinal endoscopies). The agreements in identifying anesthesia accompanied colonoscopies were high across the three studies.

Major and minor complication: We examined three most common and important complications after colonoscopy: perforation, bleeding, and infection. Perforation of colon after colonoscopy is a significant adverse outcome, detected during or soon after colonoscopy and can ultimately result in death. Currently, perforation occurs about 1 out of 1,000 to 1,400 colonoscopies. Bleeding and infection are also common complications after colonoscopy. Evidence shows that the complication rate varies by the endoscopist's characteristics and can be an indicator of the quality of care, including the skills of the endoscopist and the bowel preparation. We identified perforation and infection events using ICD-9 codes that are present either on or within 7 days after colonoscopy, and bleeding events as the presence of ICD-9 codes that are up to 7 days after colonoscopy date, since the presence of bleeding can also be a reason for visit rather than the outcome. We tested the sensitivity of including

or excluding the date of colonoscopy. Minor complications include the presence of other relevant cardiac or other GI symptoms from the literature within 7 days of colonoscopy without including the test date from inpatient, outpatient, or office claims. We examine looking up to 30 instead of 7 days after colonoscopy.

Severity of bleeding: We further examined the severity of bleeding to distinguish between major and minor bleeding based on other studies. We identified major bleeding as either 1) indication of hemoperitoneum or 2) indication of inpatient or emergency department stays with diagnosis codes for gastrointestinal or nonspecific hemorrhage or 3) indication of transfusion within seven days after the test date.

Interval cancer: An interval CRC is a CRC that appears after a negative colonoscopy. An interval CRC may occur due to genetic and clinical factors of the patient, as well as inadequate polypectomy or missed lesions. Thus, higher interval cancer rate may signal poor provider quality, to an extent the patient risk is adjusted. We identified colonoscopies attributable to the interval cancer as the test that precedes 6 to 36 months before the first diagnosis of CRC consistent with previous definitions. To measure the incidence of interval cancer, we first identified the overall CRC incidence. We used the algorithm by Ramsey *et al.* (2004) to find the first presence of diagnosis codes for CRC in inpatient, outpatient, and office visits, and the first date of presentation of diagnosis codes for CRC. Next, we calculated the time interval between the first date of diagnosis and all colonoscopy dates, and identified the colonoscopies that precede first cancer diagnosis by 6 to 36 months. For patients who had more than one colonoscopy within this period, the assignment of cancer was based on the date of the first complete colonoscopy performed. To check the sensitivity, we removed the cases where the patient does not have any confirmation colonoscopy within 6 months before the first date of CRC diagnosis to account for the long interval to confirmatory colonoscopy. There is still a concern for this approach in that the last colonoscopy before cancer diagnosis may have been diagnostic but were not indicated in claims, in which case cancer should not be attributed to the provider. It is also possible that the cancer is recurrent but we are not able to observe previous history.