

Essays on Geographic Variations in Medical Expenditures and Utilization in the Military
Health System

by

William Patrick Luan

Dissertation submitted to the Faculty of the
Preventive Medicine and Biostatistics Graduate Program
Uniformed Services University of the Health Sciences
In partial fulfillment of the requirements for the degree of
Doctor of Public Health, 2018

Distribution Statement

Distribution A: Public Release.

The views presented here are those of the author and are not to be construed as official or reflecting the views of the Uniformed Services University of the Health Sciences, the Department of Defense or the U.S. Government.



UNIFORMED SERVICES UNIVERSITY, SCHOOL OF MEDICINE GRADUATE PROGRAMS
Graduate Education Office (A 1045), 4301 Jones Bridge Road, Bethesda, MD 20814



APPROVAL FOR DOCTOR OF PUBLIC HEALTH DISSERTATION IN THE DEPARTMENT OF
PREVENTIVE MEDICINE AND BIOSTATISTICS

Title of Thesis: "Essays on Geographic Variations in Medical Expenditures and Utilization in the Military Health System"

Name of Candidate: William Patrick Luan
Doctor of Public Health
July 20, 2018

DISSERTATION AND ABSTRACT APPROVED:

DATE:

[Redacted Signature]
Dr. Douglas Robb
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Committee Chair

[Redacted Signature]
Dr. Patrick Richard
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Thesis Advisor

[Redacted Signature]
Dr. Cara Olsen
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Committee Member

[Redacted Signature]
Dr. Todd Leroux
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Committee Member

[Redacted Signature]
Dr. Jonathan Skinner
DARTMOUTH COLLEGE
Committee Member

25 July 2018

ACKNOWLEDGMENTS

I would like to thank my committee for all of their invaluable feedback, guidance, and support throughout this process. I am so incredibly fortunate to have had the opportunity to work with and learn from each of you.

I would also like to thank those who made data available for my research: Alan Sim, Lee Ann Zarzabal, and Alex Rittel at the Solutions Delivery Division and Andrew Findley and the CCQAS Service Representatives at the Defense Health Agency.

Finally, thanks to my family in particular my grandparents (the original Dr. Luan), my mom (the current Dr. Luan), and Emma. I could not imagine being here without your love and constant support. I still believe those early-childhood exposures to medicine and variations have led me on this path. A special thanks to my canine writing companions, Bunny, Charlie, and Bentley, who kept me company throughout the process.

COPYRIGHT STATEMENT

The author hereby certifies that the use of any copyrighted material in the dissertation manuscript entitled: Essays on Geographic Variations in Medical Expenditures and Utilization in the Military Health System is appropriately acknowledged and, beyond brief excerpts, is with the permission of the copyright owner.

[Signature]

A black rectangular box redacting the signature of the author.

William Patrick Luan

September 21, 2018

DISCLAIMER

The views presented here are those of the author and are not to be construed as official or reflecting the views of the Uniformed Services University of the Health Sciences, the Department of Defense or the U.S. Government.

ABSTRACT

Essays on Geographic Variations in Medical Expenditures and Utilization in the Military Health System:

William Patrick Luan, Doctor of Public Health, 2018

Thesis directed by: Patrick Richard, Associate Professor, Department of Preventive Medicine and Biostatistics

Over 40 years ago, researchers first documented geographic variations in the provision of health care in the United States. Recently, scholars have focused their investigations on understanding the causes and mechanisms of variations to help inform health policy.

This work consists of three essays focused on identifying and understanding health care variations in the Military Health System (MHS). The first paper characterizes geographic variations in health care spending and utilization among TRICARE Prime beneficiaries. I find surprisingly large variations in expenditures of up to 70% across catchment areas. By leveraging patient migration, I am able to separate patient-level determinants of variations, which movers take with them, from place-specific factors which remain constant. These results suggest that variations in the MHS are overwhelmingly attributable to place-level factors (80-90%). This study not only has landmark implications for policy in the MHS (spending could be potentially reduced by up to 32% system wide), but also motivates further research to understand what features unique to

location could cause the differences in how seemingly similar patients are treated. One plausible hypothesis is that physicians might contribute to geographic variations. Using a novel dataset of physician characteristics, I find that variations in the cost and complication rates of bariatric procedures vary widely, but that these variations are predominately explained by patients' health status and not by any observable characteristics of the surgeon. Complications result in inpatient stays, which cost 32% more on average. These findings suggest that efforts to improve surgical quality could not only result in better care for patients, but could also reduce expenditures while improving the surgical readiness of physicians. Finally, in the third paper, I use migrating physicians to understand the degree to which physicians alter their practice patterns in response to a change in their environment. Using a difference in differences design, I study physicians performing cesarean sections and find that moving physicians adjust their practice style by approximately .3 percent for every one percent change in their environment suggesting that factors unique to location could explain up to 30 percent of regional variations.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER 1: Introduction	1
Background	1
Geographic Variation in the Military	5
Measuring Outcomes	6
Risk Adjustment and Health Status	6
The Military Health System	7
Study Overview	8
CHAPTER 2: Large Spending Variations Found Among Military HMO Enrollees: A Look Using Compulsory Patient Migration.....	10
Abstract	10
Introduction	10
Methods	12
Data	13
Sample	13
Measures	15
Outcome Variables	15
Independent Variables	16
Geographic Units of Analysis	16
Statistical Analysis	17
Results	18
Sample Demographics	18
Additive Decomposition	19
Event Study	21
Discussion	22
Conclusion	25
Acknowledgements	25
Disclaimer	25
Chapter 3: Variation in Complication Rates and Bariatric Surgery Costs in the Military Health System	32
Key Points	32
Abstract	32
Introduction	34
Methods	35
Data Source and Study Population	35
Outcome Variable – Surgical Complications	36
Outcome Variable – Costs	36

Statistical Analysis.....	37
Results.....	38
Conclusion	42
Disclaimer	42
Acknowledgements.....	43
A. Average Adjusted Complication Rates by Facility	48
B. Average Adjusted Complication Rates by Surgeon	48
C. Average Adjusted Cost by Facility	49
D. Average Adjusted Cost by Surgeon	49
Chapter 4: Physician Compulsory Relocation in the U.S. Military Health System: The Case of Cesarean Sections	50
Abstract:	50
Introduction.....	50
Empirical Approach.....	54
Data	55
Sample.....	55
Measures	56
Geographic Units of Analysis.....	56
Risk Adjusting Cesarean Section Rates.....	57
Analysis.....	58
Results.....	60
Sample Demographics	60
Risk Adjusting	60
Differences in Differences Analysis	61
Discussion	62
Conclusion	65
Acknowledgements.....	65
Chapter 5: Conclusion and Discussion	73
Brief Overview of Major Findings	73
Policy Recommendations.....	75
Future Research	77
REFERENCES	80
Supplemental Appendix.....	85
Supplemental Appendix – Large Variations found among military hmo enrollees: a look using compulsory patient migration	85
Technical Description of Models.....	85
Supplemental Tables.....	88
Robustness Checks.....	93
Supplemental Appendix – Variation in complication rates and bariatric surgery episode costs in the military health system.....	102
International Classification of Disease Codes	102

Supplemental Appendix – Physician Compulsory Relocation in the U.S. Military health system: the case of c-sections	105
International Classification of Disease Codes	105

LIST OF TABLES

Table 1. Summary Characteristics of Tricare Prime Enrollees.....	27
Table 2. Additive Decomposition of Expenditures and Utilization.....	29
Table 3. Variance Decompositions of Expenditures and Utilization.....	30
Table 4. Bariatric Surgeon Cohort Characteristics	44
Table 5. Mixed Effects Logistic Model of Surgical Complication.....	46
Table 6. Multivariable Linear Mixed Model of Surgical Episode Costs.....	47
Table 7. Summary Characteristics of Sampled Military Physicians.....	66
Table 8. Logistic Regression Model of C-Section Risk Factors.....	67
Table 9. Cross-sectional Model of Adjusted C-Section Rates.....	68
Table 10. Difference-in-Differences Analysis.....	72
Table 11. Age and Sex-adjusted Spending by Census Region and Division (2007-2014).....	88
Table 12. Additional Characteristics of Movers	89
Table 13. Additive Decomposition of Expenditures and Utilization – Active Duty Sample	90
Table 14. Additive Decomposition of Expenditures and Utilization – Added Health Adjustment.....	91
Table 15. Traditional Split Sample	93
Table 16. Split Sample by Year	94
Table 17. Split Sample by Year – Active Duty Sample	95
Table 18. Alternate Outcomes – Level Spending	96
Table 19. Alternate Outcomes- Log spending plus .01	97
Table 20. Alternate Outcomes- Log spending plus 10	98
Table 21. Additive Decompositions Using Hospital Referral Regions	101
Table 23. ICD-9-CM Complication Codes.....	102
Table 24. ICD-9-CM Codes for Bariatric Procedures	104
Table 25. ICD-9-CM Codes for Deliveries.....	105
Table 26. ICD-9-CM Codes for Complications.....	105

LIST OF FIGURES

Figure 1. Event Study of Expenditures	31
Figure 2. Surgeon Procedure Choice over Time	45
Figure 3. Complication Rates and Costs by Facility and Surgeon	48
Figure 4. Variation in Average C-Section Rates by Facility	69
Figure 5. C-Section Rates over Time.....	70
Figure 6. Distribution of Changes to Cesarean Section Environments Across Moves	71
Figure 7. Event Study of Utilization	92
Figure 8. Comparison of Adjusted Spending Across Hospital Referral Regions.....	99
Figure 9. Comparison of Adjusted Spending Ranks across Hospital Referral Regions.	100

CHAPTER 1: Introduction

BACKGROUND

It is a common perception in the United States that healthcare costs are too high and that expenditures are growing at an unsustainable rate. Furthermore, evidence has shown that increased levels of spending are not necessarily correlated with better outcomes (37). The sheer magnitude and complexity of the U.S. healthcare system makes accurately quantifying the sources of inefficiency and waste difficult. This waste is estimated to be around 3% of U.S. GDP. Additionally, the healthcare system is so entrenched that it is unlikely that these inefficiencies will be shaken out through market forces and a laissez-faire approach (66). In the military setting, budgetary constraints combined with obligations to care for returning warfighters all while maintaining a medically ready force necessitate reliable, efficient care. The implications to access, quality, safety, training, and readiness are consequential and the potential cost savings of removing care which provides little to no benefit are enormous.

Due to data limitations and the structure of the U.S. health system, much of the evidence of geographic variation has been based upon Medicare claims¹. Significantly less is known about populations under 65 years of age with other payors, particularly those covered under a managed care program. Previous work has shown that per-capita risk-adjusted Medicare spending varies by at least a factor of two across Hospital Referral Regions (HRR) (37). Patient characteristics

¹ See 66.Skinner J. 2012. Causes and Consequences of Regional Variations in Health Care. In *Handbook of Health Economics*, 2:45-86. Number of 45-86 pp. for a comprehensive review)

and patient preferences fail to explain away the differences across regions (10; 35). Within Medicare, research has shown that most of the variation is driven by differences in the quantity of services between different regions (39). By contrast, however, Cooper et al (21) find that variations in spending among the privately insured tend to vary more, by a factor of three, and are largely driven by differences in prices. Furthermore, they find that private spending is weakly correlated to Medicare spending by only 0.14. Their evidence from the Health Care Cost Institute is largely supported by the work of others. Chernew et al (19) found a small negative correlation between private insurers and Medicare data. More recent evidence using multiple databases found weak positive correlations at the HRR level (20).

Researchers agree that supply-side factors are important determinants of regional variations in healthcare expenditures and utilization in the United States. These supply-side factors are roughly divided into either environmental factors such as the local insurance market structure and pricing (54), hospital capacity(38), and productivity spillovers(18), or into physician factors such as education and training(30), or persistent practice styles(30; 69). One plausible theory is that local physician practice patterns drive differences across regions. Epstein and Nicholson (31) study Cesarean Sections and find that while physicians do learn from peers they do not substantially change prior beliefs about treatment. Cutler et al (24) use a novel study design which incorporates hypothetical clinical vignettes to tease out the roles of physicians' beliefs and patients' preferences about treatment. They find limited contributions from patient demand, consistent with Baker et al (8), but find

that unfounded physicians' beliefs about treatment explain 35 percent of spending for end-of-life care and 12 percent of spending for heart attack patients. Doyle et al (30) study patients randomized to physicians in-training from two different residency programs practicing in the same facility. They find that while outcomes are comparable between the two residency programs, treatment costs are higher for patients randomized to the lower-ranked training program. Lastly, Molitor (56) studies moving cardiologists and finds that their catheterization behavior changes .6-.8 percent for every 1 percent change in practice environment. These results suggest that environmental factors such as financial or legal incentives, hospital capacity, or productivity spillovers may explain between 60-80 percent of regional disparities in physician behavior (56). Others point to supply-sensitive care as a potential explanation for variations. In this case, the availability of resource such as CT scanners or MRIs has a major influence on utilization rates. Higher levels of Medicare spending have been shown to be associated with greater use of supply-sensitive services (72).

The empirical challenge in the variation literature has been to disentangle supply-side effects from patient-level factors while using the best-available data. One strategy is to use migration as a natural experiment to separate persistent place factors from individual factors which movers take with them. Recently, research has used patient and provider migration to explain various features of healthcare including variations in hospitalizations and environmental exposures (52), diagnostic intensity (67), patient utilization (34), costs associated with fragmented care (3), and physician practice patterns (56).

Most studies of “movers” use Medicare claims as they are a robust and nationally-representative data source. Studying retiree movers, however, creates unique empirical challenges that could potentially bias results to an unknown degree and direction. Endogenous migration patterns, known colloquially as “snowbirds”, where a patient’s destination is associated with changes to health status, are inherent to the Medicare population. Similarly, healthcare providers who move could “positively match” to geographic regions that suit their practice style (56). While migration is an excellent strategy to separate effects, location is rarely exogenous to other factors of interest in the health care setting.

A unique feature of the U.S. Military, however, presents an opportunity to improve upon previous studies by removing these endogenous migration patterns. This feature, the Permanent Change of Station (PCS) system, assigns service members to their next duty station. Though periodic relocation is not unique to the military, it differs in several important ways which present a better natural experiment for quasi-experimental designs. The duty station to which individuals move is administratively determined by each of the three Services according to first and foremost the requirements of the Army, the Navy, or the Air Force and their personnel needs. Many factors are included in these decisions including authorizations, Service priorities, professional development needs, and the disposition of the force (27). Importantly, however, a service member’s health is generally not a consideration, individuals have little choice as to their final location, and family members tend to move with their sponsor. By leveraging this pseudo-random migration, I can better separate individual effects from place effects under a “mover” empirical strategy.

GEOGRAPHIC VARIATION IN THE MILITARY

Little is known about the extent of geographic variation in the MHS. The Navy Bureau of Medicine and Surgery (BUMED) asked the Center for Naval Analyses (CNA) to broadly investigate geographic variation in the delivery of health services among Navy facilities(44). These reports compared variation in procedure rates across 12 Navy facilities. The authors documented that “high-intensity” facilities diagnose ischemic heart disease over twice as frequently as “lower-intensity” facilities. These same facilities perform more than three times as many percutaneous and coronary artery bypass graft (CABG) procedures when compared to low-intensity medical centers (44). Similar results are found for other procedures such as cesarean section, routine cancer screenings, knee replacement, hip replacement, and various surgical procedures (11; 44; 45). CNA estimates that the most expensive hospital regions have a cost per beneficiary that is roughly 60% higher than the cost of the least expensive regions (11). More recent cross-sectional evidence documents similar variations in procedure rates and annual-costs per beneficiary (48). The same research group finds that procedure rates vary more in the military as compared to Medicare though per-capita spending is strongly correlated ($r=.3$) (2). These studies, while informative and useful for describing the issue, cannot provide actionable recommendations to policymakers. Longitudinal data is available for analysis, yet these studies largely adopt cross-sectional designs which limit the strength of the evidence and any causal claims. Applying more rigorous methodologies to the data advances our understanding beyond

documenting differences across regions by providing nuanced insight into the mechanisms underlying variations.

Measuring Outcomes

Though early studies focused on rates of surgical procedures and the supply of medical capital, the attention in recent years has shifted the focus towards measuring healthcare expenditures. One common critique of comparing expenditures centers on the price adjustment of Medicare payments. Payments to providers and hospitals are typically adjusted to account for cost-of-living, the disproportionate-share program which provides additional reimbursement for serving low-income individuals, and additional reimbursements to compensate for training residents. Gottlieb and colleagues used diagnostic related groups (DRG) and relative value units (RVUs) as measures of utilization to compare real Medicare expenditures and price-adjusted expenditures which represent the cost if Medicare reimbursed for national prices irrespective of location (39). In effect, any additional spending in one HRR versus another would be because more services are provided not because prices are higher. They found that variation fell only slightly after adjusting for prices suggesting that utilization is perhaps a greater driver of payment variations.

Risk Adjustment and Health Status

One of the major challenges of variation studies revolves around the concept of risk-adjusting. Standard adjustment using age, sex, and race fails to adequately capture health status with sufficient fidelity to conduct rigorous analyses. The challenge facing researchers is to answer the question of whether or not there exists regional variation after controlling for differences in health status.

The controversy surrounding risk adjustment largely deals with the uncertainty associated with deriving a patient's health status from claims data. The quality of risk adjustment is limited by the quality of the data on which it is based. In the Medicare setting, risk adjustment relies upon the diagnoses recorded in the patient's medical record. However, if a patient lives in a higher intensity area where more diagnostic tests and procedures are performed, the likelihood of a diagnosis for specific conditions increases (67). This phenomenon could potentially bias studies and lead to improper conclusions for health policy makers. Further complicating matters is the practice of upcoding whereby practices record conditions on submitted claims in such a way that risk scores and subsequently reimbursement are higher for the services performed. The degree to which practices upcode varies and introduces more variability and uncertainty into claims data.

The Military Health System

The Military Health System (MHS) is a global, comprehensive, integrated system responsible for caring for all active and retired members of the military as well as their family members through one of the nation's largest healthcare systems with total spending of over \$50 billion per year. TRICARE is the health care program responsible for administering health benefits for active duty, retirees, and dependents. It is comprised of two overarching programs, the direct care system and the purchased care system, which work in tandem to meet the needs of the population. The direct care system includes roughly 60 inpatient acute care hospitals, 385 stand-alone medical clinics, and 350 stand-alone dental clinics. These facilities are staffed and operated by the Department of Defense with services provided at no cost to eligible Active Duty Service

Members (ADSMs) and their families and at extremely low cost to other eligible beneficiaries who pay a nominal annual enrollment fee and network copayments. The second program, the Purchased Care System, allows the majority of beneficiaries to also receive care from civilian providers in the community (25).

The MHS faces many of the same challenges as civilian health systems including changing demographics, shifting policies, new standards for access and quality, tightening budgets, and advances in science, clinical care, and administration. The Quadruple Aim has served as the MHS strategic framework since 2009. It consists of four pillars: Increased Readiness, Better Health, Better Care, and Lower Cost that guide MHS strategic initiatives, delivery, and reform. Each aim of the strategic framework, adapted from the Triple Aim developed by the Institute for Healthcare Improvement, is relevant to all aspects of healthcare delivery in the system and to this research.

STUDY OVERVIEW

Each of the following three chapters seeks to understand an aspect of geographic variation in expenditures and utilization in the Military Health System. The first paper studies TRICARE Prime beneficiaries to quantify the magnitude of variations and their sources. By studying moving patients, I can separate variation into a patient component and a place component. Results from this work motivate the other two papers to understand two possible sources of supply-side variations at the provider level. The second paper studies variations in complication rates and the inpatient costs of bariatric surgery in the cross-section. I study 38 surgeons to quantify variations in complication rates and costs. The third paper similarly employs a mover strategy but instead considers 496 migrating physicians attending

8,650 births. Observing changes in Cesarean section rates around moves allows us to understand if the practice environment influences physician practice patterns. Collectively, these three works demonstrate the magnitude and sources of variations in utilization and spending across the MHS.

CHAPTER 2: Large Spending Variations Found Among Military HMO Enrollees: A Look Using Compulsory Patient Migration

ABSTRACT

Previous studies have documented large regional spending variations in both the public and private sectors using Medicare and private-payer data. Less is known about care provided to the 9.4 million eligible military personnel and their families. We use the near-random migration of military personnel to different regions of the country to quantify the roles of patient demand and provider supply. Using TRICARE Prime enrollee claims, we find that regional spending per enrolled beneficiary varies by up to 70 percent when comparing the highest to lowest quintiles and shows no correlation to Medicare spending. Ninety percent of geographic variation in spending is attributable to place-specific factors and about 10% is attributable to patient factors. These results from a large, publicly-funded health system suggest that place-factors, such as local market structure, distribution of capital, or productivity spillovers, determine the majority of spending variations among patients with HMO-like coverage.

INTRODUCTION

There is an abundance of evidence documenting regional variations based largely upon fee-for-service Medicare data or private insurance claims. There are many mechanisms hypothesized to be at least partially responsible for differences between regions including: supplier induced demand, defensive medicine and malpractice pressures, the selective sorting of patients according to their preferences, or systematic differences in patient health status.

Recently, research has used patient and provider migration to explain various features of healthcare including variations in diagnostic intensity (67), patient utilization (34), costs associated with fragmented care (3), and physician practice patterns (56). Changes around moves present a natural experiment to separate individual-factors, which movers take with them, from persistent place factors. Most studies of “movers” in healthcare use Medicare claims as they are widely-available source of rich data. One challenge in this context is that a mover’s destination could be associated with changes to health status or so-called “snowbird” migration patterns among retirees. On the provider side, it is hypothesized that physicians could “positively match” to geographic regions that support their practice style exacerbating variations in patient care (56). For instance, a physician with an aggressive treatment style may choose to practice in a higher intensity region. These potentially endogenous migration patterns are inherent to Medicare data.

In this study, we consider the Military Health System (MHS), a nationally-representative publicly funded system with a predominately working-age population, a workforce of salaried military physicians who do not face individual malpractice liability, and administratively determined migration patterns for both patients and providers.

This work uses the near-random migration of TRICARE Prime beneficiaries to separate the roles of supply and demand factors driving geographic variations. We build upon previous work by leveraging the mandatory military system of personnel relocation for both patients and providers, to remove endogenous migration patterns while studying a young, healthy population covered through a managed care program. With a better understanding of the predominant determinants of variations, further research can isolate

hypothesized mechanisms of variations such as physician incentives, physician human capital, or patient preferences. These insights will allow policymakers to tailor healthcare reforms to provide better care at lower cost.

METHODS

The Department of Defense's MHS creates a compelling environment to advance the conversation about variations. It is a global, comprehensive, integrated system that includes combat medical services, peacetime healthcare delivery, public health services, medical education and training, and medical research and development. It is responsible for caring for active and retired members of the military as well as their family members through one of the nation's largest integrated healthcare systems responsible for caring for 9.42 million eligible beneficiaries with total spending of over \$50 billion per year (26). The MHS is comprised of two overarching programs: the direct care system and the purchased care system that work in tandem to meet the needs of the population. The direct care system includes roughly 60 inpatient acute care hospitals, 385 stand-alone medical clinics, and 350 stand-alone dental clinics (26). These facilities are staffed and operated by the Department of Defense. The second program, the purchased care system, which is executed by TRICARE, allows the majority of beneficiaries to also receive care from contracted civilian providers in the community. This dual-mechanism allows for operational continuity during wartime when military facilities are unable to meet increased demand.

At the time of this study, TRICARE consisted of three primary health plans for beneficiaries. TRICARE Prime is the managed care Health Maintenance Organization option where all enrollees have a primary care manager responsible for their care and any

necessary referrals. TRICARE Standard was a basic fee for service program and TRICARE Extra was its complementary network benefit. Active duty service members must enroll in TRICARE Prime at no cost to them. These plans all feature low out-of-pocket costs and standardized, universal benefits for dependents and retirees.

Studying MHS beneficiaries is complementary to previous work on Medicare patients. In addition to the uniform coverage provided by TRICARE and the breadth of the MHS network of care, structural features of the delivery of care to the military population allows us to isolate hypothesized mechanisms of spending and utilization variations. One such feature, the Permanent Change of Station (PCS) system, rotates personnel between duty stations according to the personnel needs of the three branches of the military: The Army, Navy, and Air Force. Many factors are included in these decisions including authorizations, Service priorities, professional development needs, and the disposition of the force (27). Importantly, however, individuals have very little choice as to their relocation destination, health status is generally not factor, and family members often move with their sponsors for PCS moves (52).

DATA

We analyze data from the Department of Defense's Military Health System Data Repository (MDR) database. The MDR is a centralized data repository integrating MHS healthcare data worldwide from direct care healthcare facilities as well as from purchased care claims. This system contains robust beneficiary, clinical, and administrative data.

Sample

For this analysis, we use MDR data for the years 2007 through 2014. We restrict our sample to adults aged 18-64 who were continuously enrolled in TRICARE Prime. We

exclude those who were dually eligible for Medicare or another health insurance plan.

We exclude activated National Guard and Reserve members whose unique operational status could lead to incomplete observation of healthcare use.

To eliminate deployment-related health shocks, we further restrict our analysis to individuals who lack any oversea medical claims and whose primary duty station was within the United States for the duration of the study period. Patients were assigned to TRICARE Prime catchment areas based upon the Defense Medical Information System (DMIS) ID, a facility identifier, where they had the majority of their visits for the year or their assigned DMIS ID based on their home address if they did not receive any care for that year. Movers were categorized by a change in their DMIS ID between consecutive years if the distance of the move was greater than 40 miles. We used a 40-mile threshold as TRICARE determines catchments areas to be the 40-mile radius around a military treatment facility (MTF) or Base Realignment and Closure (BRAC) site. Individuals who move more than once during our study period are excluded; multiple-movers comprised 26% of the total sample.

We conduct a sub-analysis consisting of a sample of active duty service members exclusively to capitalize on the natural experiment, but generally relax this restriction for our main results as policymakers are concerned with all beneficiaries and dependents tend to move with their military sponsors. “Retirees” refer to individuals who have retired from active duty but continue to receive coverage through TRICARE. Results from our active duty sample exhibit a similar place-effect to our main decomposition and can be found in the Supplemental Appendix.

MEASURES

Outcome Variables

Our primary outcome variable is comprised of an individual's annual healthcare expenditures. For costs accumulated in the purchased care system, we sum the total amount paid by TRICARE and the total amount of patient cost-share. For the direct care system, we sum system-generated costs assigned to encounters. We then combine the figures into a single measure of annual expenditures for each patient. Due to the skewed distribution of the expenditure data, for our main model specifications, we add one and take the natural log to reach a more normal distribution. We verify that other transformations of outcomes do not substantially change our results in the Supplemental Appendix.

As a robustness check, we also estimate equations using a quantity-based measure of utilization applying average prices. For this measure, we follow in the spirit of Gottlieb et al (39) by using Medicare Severity Relative Weighted Product (MS-RWP) to represent the amount of effort necessary to perform inpatient utilization and use the sum of relative value units (RVUs) to represent physician utilization in an outpatient setting. The MS-RWP is a Medicare Severity Diagnosis Related Group (MS-DRG) based measure of the relative costliness of a given episode. It consists of the MS-DRG weight and an adjustment for length of stay outliers (57). To create a single measure of quantity-based utilization, we sum a patient's annual inpatient utilization as measured by MS-RWPs and multiply by the average cost per MS-RWP in our sample (\$853.03) and add their annual outpatient utilization as measured by total RVUs multiplied by the average cost per RVU in our sample (\$49.34). Differences in this measure using fixed, average prices are due to increased or decreased service use and not variations in price or reimbursement. We also

add one and take the natural log for our quantity-based measure to account for skewness when estimating equations.

Independent Variables

For our main model specification, we include fixed effects for patients, calendar year, TRICARE catchment area, and a relative year fixed effect term for movers. We also include MDR-generated age-bins as time-varying covariates. In an alternate model specification shown in the Appendix, we include Charlson comorbidity score and a MDR-generated injury flag to better account for any exogenous health shocks^{Error!}

Bookmark not defined.(62)

Geographic Units of Analysis

Unlike much of the variation literature which uses Hospital Referral Regions (HRRs) as the unit of analysis, we utilize administratively-determined TRICARE catchment areas. Some catchments are anchored by a DoD-operated treatment facility while others are not. Additionally, the range of services and sizes of DoD facilities can vary from large medical centers to smaller community hospitals and outpatient clinics. Patients are administratively assigned to catchments and eligible for TRICARE Prime if they live within 40 miles of a MTF or BRAC site, though they can receive care either through the direct care system or with a contracted civilian provider through the purchased care system. Care delivered in military facilities is prioritized through a “right of first refusal” process depending upon capacity and location. If the DoD facility cannot perform the necessary service or procedure, then patients are referred to network physicians or hospitals in the community. Unlike HRRs, multiple catchment areas can represent the same healthcare market for tertiary hospital care.

STATISTICAL ANALYSIS

We first report descriptive statistics of our cohort stratified by “mover” status. Additional characteristics of movers as well as age and sex adjusted regional spending tables can be found in the Supplemental Appendix. Our main results come from an additive decomposition of annual healthcare utilization. Following Finkelstein et al (34), we regress natural-log transformed healthcare expenditures using linear models which include year, catchment, patient fixed effects, a relative year fixed effects for movers and time-variant age bins. Relative year fixed effects are normalized to zero for non-movers. We assume that the difference in observed utilization between any two catchments is equal to the sum of a patient- and place-component(34). We then let the patient-component comprise the sum of the individual patient fixed-effect, relative-year fixed effects, and time-variant patient observables. The place-component is equal to the estimated catchment fixed effect. We compare different sets of high- and low- utilization catchments by presenting differences in the simple averages of the patient- and place-components as well as the overall differences in utilization. To supplement our main results, we conduct a variance decomposition of patient and place effects. We also estimate the amount of the variance that would be reduced if average patient characteristics and place effects were equalized across all catchments.

To present an alternate depiction of changes to healthcare spending, we conduct an event study using a sample of movers only. Again, following Finkelstein et al (34), we scale a mover’s annual health expenditures so that the direction and magnitude of the jump are informative for a variety of destination and origin pairs. We do this by calculating the difference in average ln-transformed expenditures between the mover’s destination and origin catchments and interacting this value with relative-year dummy

variables. These scaled relative year coefficients become the parameters of interest measuring changes in an individual's expenditures relative to the year of their move scaled to the difference between their destination and origin catchments. We then model a mover's ln-transformed annual healthcare utilization using linear fixed effects with an individual, a catchment of origin, and year fixed effects, patient age bins, and our scaled relative-year coefficients. We present plots of the relative year coefficients with 95% confidence intervals calculated from robust standard errors. A scaled value of 0 means that the individual's expenditures are equal to average expenditures at the origin catchment, while a value equal to 1 means the individual's expenditures are equal to those at the destination catchment.

RESULTS

Sample Demographics

Characteristics of our final analytic sample are summarized in Table 1 stratified by "moving" status. Overall, movers tend to be younger, single, male, and on active duty. There are also slight differences in the Service composition as there are more movers associated with the Air Force and slightly fewer associated with the Navy and Army. Rank and region of first observation are similar across movers and non-movers. Both groups have roughly the same number of comorbidities and the same number of years without any healthcare expenditures. Age- and sex- adjusted per-capita annual spending is the same across regions (see Appendix).

Additional information on movers can be found in the Appendix. Briefly, the average move distance was 2,191 miles and movers were less likely to relocate in 2010,

2012, and 2013². The geographic distribution of movers largely stayed constant though there were slightly fewer movers located in the Northeast and West by the end of our study in 2014.

Additive Decomposition

Our main results are presented in Table 2 which shows two additive decompositions of the differences between high- and low-spending catchments. Each of the columns represent a sample of catchments ranked and separated by average spending per capita. For these different collections of catchments, we present the simple average of the overall difference in ln- transformed expenditures as well as the shares attributable to patient and place.

The first column compares regions that fall above median spending to those that are below. Overall, we see a difference in ln-transformed spending of 0.30 (or a proportional difference of 35% in level spending) between the average spending of above median regions as compared to those below the median. Catchments in the highest quintile spend approximately 76% percent more on average than catchments in the lowest quintile. We find that 90 percent of this difference is attributable to place. As we move across columns, we see that the differences in spending widen when comparing regions in both the top and bottom quintiles and deciles. Interestingly, however, the share of difference attributable to place remains constant around 90 percent.

To test the robustness of these results, we re-run the model using our quantity-based utilization measure. This measure applies average nationwide costs from our sample to RVUs and MS-RWPs to better reflect resource use devoid of any variations in

² The relatively large move distance is due to medical facilities in Hawaii and Alaska.

price, reimbursement, or cost allocation that may differ between facilities or between the Purchased Care and Direct Care systems. Therefore, any differences in spending under constant prices are due to differences in utilization because more or less services were provided. This alternate specification shows similar results suggesting utilization differences are important determinants of variations though with a more modest place estimate of about 80 percent.

We replicate the same decompositions but with a restricted sample consisting exclusively of active duty service members. We restrict to active duty movers and non-movers to capitalize on the experiment of the PCS system as well as this population's homogeneity of health status. These results are presented in Table 14. In this sample, we find greater differences in expenditures between high and low spending regions. Comparing expenditures per capita within the highest spending 20 percent of regions to the bottom 20 percent, we see an average difference in ln-expenditures of 1.10 or a proportional difference of 200% in level-spending. The share attributable to place increases slightly within the active duty population to around 96 percent. Similarly, when decomposing utilization across the active duty population, the place share increases to about 83 percent as compared to 80 percent when using the full sample.

To control for shocks to patient health, which may predicate or coincide with relocation, we ran a final decomposition on the full sample using an alternate specification of the model which included additional time variant variables in addition to age bins. This alternate model included an injury flag to account for any acute health shocks as well as variables for Charlson Comorbidity score to control for patient health

changes over time. The addition of these additional variables did not fundamentally change the results (see Table 15).

Lastly, for our main regression specifications modeling expenditures and utilization, we conduct a variance decomposition that can be found in Table 3. This effectively allows us to better understand which independent variables explain the variability in spending. The fourth row correlates average catchment and patient effects as a randomization test of the PCS system. If this term is close to zero, as we find, then we can be sure that, on average, high-cost patients are not differentially sorting to high-cost facilities. The final two rows estimate how much of the variance in ln expenditures would be eliminated if average patient characteristics and average place effects were equalized across catchments. We repeat the same variance decomposition for the quantity-based measure of utilization which shows similar results.

Event Study

We present our primary event study results in Figure 1 which plots the model's scaled relative-year coefficients with 95 percent confidence intervals calculated from robust standard errors. The distribution of the difference between a mover's destination and origin catchments is roughly symmetrical and normally distributed suggesting that moves from high to low intensity regions are as common as from low to high. Our data shows that prior to a move an individual's spending is roughly equal to the average at their origin. There is then a sharp discontinuity around the time of the move when a mover's spending quickly adjusts to their new location upon which it gradually plateaus over time. We show results from an event study of our quantity-based outcome in Figure 7.

DISCUSSION

Researchers largely agree that supply-side factors are an important driver of variations, patient preferences and observables explain relatively little of the variations, and differences in patient health status may be important but measurement error is a perennial challenge (8; 17; 24; 66). Our study exploits the natural experiment of near-random migration of military personnel. We find substantial evidence of variations (76% when comparing quintiles) in the Military Health System using our sample of TRICARE Prime enrollees. We find that 80-90 percent of variations in healthcare expenditures are attributable to place-specific factors, with the remaining 10-20 percent attributed to patients.

These results present different findings from previous studies using Medicare data which found that 40-50 percent of variations can be explained by patient factors, but they do support much of the geographic variation narrative that supply-side factors are the driving force behind variations and should be the focus of future in-depth study. Furthermore, regional spending per enrolled beneficiary varies by up to 70 percent when comparing the highest to lowest quintiles. Holding quality of care fixed across areas, our models suggest that if other regions reduced spending to the same level as the catchments in the lowest quintile, we could expect savings of up to 32%.

Why do variations in the military service differ from those in the Medicare program? First, we would expect patients to be healthier overall, as active duty service eligibility is contingent upon operational readiness which includes physical fitness standards and the exclusion of certain health conditions. Many medical services are also mandatory for active duty members which may reduce any variation in preferences for care. However, the relatively small contribution of patient preferences to variations is

consistent in magnitude to other research using Medicare data (8), but our work presents a smaller patient share overall than found by Finkelstein and colleagues in the Medicare population (34). The large variations in spending could be partially explained by low TRICARE out-of-pocket costs which, on the one hand, limit concerns over financial barriers to seeking care but also remove any checks to overutilization.

Under the Health Maintenance Organization structure of TRICARE Prime, it is conceivable that physicians in their roles as gatekeepers contribute to regional variations. Recently, researchers have focused on understanding characteristics of physicians, which may also cause the differences in care seen across regions. Physician human capital, physician beliefs, and practice style have been previously shown to influence variations in treatment intensity (24; 56). Furthermore, physician training and education has been shown to influence treatment costs and diagnostic efficiency (30). Evidence also suggests that practice style is persistent and learned, with less variation among specialists (16). These factors may become more important under a HMO-like system structure where primary care managers play a large role in a patient's treatment decisions. Further study is necessary to quantify the exact contributions of physicians and their practice patterns as drivers of variations, but the MHS proves an ideal setting to isolate many of the hypothesized drivers in a data rich environment.

This work capitalizes on several key strengths of the setting. First, the randomness of the PCS system removes any endogenous migration patterns which could bias our estimates of the contributions of patient and place. Second, the homogenous health status of the TRICARE Prime population minimizes concerns of unobserved health, particularly among the active duty population. Third, the MDR is a large, centralized data set which

provides robust data for analysis. Nevertheless, there are several limitations associated with this study. First, owing to data restrictions associated with the de-identification process, encounter data was limited to calendar year which may cause additional noise around moves. When running our main models, we excluded data from the year of move, but this could be insufficient for precisely timing the move. Other limitations of the de-identification process include limited demographic data such as restricting rank data to a dichotomous “officer” or “enlisted” which limits our ability to test for changes to preferences over time as socioeconomic status changes. Second, other demographic variables such as race are unreliable in the MDR though many of these are time-invariant and would be controlled for by our patient fixed effect. Third, military treatment facilities are allowed to allocate costs as they see fit; this means, for example, that purchased care claims lack the administrative costs associated with administering TRICARE, while these same administrative costs are included in direct care encounter data. To address this problem, we present quantity-based utilization measures using average prices, so that observed variations are the consequence of utilization changes, rather than patients shifting across accounting systems.

Like many other studies in the variation literature, caution must be exercised when drawing conclusions about efficiency and value using the results from this study. Our study does not control for quality of care. There is a previous literature based on Medicare data that finds mixed results; on net it appears there is little association between spending and health outcomes (36; 60). Measuring quality variation across military facilities is a topic for in-depth future research.

Finally, it is important to consider the results of the study in the context of the mission and structure of the MHS. Value in the MHS is not the same as value in the private sector. Leadership is willing to accept a certain degree of inefficiency in exchange for operational readiness in the case of war, the so-called “cost of readiness.” Policymakers should clarify and consider the optimal allocation of resources and an acceptable cost threshold for readiness.

CONCLUSION

Our study documents large spending variations in a publicly funded HMO with the majority of variations attributable to supply-side determinants. Considerable savings could be realized to tax-payers and the Military Health System through a better understanding of supply-side variations. Though further study is necessary to quantify the contributions of specific determinants, the MHS provides an ideal setting to isolate many of the hypothesized mechanisms in a data-rich environment. These insights will equip both policymakers and clinicians to produce better care at lower cost.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Alan Sim, Lee Ann Zarzabal, and Alexander Rittel at the Defense Health Agency Solutions Delivery Division for their assistance pulling data.

DISCLAIMER

The contents, views or opinions expressed in this publication or presentation are those of the author(s) and do not necessarily reflect official policy or position of

Uniformed Services University of the Health Sciences, the Department of Defense (DoD), or Departments of the Army, Navy, or Air Force. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Table 1. Summary Characteristics of Tricare Prime Enrollees

Note: Summary statistics are provided for the 2007 calendar year.

		Non-movers		Movers	
		Count	Percentage	Count	Percentage
Age Group					
	18-24	6,546	2.34	12,018	10.46
	25-34	26,394	9.42	31,443	27.37
	35-44	78,992	28.19	34,655	30.16
	45-64	168,249	60.05	36,780	32.01
Gender					
	Female	140,287	50.07	50,746	44.17
	Male	139,894	49.93	64,150	55.83
Sponsor Service					
	Army	91,778	32.76	36,270	31.57
	Coast Guard	5,313	1.9	3,282	2.86
	Air Force	93,515	33.38	45,811	39.87
	Marine Corps	17,523	6.25	6,118	5.32
	Navy	64,393	22.98	19,248	16.75
	Navy Afloat	4,671	1.67	3,355	2.92
	Other	2,988	1.07	812	0.71
Beneficiary Category					
	Dependents of AD, Retiree, or Survivor	126,741	45.24	42,856	37.3
	Retirees	118,381	42.25	26,693	23.23
	Active Duty	35,059	12.51	45,347	39.47
Marital Status					
	Single	27,285	9.74	16,994	14.79
	Married	252,896	90.26	97,902	85.21
Rank/Sponsor Rank					
	Enlisted	222,627	79.46	89,680	78.05
	Officer	49,013	17.49	21,538	18.75
	Warrant Officer	8,535	3.05	3,656	3.18
	Cadet	5	0	22	0.02
Region of First Observation					
	Northeast	9,022	3.22	5,849	5.09
	Midwest	23,340	8.33	8,970	7.81
	South	176,161	62.87	67,020	58.33
	West	71,658	25.58	33,057	28.77
	Total	280,181	100	114,896	100
Annual per Capita Healthcare Expenditures (all sources, unadjusted dollars)					

Mean	\$ 4,284.31	\$ 3,865.66	t=30.78	p=<.001
Standard Deviation	(11,726.94)	(8,903.74)		
Median	\$ 1,891.86	\$ 1,623.03		
Geometric Mean	6.96	6.81	t= 48.16	p=<.001
Standard Deviation	(2.48)	(2.50)		
Share of Person-Years with 0 utilization	0.081	0.085		
Average Number of Comorbidities				
Mean	0.750	0.786	t=-15.17	p=<.001
Standard Deviation	0.702	0.613		

Table 2. Additive Decomposition of Expenditures and Utilization

Note: Table is based upon estimation of our main model specification using either ln-transformed annual per-capita expenditures (Expenditures) or ln-transformed quantity-based expenditures (Utilization) as outcomes. The models both include patient, catchment, calendar year fixed effects, and relative year fixed effects for movers which are normalized to zero for non-movers. We also include time-variant MDR-generated age bins. The “expenditure” model had an adjusted R-squared value of .4647 and the “utilization” model’s adjusted R-squared was .4366. The sample includes all movers and non-movers. n=3,160,616 patient years

		Above & Below Median Catchments	Top & Bottom 20% Catchments	Top & Bottom 10% Catchments
Difference in Average Ln Expenditures				
	Overall	.2992	.5645	.7673
	Due to Place	.2708	.5293	.7083
	Due to Patients	.0284	.0352	.0590
Share of Difference				
	Patients	.0950	.0623	.0769
	Place	.9045	.9377	.9231
Difference in Average Ln Utilization				
	Overall	.2616	.4685	.6288
	Due to Place	.2087	.3686	.5092
	Due to Patients	.0529	.0999	.1196
Share of Difference				
	Patients	.2021	.2133	.1902
	Place	.7979	.7867	.8098

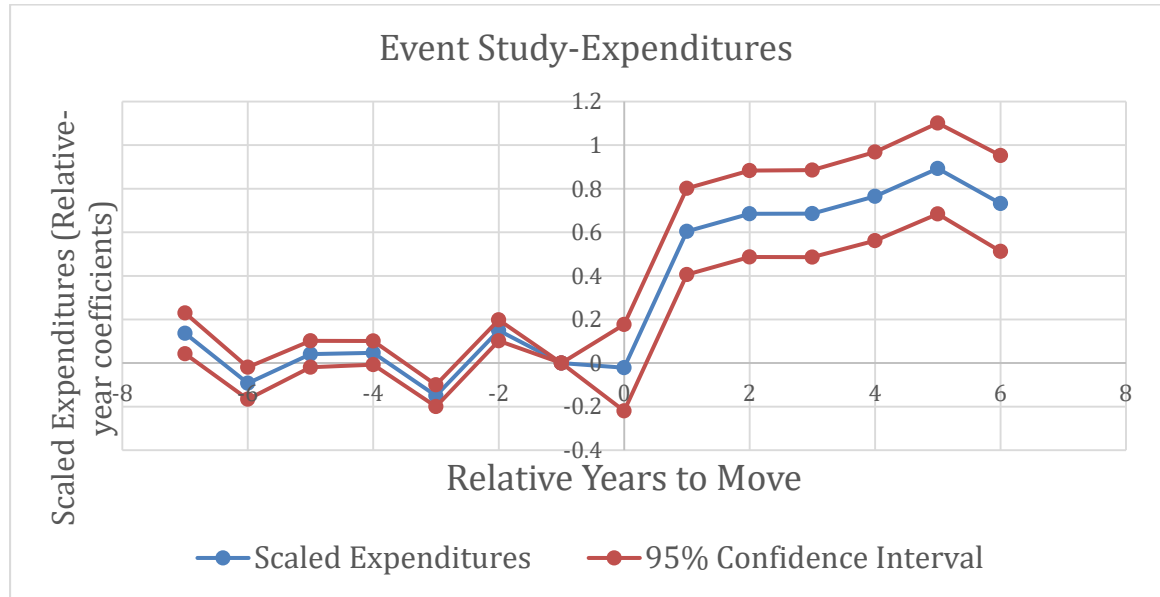
Table 3. Variance Decompositions of Expenditures and Utilization

Note: The variance decompositions use the same model specifications used in our additive decompositions. Again, we ln-transformed annual per-capita expenditures (Expenditures) or ln-transformed quantity-based expenditures (Utilization) as outcomes. The sample includes all movers and non-movers. n=3,160,616 patient years

Expenditures		
Cross-Catchment variance of average:		
Ln Expenditures		.0460
Catchment effects		.0466
Patient Effects		.0053
Correlation of Average Catchment and Patient Effects		-.0002
Share of variance would be reduced if:		
Catchment effects were made equal		.8858
Patient effects were made equal		-.0141
Utilization		
Cross-Catchment variance of average:		
Ln Utilization		.0310
Catchment effects		.0250
Patient Effects		.0073
Correlation of Average Catchment and Patient Effects		-.0006
Share of variance would be reduced if:		
Catchment effects were made equal		.7652
Patient effects were made equal		.1950

Figure 1. Event Study of Expenditures

Note: The event study plots coefficients from our parameterized relative-year coefficients. The dependent variable is an individual's annual expenditures. The models include patient, catchment, calendar year fixed effects, and relative year fixed effects interacted with $\hat{\delta}_i$, the difference in average expenditures between a beneficiary's destination and origin. We also include time-variant MDR-generated age bins. The coefficient for relative year -1 is normalized to zero. The orange lines represent 95% confidence intervals calculated from robust standard errors. The sample includes all movers. N= 919,168 patient years



Chapter 3: Variation in Complication Rates and Bariatric Surgery Costs in the Military Health System

KEY POINTS

Question: To quantify variations in bariatric surgery complications and costs within the U.S. Military Health System.

Findings: In this analysis of military encounter data from 1,277 patients who underwent bariatric weight-loss surgery, there was large variation in complication rates (49.4% CV) and unadjusted procedure costs (25.9% CV) across facilities. After controlling for patient health, variations in complication rates and episode costs fall to 12.6% CV and 4.4% CV respectively.

Meaning: We find large variations in complication rates even after controlling for patient health. Efforts to improve surgical quality would impact costs, complications, and surgical competency of uniformed physicians.

ABSTRACT

Importance: Within the Military Health System (MHS), facilities have struggled to meet minimum recommended volume thresholds for certain procedures.

Understanding variations in complication rates and cost can help policymakers tailor policy to target improvement.

Objective: To quantify the variation in bariatric surgery complication rates and costs across a sample of military hospitals.

Design: A retrospective cohort of 38 military surgeons practicing in 21 hospitals from 2007-2014 who performed 1,277 bariatric surgeries. Analysis was performed in the spring of 2018.

Setting: 21 military hospitals that comprise part of the Military Health System. Data from the Centralized Credentials and Quality Assurance System (CCQAS), which provides education and training characteristics of physicians, was linked to patient encounter data from the Military Health System Data Repository (MDR).

Participants: Physicians were included if they performed at least 5 bariatric surgeries over the study period. Patients were included if they had a diagnosis of obesity (Body Mass Index greater than 30) and underwent a bariatric weight-loss surgery.

Main Outcomes and Measures: Inpatient costs and complication rates were calculated and summarized across surgeons and facilities.

Results: We find evidence of large variations in both complication rates and costs per admission. Overall, we found a 15.5% complication rate across the sample. When comparing averages across facilities, we find large variation in complications (49.4% CV) and procedure costs (25.9% CV). Controlling for patient comorbidities, BMI, and year attenuates much of the variation (12.6% CV-complications, 4.4% CV-cost), but cannot completely explain differences across facilities. Our model suggests that complications cost 32% more than complication-free surgeries on average suggesting that quality-improvement efforts could potentially yield large savings.

Conclusions and Relevance: There are large variations in complication rates and episode costs in the MHS. Policymakers should target efforts to improve surgical quality across facilities and physicians. Surgical quality improvement initiatives could produce savings to the MHS through reduced complications and improved surgical readiness.

INTRODUCTION

With a rise in obesity rates, bariatric surgery has become an increasingly common surgical procedure (47; 64). While mortality rates have decreased over time to .1-.2%, serious complication rates remain a challenge (47). Furthermore, small caseloads and low event rates make reliably differentiating between both low- and high-performing surgeons and facilities difficult (51).

The magnitude and determinants of variation in the costs of surgical procedures are not well characterized. Research has shown that the costs of common surgical procedures can vary widely by up to 28 fold even after adjustment for patient observables and surgical complexity (70). Similarly, complication rates have been shown to vary significantly across facilities by up to a factor of 17 irrespective of surgical volume or accreditation as a surgical Center of Excellence (28; 43). These surgical complications have been shown to be significant drivers of costs to hospitals and health care payers (42; 61). A large vein of research has found that surgical volume is closely associated with better outcomes and lower costs for some procedures. While the literature relating volume to quality is quite extensive, there are few studies investigating the relationship between surgical cost and quality. Work using Medicare data, which has more standardized prices, has found that surgical complications and quality are important determinants of costs (12; 58; 65). Lower-quality, lower-volume facilities have been shown to receive higher payments for 30-day readmissions and post-discharge ancillary care (63). This suggests that efforts to target quality improvement may reduce expenditures.

For the Department of Defense's Military Health System (MHS), these challenges pose a two-pronged quandary. On one hand, due to the relative youth and health of the

military population, the system often lacks the patient caseload to meet recommended minimum volume thresholds at some facilities (13). On the other, leaders within the MHS must balance the organization's role as a payer against its obligation to maintain the highest readiness and competency standards for its surgeons all while providing the safest care possible to its patients. In other words, policymakers need to understand if the "business case" for improving surgical quality and efficiency translates to the MHS and aligns with other operational goals. As the MHS shifts to a centrally managed structure under the Defense Health Agency, an understanding of surgical variations could help policymakers to consider ways to standardize and improve surgical quality and outcomes. Given these challenges, we sought to evaluate variations in complication rates and the costs of bariatric surgery, considered an important training surgery for deploying medical personnel, in a population of uniformed surgeons practicing in military hospitals across the country.

METHODS

Data Source and Study Population

We analyze data from the Department of Defense's Military Health System Data Repository (MDR) database. The MDR is a centralized data repository integrating MHS healthcare data worldwide. This system contains robust beneficiary, clinical, and administrative data. We also extract provider training and education characteristics from the Centralized Credential and Quality Assurance System (CCQAS), a web-based, worldwide credentialing, privileging, and risk management system for the Defense Health Agency. CCQAS is a regularly updated database compiling information on

uniformed providers including graduate education, post-graduate training, and any additional qualifying medical training.

We study a sample of Active Duty physicians who practiced and were stationed within the continental United States for the years 2007 through 2014. We then match bariatric surgeries (identified from ICD-9-CM codes) from the MDR to our sample of providers over the same time period. For this analysis, we focus on bariatric procedures including: Roux-en-Y Gastric Bypass, Sleeve Gastrectomy, Gastric Banding, and Biliopancreatic Diversion with Duodenal Switch. We exclude surgical revisions from our sample. Practice location is determined by the Defense Medical Information System (DMIS) ID, a facility identifier. We excluded physicians who performed fewer than 5 surgeries over the 8-year period of observation resulting in a 6% decrease in sample size. Our final patient sample consisted of 1,277 unique TRICARE beneficiaries who received bariatric surgery from 2007 through 2014. Due to data restrictions resulting from de-identification, we have limited patient demographic information.

Outcome Variable – Surgical Complications

We calculate complication rates using previously described ICD-9-CM codes (9; 64). We include a full list of the codes used in the Supplemental Appendix. We then create an indicator variable for the presence of any recorded complication. For this analysis, we only consider complications which occurred during the initial inpatient admission.

Outcome Variable – Costs

This outcome variable is comprised of a surgical encounter's cost. For this analysis, we use MDR generated Patient-Level Cost Accounting (PLCA) figures which

best capture the true cost to the MHS of a surgical episode. PLCA allocates patient care, support, and overhead costs to clinical services. These costs are centrally administered and regularly audited for accuracy. Due to the right-skewed distribution of expenditure data, for estimating equations, we take the natural log to reach a more normal distribution appropriate for use with linear models.

STATISTICAL ANALYSIS

First, we present summary statistics of our physician cohort including demographic and educational characteristics. We then summarize counts of different types of bariatric surgery and average costs stratified by facility. For unadjusted costs presented in summary tables and figures, we deflate costs to 2014 dollars using the Medical Care component of the Consumer Price Index (CPI-U) (1; 15).

We use multivariable mixed-effects logistic models to adjust complication rates for patient-level factors. The model includes the number of Elixhauser comorbidities, patient body mass index (BMI) bins, and calendar year as covariates. To account for clustering and small samples, we include random effects at both the facility and individual surgeon levels to shrink our estimates. We then use this model to calculate adjusted complication rates for both surgeons and facilities. To model spending, we use linear mixed-effects models to estimate ln-transformed spending per surgical episode including fixed-effects for year, the number of Elixhauser comorbidities, patient BMI bins, and an indicator for the presence of a complication. This multilevel model also includes random effects at both the facility and physician levels. Note that when summarizing variations in spending at the facility and physician levels in subsequent figures we remove the complication indicator from the model to capture real variation in

complications that could be causing differences in spending across hospitals. We apply a Bonferroni correction to adjust confidence intervals for multiple comparisons when plotting costs and complications rates. Analysis was conducted using STATA Version 15.

RESULTS

We analyzed a total of 1,277 bariatric surgeries. Characteristics of the surgeons in our sample are shown in Table 4. Overall, surgeons are more likely to be affiliated with the U.S. Army, tend to be more senior officers, possess a MD, and be trained as a general surgeon. Individual surgical volumes vary substantially and are relatively low as compared to civilian medicine with about 12 surgeries per year. Figure 2 summarizes counts of the type of bariatric surgery by the 21 military hospitals in our sample. The blue line overlaying the histogram depicts the average unadjusted cost per surgical episode for each facility deflated to 2014 dollars. Laparoscopic Roux-en-Y Gastric Bypass and Laparoscopic Sleeve Gastrectomy were the most common types of procedures performed. Open surgeries were less common in our sample, consistent with trends in clinical practice towards less invasive procedures (49). Unadjusted, untransformed surgical episode costs exhibited variation in the sample with mean costs of \$22,554 and a coefficient of variation across hospitals of 26%. We find that procedure choice, surgical volume, and cost are all heterogeneous across facilities. The most frequently performed type of bariatric surgery varied across hospitals as did overall bariatric volume in the sample. Overall, the complication rate in our sample was 15.5%, in-range with comparable estimates among civilian facilities (12.3-18.7%) when using the same set of ICD-9 codes of complications(9).

To better account for differences in patient health, we present results from our models of complications and cost in Tables 5 and 6 respectively. In Table 5, we present odds ratios for variables included in the model. Multiple comorbidities strongly increase the odds of a surgical complication. We then use this model to calculate adjusted complication rates for each physician and hospital which we present in Figure 3 Panels A and B. Along with point estimates, we also present corresponding 95% confidence intervals corrected for multiple comparisons. The orange line represents the overall sample average. We find evidence of significant variation in complication rates even after adjustment when considering both individual surgeons (17% CV) and hospitals (12.6% CV).

Next, we consider our model of spending. We use natural-log transformed costs as an outcome to better account for the right-skewed distribution of our data (55). This transformation means that regression coefficients must be exponentiated in order to be easily interpreted. The last column in Table 6 presents exponentiated coefficients representing the percent change in untransformed dollars of a one unit change in the independent variable. Our model shows that comorbidities and complications explain a large share of the differences in spending. For instance, surgical episode costs are 30% higher for patients with 6 Elixhauser comorbidities relative to those without a single comorbidity. Similarly, complications result in 32% higher episode costs on average when controlling for other variables. The variance components of the random effects estimate the amount of the overall variance in costs attributable to facilities and individual physicians nested within facilities. Differences between hospitals explain 38% of the total variance while differences between surgeons explain 2%. While the total

variance is relatively small, hospital-level effects explain a greater share of the variance than physicians.

Figure 3 Panels C and D present physician- and facility- average adjusted costs and corresponding 95% confidence intervals. Note that when calculating these adjusted costs, we do not control for complications to capture any heterogeneity in complication rates which may be driving spending differences. Adjusted costs are estimated with fixed effects for patient BMI bins, the count of Elixhauser comorbidities, and calendar year with random effects at the hospital and surgeon levels. The orange horizontal line represents the sample average adjusted cost. Ninety-five percent confidence intervals are again adjusted for multiple comparison testing using a Bonferroni correction. We find that several facilities consistently provide more efficient than average care while other hospitals appear more expensive. This pattern also holds when comparing across physicians.

In this study, we aimed to estimate variations in the costs and complication rates of bariatric procedures within the Military Health System. Upon controlling for patient health through comorbidities and BMI, we find that variations persist with respect to both costs and complication rates. The differences across both hospitals and surgeons suggest that there is room for sustained quality improvement. Furthermore, efforts to standardize and improve surgical quality could not only lead to better patient outcomes and improved surgical readiness, but could also decrease costs to the Military Health System.

This study had several strengths that give credence to our findings. The Military Health System Data Repository is a robust data source particularly for encounters which occur in DoD-operated facilities. These encounters have a distinct advantage to claims in

that we can observe the true cost to the DoD by using patient-level cost accounting. Our modeling approach which uses random effects at both the hospital and surgeon levels helps accounts for clustering and small sample sizes that may lead to imprecise estimates. Random effects allow us to shrink our estimates for reliability and consistency.

The MDR contains robust beneficiary data; however, owing to data restrictions and de-identification, we could only account for clinical factors recorded in the MDR and could not model patient demographics such as age, race, or sex. Additionally, we were limited to investigating complications which occur during the initial inpatient admission and cannot examine other quality measures such as readmissions or 30-day complications. Lastly, our study may not be directly generalizable to other settings. Though the MHS faces many similar challenges as the civilian health care sector, research has shown that patterns of variation are vastly different between private and public payers (21).

The Military Health System has come under considerable pressure in recent years to meet minimum thresholds for surgical volume. Our study suggests that despite the relatively low volume observed in this sample our overall bariatric complication rate (15.5%) is comparable to rates among civilian centers when using the same set of ICD-9 codes for complications (12.3-18.7%)(9). These results support work by Dimick and colleagues who suggest that bariatric procedure volume is not strongly predictive of quality or outcomes (29).

In light of the landmark reorganization of the MHS under the Defense Health Agency, policymakers have a unique window of opportunity to shape the future of military medicine. In this context, policymakers should prioritize efforts to reduce

surgical complications and improve quality through sustained and standardized quality initiatives. It is also important to remember the MHS's perspective as a payer. Annual budgets are not created using risk-adjusted dollars, and costly complications can contribute to rising health care costs. Policymakers can, however, align their obligations to provide cost-effective, high-quality care with the need for competent surgeons capable of deploying to the battlefield. Future research should focus on evaluating a unified strategy for the Military Health System which best combines the "business case" for quality improvement with other operational goals such as surgical readiness and a healthy beneficiary population.

CONCLUSION

In summary, as the Military Health System seeks to promote cost-efficiency, surgical competency, and high-quality care, policymakers and clinicians alike must grasp the dynamics of high-risk, low volume procedures such as bariatric surgery. To efficiently improve quality and reduce costs, policymakers must understand the magnitude and determinants of variations. This study suggests that standardized improvements to surgical quality would benefit all parties through safer care at lower cost.

DISCLAIMER

The contents, views or opinions expressed in this publication or presentation are those of the author(s) and do not necessarily reflect official policy or position of Uniformed Services University of the Health Sciences, the Department of Defense (DoD), or Departments of the Army, Navy, or Air Force. Mention of trade names,

commercial products, or organizations does not imply endorsement by the U.S. Government.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Alan Sim, Lee Ann Zarzabal, and Alexander Rittel at the Defense Health Agency Solutions Delivery Division for their assistance pulling MDR data. We would also like to thank Dr. Andrew Findley and the CCQAS Service representatives for making their data available.

Table 4. Bariatric Surgeon Cohort Characteristics

Service	Count	Percent
Air Force	9	24%
Army	18	47%
Navy	11	29%
Rank		
O-4	7	19%
O-5	13	34%
O-6	18	47%
Medical Degree		
DO	3	8%
MD	35	92%
Recorded Specialty (2017) ^a		
Gastroenterology	1	3%
General Surgery	35	91%
Trauma Surgery	1	3%
Surgical Oncology	1	3%
Total	38	100%
	Mean	SD
Years of Post-Graduate Experience	13.02	3.95
Number of Bariatric Surgeries per Year	12.45	7.84
Physician-Years of Observation	5.44	1.71

Figure 2. Surgeon Procedure Choice over Time

Note: Unadjusted costs used in summary tables are adjusted to 2014 dollars using the Medical Care component of the Consumer Price Index for Urban Consumers (CPI-U)

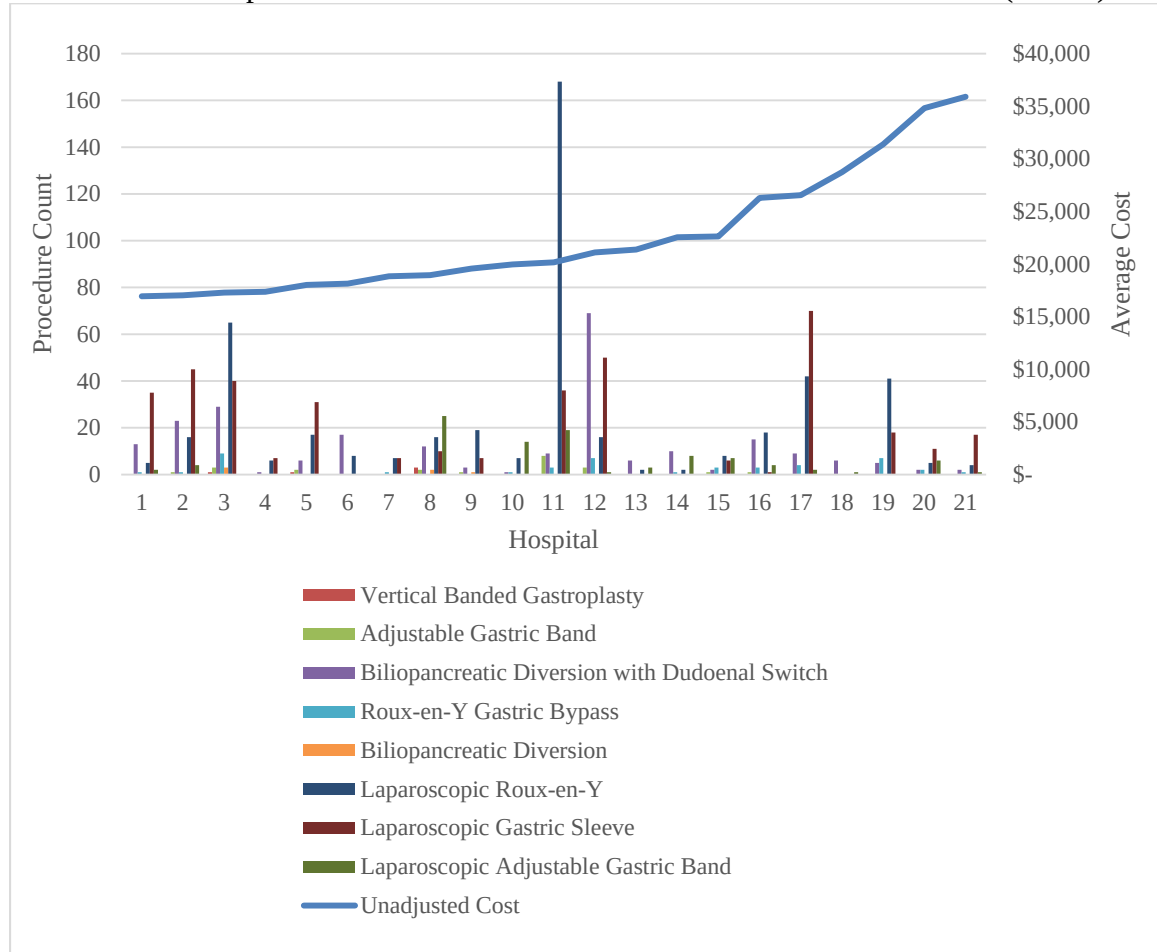


Table 5. Mixed Effects Logistic Model of Surgical Complication

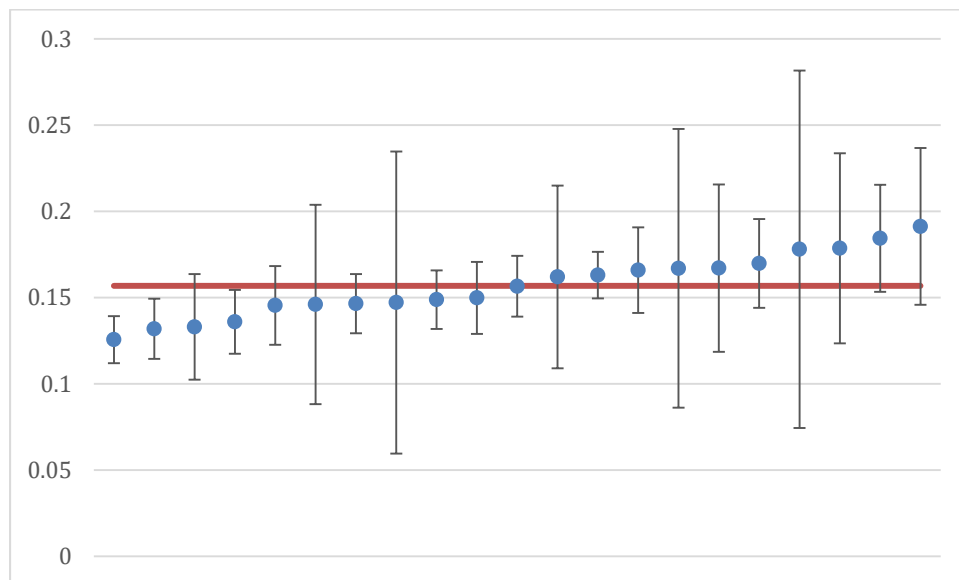
	Odds Ratio	Standard Error	P-value
BMI 30-39	Reference		
BMI>40	1.037	.189	.840
Elixhauser Comorbidity Count			
0	13.099	8.314	<.001
1	Reference		
2	1.022	.240	.925
3	1.756	.404	.015
4	1.023	.323	.941
5	2.995	1.213	.007
6	5.623	4.403	.027
Calendar Year			
2007	Reference		
2008	.432	.183	.047
2009	.748	.277	.433
2010	.662	.244	.262
2011	.754	.265	.421
2012	.342	.137	.006
2013	.651	.235	.234
2014	.429	.162	.025
Constant	0.123	0.029	<.001
Random-Effects Levels			
Facility	Variance	SE	
Physician	8.94E-36	4.00E-19	
	0.159	0.089	

Table 6. Multivariable Linear Mixed Model of Surgical Episode Costs

	Coefficient	SE	P-value	Proportional Change (%)
BMI 30-39	Reference			
BMI>40	0.024	0.015	0.121	2.4
Elixhauser Comorbidity Count				
0	-.135	.071	.056	-12.6
1	Reference			
2	-.019	.018	.313	-1.8
3	-.002	.020	.927	.2
4	0.072	.025	.004	7.5
5	0.044	.042	.294	4.5
6	0.260	.088	.003	29.7
Complication Indicator	.277	.019	<.001	31.9
Cy				
2007	Reference			
2008	-0.033	0.037	0.360	-3.2
2009	-0.094	0.035	0.008	-9.0
2010	-0.069	0.035	0.050	-6.7
2011	-0.078	0.034	0.022	-7.5
2012	0.062	0.034	0.071	6.4
2013	0.091	0.034	0.008	9.5
2014	0.085	0.034	0.012	8.9
Constant	9.654	0.089	<.001	
Random-Effects Levels	Variance	SE		
Facility	0.038	0.013		
Physician	0.002	0.002		
Residual	0.059	0.002		

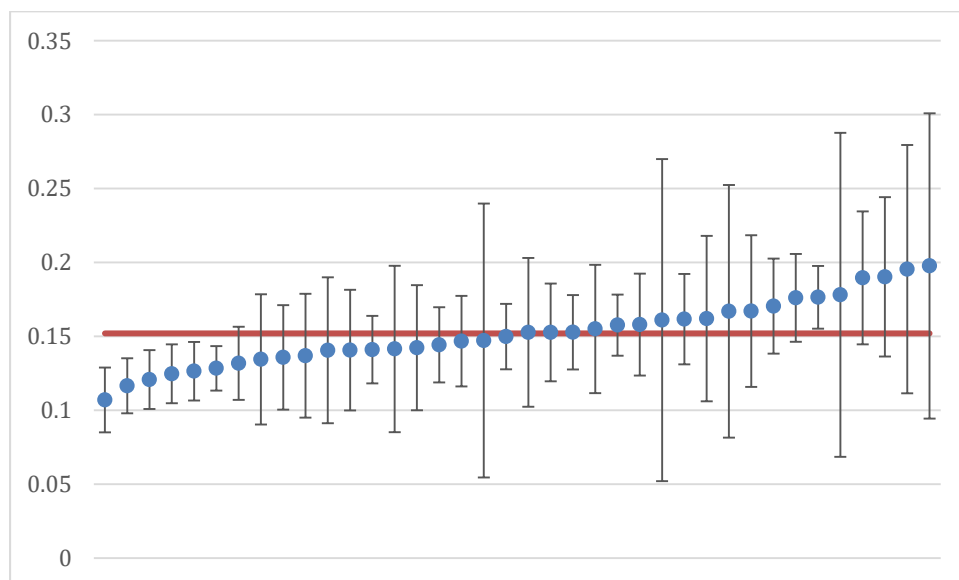
Figure 3. Complication Rates and Costs by Facility and Surgeon

A. Average Adjusted Complication Rates by Facility



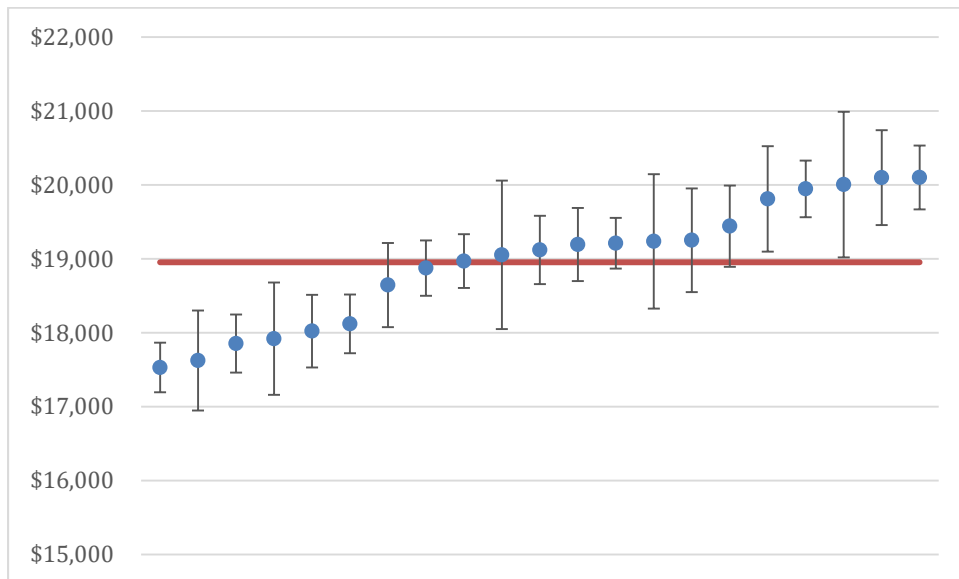
Note: An observation represents a facility's average adjusted complication rate with accompanying 95% confidence intervals. Complication rates are adjusted for calendar year, Elixhauser comorbidities, and patient BMI bins. The orange line represents the sample average complication rate. The coefficient of variation is 12.6%

B. Average Adjusted Complication Rates by Surgeon



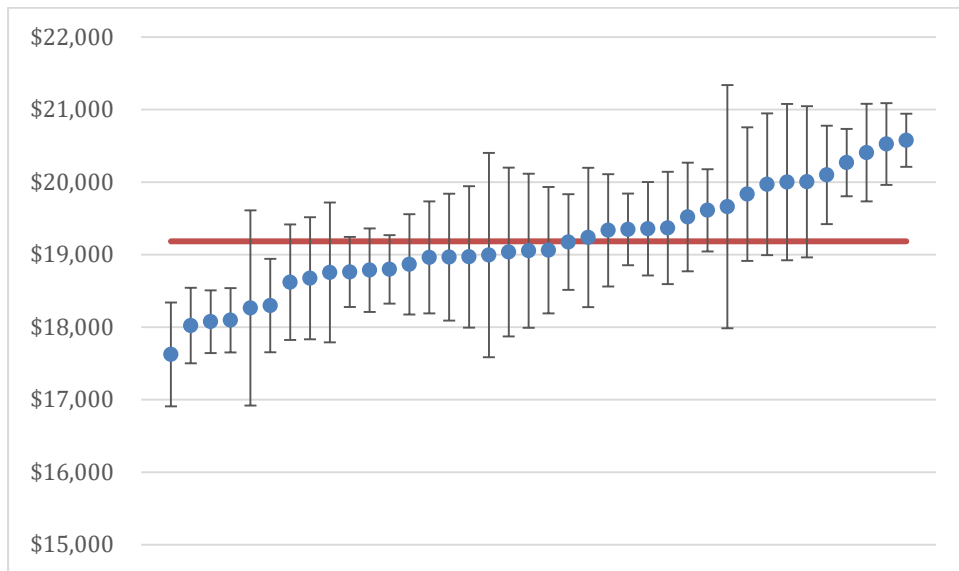
Note: An observation represents a surgeon's average adjusted complication rate with accompanying 95% confidence intervals. Complication rates are adjusted for calendar year, Elixhauser comorbidities, and patient BMI bins. The orange line represents the sample average complication rate. The coefficient of variation is 17.0%

C. Average Adjusted Cost by Facility



Note: An observation is a facility's average surgical cost with accompanying 95% confidence intervals. Costs are adjusted for calendar year, Elixhauser comorbidities, and patient BMI bins. The orange line represents the sample average adjusted cost. The coefficient of variation is 4.4%.

D. Average Adjusted Cost by Surgeon



Note: An observation is a physician's average surgical cost with accompanying 95% confidence intervals. Costs are adjusted for calendar year, Elixhauser comorbidities, and patient BMI bins. The orange line represents the sample average adjusted cost. The coefficient of variation is 3.8%.

Chapter 4: Physician Compulsory Relocation in the U.S. Military Health System: The Case of Cesarean Sections

ABSTRACT:

Supply-side factors are thought to be the predominant force driving geographic variations and previous studies have documented large variations in physician practice style when treating seemingly similar patients. We exploit the natural experiment of the near-random migration of military physicians to separate the contributions of physician-specific factors from the practice environment in the case of Cesarean sections. We find that physicians modify their C-section behavior in response to new environments across a move. We estimate that 28-37 percent of regional variations in physician practice patterns could be explained by environmental factors unique to the practice setting.

INTRODUCTION

Researchers have established that large healthcare variations exist across the United States as measured by per-capita expenditures, utilization, or rates of certain procedures (66). This extensive vein of literature has shown that seemingly-similar individuals receive significantly different amounts and types of care depending on where they receive that care (71). Similarly, researchers have documented variations in healthcare delivery and physician practice style. Adjusting for underlying patient illness, demographic differences, measured preferences, and other observable factors does not explain the variations (8; 10; 73). As well as spending, quality of care, is similarly varied across geographic regions with different patterns across the publicly and privately insured (21; 53). In addition to saving money, variations in quality might be indicative of opportunities to streamline and improve care.

Despite the extensive research documenting variations, much less is known about the mechanisms driving variations. Supply-side factors have been increasingly thought to be the predominant force exacerbating variations. These supply factors roughly divide into either environmental factors such as the local insurance market structure and pricing (54), hospital capacity(38), and productivity spillovers(18), or into physician factors such as education and training(30), or persistent practice styles(30; 69). The challenge for researchers has been disentangling physician factors from place factors using the best available data and methodologies. Many physician factors such as education and training may form endogenously with place (56). Furthermore, physicians may positively match to geographic regions which fit their practice style thereby overestimating contributions of place-factors to variations. Lastly, as noted by Song et al (67), diagnostic practices vary regionally with higher intensity regions being more likely to record diagnoses. This diagnostic intensity bias presents a challenge when risk adjusting using claims databases.

Researchers have addressed these challenges by using migration as a natural experiment to separate persistent place factors from individual factors which movers take with them. These studies aim to explain hypothesized mechanisms of variations including diagnostic intensity (67), patient utilization (34), costs associated with fragmented care(3), and environmental impacts to physician behavior (56). Our study is most similar to work estimating the effect of the practice environment when investigating catheterization behavior of cardiologists who have recently moved. That study found that 60-80 percent of regional disparities in physician behavior are attributable to the practice environment (56). One common threat to validity of these previous studies is that physician migration may be endogenous to unobserved factors.

This work builds upon the literature by leveraging the near-random migration of physicians in the U.S. Military Health System (MHS) through the Permanent Change of Station (PCS) system. More precisely, we investigate how physicians who move change their Cesarean section behavior before and after moves when treating seemingly similar patients. Studying C-sections is valuable for several reasons. First, C-sections are a common, well-studied surgery that has been previously shown to vary by up to tenfold in ways that cannot be explained by patient health or preferences in the civilian healthcare system (50). There is also evidence that higher-intensity areas perform C-sections in less medically appropriate populations with associations between procedure use and greater provider density, capacity of the local health care system, and malpractice pressure (6). Recent evidence also suggests that birth outcomes would be improved if low-risk patients received fewer C-sections while rates should increase for high-risk patients (23).

Within the Military Health System, C-sections have been shown to vary widely when comparing age- and race-adjusted rates (48). Second, within civilian health care there has been pressure from the Joint Commission and other organizations to reduce growing C-section rates (68). One important study found that patients who are themselves physicians and are presumably better informed about medical risk had lower C-section rates than comparable mothers (46). For the MHS, obstetrics represents the largest medical service line and is an important area of focus as budgetary constraints increase (48). Lastly, claims for birth encounters, particularly in the military's data-rich environment, contain sufficient information to adequately risk-adjust rates for fair comparisons across patient populations making it ideal for studies of this nature.

Studying C-sections in the context of the MHS provides additional clarity and unique advantages for observing physician decision making from claims databases. One theory behind variations is that defensive medicine encourages physicians to induce additional care to protect themselves from potential lawsuits. While there is empirical evidence refuting the occurrence of defensive medicine particularly in the case of C-sections (see Baicker et al (7) and Currie and MacLeod (22)), for our study, military physicians are not subject to personal malpractice suits for care provided at military treatment facilities. Therefore, we can be confident that defensive medicine is not influencing physician decision making. Furthermore, uniformed physicians are salaried according to rank, time in service, and specialty. Compensation is not tied to the number or type of procedures performed. While Gruber et al (40) and Gruber and Owings (41) present early evidence of inducement of C-sections for financial gain, we should not see this same phenomenon in the MHS.

Next, Cutler et al (24) demonstrate the importance of physician beliefs about treatment as determinants of regional variations. Similarly, Epstein and Nicholson (31) found that within market variation exceeds cross-market variations and that while physicians may learn from peers, they do not substantially change previously-held beliefs. Through other MHS databases, we can use education and training as a proxy for physician beliefs to see if there is a correlation to practice style. Lastly, healthcare benefits provided to military service members and their dependents offer generous coverage and all feature low out-of-pocket costs irrespective of pay-grade or occupation minimizing concerns over access or a patient's ability to pay influencing a physician's decision making.

EMPIRICAL APPROACH

We exploit physician migration to analyze how changes in location and the clinical environment impact physician practice. We observe physician C-section behavior changes across moves and compare it to rates from other physicians practicing in the moving physician's origin and destination. We then implement a linear difference-in-differences model to estimate how a change in environment impacts physician behavior. A key assumption here is that nothing besides the physician's environment changes across the move. If other factors that are equally correlated with a change in environment and physician behavior coincide with the move, then this potential endogeneity could bias our results. A second key assumption is the parallel trends assumption which states that broader physician trends in behavior would have equally impacted physicians regardless of whether or not they moved. We then take the following equation to the data:

$$\begin{aligned} (\text{Cesaran section})_{jt} &= (\text{Origin Catchment FEs})_j + (\text{Calendar Year FEs})_t \\ &+ (\text{Relative Year FEs})_j + (\text{Patient Risk Adjusters})_i \\ &+ (\text{Physician Characteristics})_j + \beta \Delta_j * T_{\text{post-move}} + \epsilon_{i,j,t} \end{aligned}$$

We estimated the regression where the dependent variable,

$(\text{Cesaran section})_{ijt}$, is an indicator for whether patient, i , received a C-section performed by physician j , in year, t . We include fixed effects for a physician mover's first catchment, year relative to move, and for calendar years. We include a panel of indicator variables for clinical risk factors for C-section. These factors derived from the literature are the same factors used in our logit model used to risk-adjust catchment rates presented in Table 9. We also include a set of indicators for physician characteristics and time-variant clinical experience bins. Δ_j denotes the difference in average practice style between a mover's destination and origin catchments. We interact Δ_j with an indicator

$T_{post-move}$ which signifies encounters after a physician moves. The parameter of interest is then β which describes the change in an individual physician's practice style before and after they move scaled to the average difference in practice styles between the movers destination and origin catchments. Since selective migration is less of a concern, we can also test alternate specifications of our model by including physician fixed effects to measure practice changes across a move relative to prior behavior.

DATA

We analyze data from the Department of Defense's Military Health System Data Repository (MDR) database. The MDR is a centralized data repository integrating MHS healthcare data worldwide from direct care healthcare facilities as well as from purchased care claims. This system contains robust beneficiary, clinical, and administrative data. We also extract provider training and education characteristics from the Centralized Credential and Quality Assurance System (CCQAS), a web-based, worldwide credentialing, privileging, and risk management system for the Defense Health Agency. CCQAS is a regularly updated database compiling information on graduate education, post-graduate training, and any additional qualifying medical training.

SAMPLE

For this analysis of cesarean sections, we study active duty physicians who practiced and were stationed within the United States for the years 2007 through 2014. We then match inpatient deliveries (identified from ICD-9-CM codes) from the MDR to our sample of providers over the same time-period. Each provider has a unique identifier that allows us to follow them over time and across practice environments. Practice location is determined by the Defense Medical Information System (DMIS) ID, a facility

identifier, where the majority of clinical encounters occurred for the year. Moving physicians were categorized by a change in their DMIS facility ID between consecutive years. The same identifier allows us to link providers to their educational and training history recorded in CCQAS. We excluded physicians who moved more than once and physicians who attended fewer than 5 births over the period of the study. Over the 7-year period of our analytic sample, we observe 8,650 births attended by 496 physicians across 37 catchment areas.

MEASURES

Geographic Units of Analysis

Unlike much of the variation literature which uses Hospital Referral Regions (HRRs) as the unit of analysis, we utilize administratively-determined TRICARE catchment areas as they are the relevant geographic unit for DoD policymakers. Since we are analyzing direct care encounters exclusively, the clinical encounters we observe all occur in military community hospitals or facilities). Unlike HRRs, multiple catchment areas can represent the same healthcare market for tertiary hospital care and are closer to measuring changes at the facility level than at the market level. While catchments may result in more precise measurements of physician behavior than aggregated HRR measures, they could be more susceptible to a single physician influencing our characterization of regional practice patterns. Individual behavior is often highly correlated to group behavior leading to misleading conclusions about true associations (4). We address this by calculating leave-out averages of C-section rates for each catchment area which exclude a physician's own patients when calculating Δ_j for our

estimating equations (56). MTF coordinates were geocoded using street addresses from the Uniform Business Office DMIS Master File and the Google Maps Developer API.

Risk Adjusting Cesarean Section Rates

Our primary outcome variable is comprised of a physician's cesarean section rate measured from 0 to 1. We analyze changes to this rate over time and across practice settings. To account for regional differences in patient severity, we use a logit model of the probability of C-section controlling for established medical risk factors (23; 31). This probability measure is purely based on clinical presentation of comorbid conditions and not any socio-economic factors. Many of these risk factors are either emergent, such as antepartum hemorrhage or a malposition fetus, or are subject to less clinical uncertainty, such as severe hypertension or preterm gestation, and therefore limit concerns over regional diagnostic intensity biases raised by Song et al (67). We use these clinical factors to adjust catchment C-section rates for the underlying health of the population and to allow for fair comparisons across regions. We use our C-section risk values from our logit model to risk adjust crude and leave-out catchment C-section rates when calculating Δ_j for our model. We use the logit model to calculate predicted values which we then use to risk-adjust catchment C-section rates.

We estimate equations using average estimates of catchment C-section rates over our samples period. Constant estimates are advantageous in that they can be calculated over a larger sample giving more precise estimates, but they will not pickup differential changes to C-section growth rates over time. This could occur, for instance, if catchments anchored by medical centers perform more C-sections over time than do catchments anchored by community hospitals. As noted in Molitor (56), rates in levels are not

directly comparable across years due to secular trends, but regional differences in rates are relatively constant across years. C-section rates increase from 30.4% in 2007 to 37.8% in 2014, while the average differences in catchment rates remain relatively constant between 5.0% (2011) – 7.7% (2008).

ANALYSIS

We first report descriptive statistics to describe characteristics of our physician cohort stratified by “mover” status in Table 7. To provide a better picture of our risk-adjustment, we present coefficients, standard errors, and marginal effects for our logit model of the clinical probability of C-section in Table 8. Overall, this model predicts well with a McFadden’s pseudo R-squared of .32. We then use these expected C-section risk to adjust C-section rates using a mixed-effects linear model to investigate variation in the cross section among non-movers in Table 9. In this model, we include fixed effects for calendar year and control for C-section risk using predicted values from the logit model. To account for clustering and any small samples, we also include random effects terms at the facility and individual physician levels. We then examine variation in average adjusted C-section rates at the facility level in Figure 4 and in adjusted rates over time in Figure 5.

Our main results come from our differences in differences analysis. First we calculate Δ_j using constant risk-adjusted leave-out catchment average C-section rates which excludes a physician’s own patients. Figure 6 shows the distribution of Δ_j , our measure of changes to a physician’s practice environment, calculated by subtracting the average C-section rate of the physician’s origin catchment from that of their destination

catchment. To avoid the spurious correlation issues previously discussed, our estimating equations use leave-out averages to characterize Δ_j .

We summarize results from multiple specifications of our model in Table 10. Each column represents a separate regression. All models have calendar year fixed effects, year relative to move fixed effects, the same set of patient clinical risk adjustment dummies used in the logit model, and time-variant physician experience bins. Alternate specifications are summarized in the table. Column 1 includes origin catchment fixed-effects and explicitly controls for a mover's Δ_j . This specification allows us to observe any “positive matching” of practice styles. A positive, significant coefficient of this term would indicate that a mover has a more intense practice style upon moving to a higher-intensity region. In Column 2, we instead include physician fixed effects. This specification effectively investigates within variation using a physician's behavior before the move as a baseline comparison to after the move. The regression in Column three builds upon column one but includes additional physician observables including dummy variables for whether a physician is a MD or DO, a foreign medical school graduate, medical specialty, and Service branch (Army, Navy or Air Force). Columns 4 and 5 include physician fixed effects but stratify movers by years of clinical experience. Column 4 presents a regression upon movers with less than 10 years of experience and Column 5 uses a population with 10 years or more of experience. In all regressions, we calculate clustered standard errors at the physician and catchment levels.

RESULTS

Sample Demographics

Characteristics of our final analytic sample of physicians are summarized in Table 7 stratified by “moving” status. Note that the panel is unbalanced and not all physicians perform C-sections in a given year. Movers are slightly more likely to hold MDs as opposed to DO medical degrees and are on average less experienced at the time of delivery. Movers also attended more deliveries on average than non-movers. Movers moved an average of 1,289 miles.

Risk Adjusting

We present results from our logit model used to risk adjust catchment C-section rates for the health of the underlying population. We present odds ratios, their respective standard errors, and the calculated marginal effects in Table 8. All variables used in the model are presented in the table and are based upon risk factors previously used in the literature (5; 23; 32). As the model shows, a previous C-section, disproportioned fetus, malpositioned fetus, and antepartum hemorrhage are the greatest predictors of a C-section. We then use the logit model to calculate predicted values which we then use to risk-adjust catchment C-section rates. This model is based upon clinical risk factors alone and measures observed clinical practice in our sample.

Our cross-sectional model of non-movers is presented in Table 9. The random-effects terms suggest that differences between facilities and physicians explain little of the unexplained variance in C-section rates. We then use risk-adjusted rates from this equation to examine variation in C-section rates among non-movers in the cross-section. We find evidence of moderate variation across facilities (16% CV). Furthermore, several

facilities have consistently higher or lower C-section rates. We present raw and risk-adjusted catchment average C-section rates by calendar year in Figure 5. These rates from our sample increased over time from an average of 32.6% in 2007 to 37.7% in 2014 following the larger secular trends in growth. This places our rates in-line with the broader civilian population which saw increases from 20.7% in 1996 to 32.7% in 2013 (59). In Figure 6, we show time-invariant averages of the differences between a mover's destination and origin catchments, Δ_j . We present the leave-out risk adjusted Δ_j for movers which we use for estimating equations. This measure excludes each moving physician's own patients when calculating catchment averages. For comparison, we also present risk-adjusted Δ_j which do include a moving physician's patients.

Differences in Differences Analysis

Our main results from the differences in differences analysis is presented in Table 10. The first column presents estimates of changes to practice style relative to other migrants from the same catchment of origin. The Δ_j coefficient presented in the first row estimates the degree of selective migration. In other words, a positive coefficient of this term would suggest that physicians who move to higher intensity catchment perform more intensely relative to their peers before the move. In our model, the Δ_j coefficient of .105 was suggestive of very mild selective migration, but failed to be statistically significant. The $\Delta_j \times \text{Post-Move}$ interaction term is our parameter of interest. It measures how much a physician changes their C-section practice style in response to a change in practice environment. We estimate an effect of .351 which means that for every 1 percent change in the C-section intensity of a physician's environment upon a move corresponds to a .351 percent change in a physician's own C-section behavior.

Column 2 presents a regression which includes physician fixed effects in place of Δ_j and origin catchment fixed effects to test within variation using each physician's behavior before the move as a baseline. We estimate a difference in differences coefficient of .282 which is similar in magnitude and interpretation to the estimate from Column 1. A one percent change in practice environment is associated with a change in physician C-section behavior of .282 percent. Our third regression specification leverages physician characteristics from CCQAS to explicitly estimate physician characteristics and incorporate human capital changes over time such as additional clinical training or any changes in specialty in addition to clinical experience. The coefficient of the parameter of interest is .372 and again in-line with our other specifications. Also of note, when controlling for additional physician characteristics, the Δ_j term indicative of selective migration practically falls to zero. The final two columns present regressions which stratify movers by their years of post-graduate experience. Column 4 presents evidence from movers with less than 10 years of clinical experience and finds no significant evidence of changes in practice style across move. The last regression in Column 5 looks at more experienced movers exclusively. We estimate a difference-in-differences coefficient of .9 suggesting that more senior physicians alter their practice style almost one to one with changes in their practice environment.

DISCUSSION

This work seeks to separate physician-level factors from factors which derive from a physician's practice environment in order to explain variations in practice style across regions. Environmental factors could include hospital capacity and bed availability, availability of specialists, or competition between hospitals. Physician

factors could include physician beliefs about treatment, education and training, or learned experience. We find that physicians modify their C-section behavior in response to new environments when controlling for patient health. We estimate that physician behavior changes roughly .28-.37 percentage points for every 1 percent change in practice environment. These results suggest that around 28-37 percent of regional variations in physician practice patterns could be explained by environmental factors unique to the practice setting. The difference in changes across moves can be attributed to physician-specific factors. Molitor (56) similarly investigated migrating cardiologists who changed their catheterization behavior by .6 -.8 percent for every 1 percent change in Hospital Referral Region practice intensity.

The remainder of the variation unexplained by the practice environment could be due to a variety of factors. Cutler et al (24) demonstrated that physician beliefs are an important driver of variations. Physician training and education has been shown to influence treatment costs and diagnostic efficiency (30). Evidence also suggests that practice style is persistent and learned, with less variation among specialists, when investigating hospital house staff in-training (16). Taken together, our results affirm that environmental factors do influence provider behavior, but that it is generally not a one-for-one response to their practice setting. The differences in magnitude of effect between our study and its nearest comparison of cardiologists could be due to differences in treating Acute Myocardial Infarction which require specialists and specialized laboratories, the fee for service environment of Medicare, or differences in practice environments between civilian and military healthcare.

This work adds evidence of persistent place-effects for a widely-used surgical procedure which has been the target for reductions by policy-makers. Within the Military Health System, obstetrics is a large service-line and a significant cost driver. Our results demonstrate that physician behavior changes in response to a new clinical environment. An alternative interpretation of the results is that a large degree of physician behavior is also modifiable. When examining medical risk and whether a mother received C-section in our study, we observed high-risk mothers who did not receive C-sections and other low-risk mothers who did, consistent with work by Currie and MacLeod (23). This suggests that perhaps procedures could be better allocated to patients and that there is room to modify physician decision making.

Policymakers should consider that physicians and the practice environment are important sources of variation. While researchers and policymakers tend to observe data in the aggregate, practicing physicians generally only observe their patients and their own decision-making. Providing standardized decision-making aids such as simple algorithms could be a cost-effective way to enhance and complement physician expertise to reduce unwarranted variations and improve outcomes.

There are several limitations in this study. Owing to data restrictions associated with the de-identification process, encounter data was limited to calendar year which may cause additional noise around moves. Furthermore, key patient demographics such as age, rank, and sponsor status were excluded to eliminate concerns over re-identification by triangulation; age in particular is an important risk factor for C-section. We addressed this limitation in part by including medical codes into our logit risk model for elderly primigravida and multigravida though these conditions may be inconsistently recorded in

medical records. It is also worth mentioning that physicians in this sample are military officers, and thus are often required to attend to duties that pull them away from clinical care, leading to lighter caseloads than civilian counterparts. Furthermore, the three branches of the services over this time-period developed training agreements with civilian facilities whereby military physicians practice in civilian facilities, yet we do not observe the care provided in civilian facilities. These factors are likely more important for studying workload comparisons to the civilian sector, but it is not clear that they would bias our results.

CONCLUSION

Our study shows that the physician factors are a predominant driver of practice variations across geographic regions. By contrast, environmental factors unique to place explain about 30 percent of variations in practice styles. Our study suggests that policies aimed at augmenting physician decision-making could reduce unwarranted variations while improving outcomes and reducing unnecessary expenditures.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Alan Sim, Lee Ann Zarzabal, and Alexander Rittel at the Defense Health Agency Solutions Delivery Division for their assistance pulling MDR data. We would also like to thank Dr. Andrew Findley and the CCQAS Service representatives for making their data available.

Table 7. Summary Characteristics of Sampled Military Physicians

	Physician Characteristics				Number of Encounters			
	Non-Mover		Mover		Non-Mover		Mover	
M.D.	314	79.49%	87	84.47%	4526	85%	2954	89%
D.O.	81	20.51%	16	15.53%	806	15%	364	11%
Total	395	100%	103	100%	5332	100%	3318	100%
Foreign Medical Graduate	3	0.70%	0	0.00%	88	2%	0	0%
Highest Rank Attained:								
0-3	17	4.30%	0	0.00%	39	0.73%	0	0.00%
0-4	161	40.76%	39	37.86%	1338	25.09%	862	25.98%
0-5	118	29.87%	46	44.66%	2046	38.37%	1858	56.00%
0-6	99	25.06%	18	17.48%	1909	35.80%	598	18.02%
Total	395	100%	103	100%	5332	100%	3318	100%
Service:								
Air Force	96	24.30%	25	24.27%	1159	21.74%	911	27.46%
Army	163	41.27%	39	37.86%	2084	39.08%	806	24.29%
Navy	136	34.43%	39	37.86%	2089	39.18%	1601	48.25%
Total	395	100%	103	100%	5332	100%	3318	100%
Average Years Post-Medical School	Mean	SD	Mean	SD				
	10.79	5.56	9.92	4.93				
Average number of Births Attended	Mean	SD	Mean	SD				
	52.62	46.77	70.22	41.72				
Average Move Distance (miles)	0	0	1288.92	1012.959				

Table 8. Logistic Regression Model of C-Section Risk Factors

Note: Table is based upon a logit model of patient risk. Marginal effects measure discrete change in the predicted probability as the binary independent variable changes from 0 to 1. Number of observations= 8,650

Logit Model of C-section Risk			
	Odds Ratio	SE	Marginal Effect
Previous C-section	9.737	1.66	0.441
Malpositioned Fetus	4.520	0.432	0.248
Antepartum Hemorrhage	4.168	1.064	0.248
Hypertension Complication	1.719	0.142	0.085
Pre-term Labor	1.390	0.160	0.041
Prolonged/Late Labor	1.422	0.136	0.036
Excessive Edema	1.176	0.275	0.029
Maternal Infection	1.154	0.193	0.011
Diabetes	3.394	0.875	0.2
Cardiovascular Disorder	0.814	0.691	0.01
Multiple Gestation	3.035	0.875	0.188
Disproportion of Fetus	18.908	10.682	0.478
Soft Tissue Abnormality	3.618	0.495	0.242
Fetal Abnormality	1.100	0.291	0.022
Polyhydramnios	2.398	0.631	0.143
Oligohydramnios	1.431	0.244	0.053
Membrane Rupture	0.888	0.121	-0.017
Amniotic Infection	3.074	0.523	0.184
Advanced Maternal Age (>35)	0.312	.229	-0.141
Asthma	0.840	.121	-0.009
Herpes	1.259	.339	0.062

Table 9. Cross-sectional Model of Adjusted C-Section Rates

Note: We present our mixed-effects model used to characterize C-section rates in the cross-section. We include fixed effects for year and C-section risk using output from our logit model. We include random effect terms at the facility and individual levels. We estimate this equation on all non-movers in our sample n=4,520 births.

	Coefficient	SE	P-value
Calendar Year			
2008	-0.006	0.021	0.784
2009	0.009	0.022	0.664
2010	0.004	0.022	0.852
2011	-0.009	0.023	0.711
2012	-0.015	0.023	0.495
2013	-0.02	0.022	0.366
2014	-0.036	0.024	0.136
Predicted C-Section Risk	0.98	0.018	<0.001
Constant	0.027	0.018	0.142
Random-Effects Levels			
	Variance	SE	
Treatment Catchment	0.001	0.0007	
Physician	0.0004	0.0007	
Residual	0.139	0.003	

Figure 4. Variation in Average C-Section Rates by Facility

Note: We use our cross-sectional model to calculate adjusted C-section rates to characterize practice patterns across facilities. Each observation is the average adjusted C-section rate with corresponding 95% confidence intervals. The average adjusted rate is .363 across all facilities. The Coefficient of Variation for C-sections across facilities is 16%. We estimate this equation on all non-movers in our sample $n=4,520$ births for all facilities.

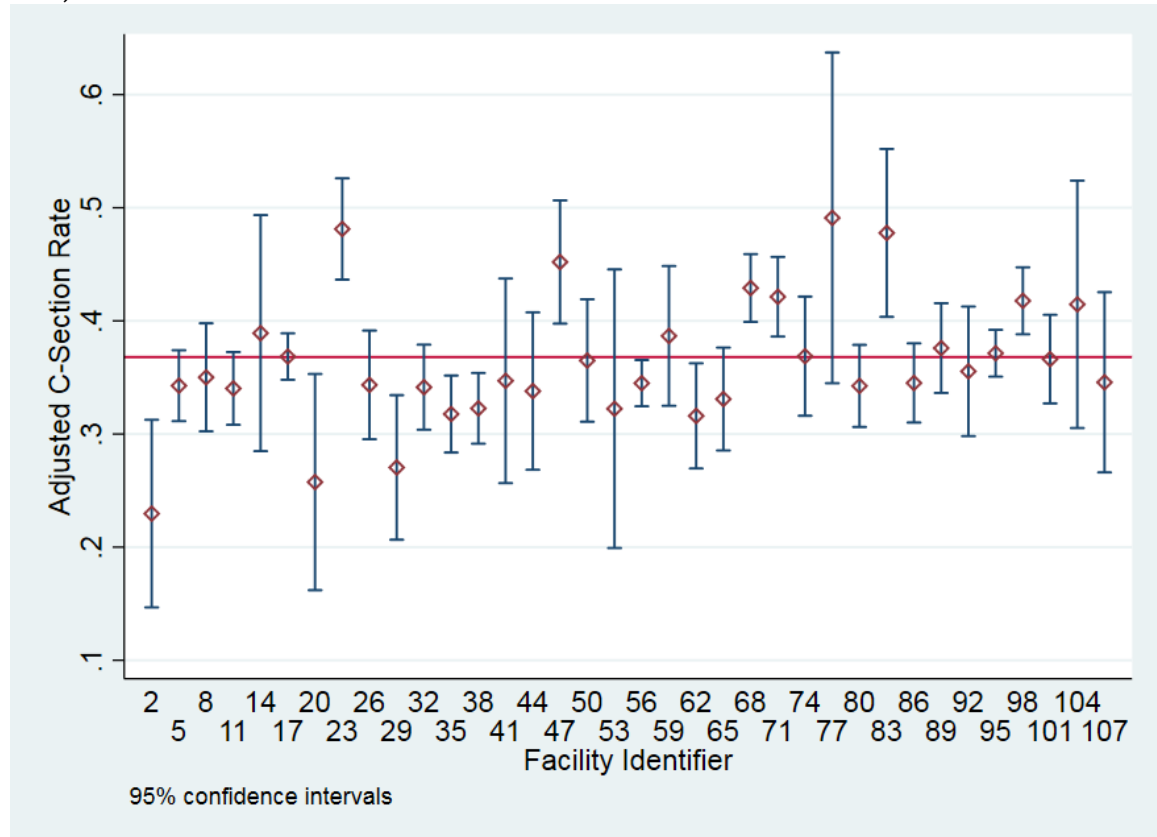


Figure 5. C-Section Rates over Time

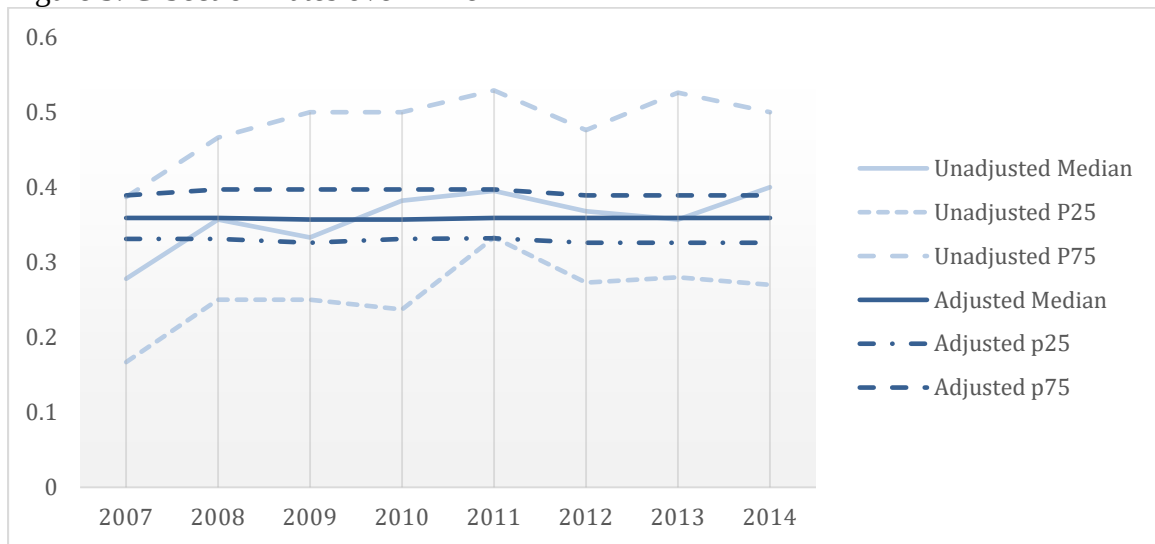


Figure 6. Distribution of Changes to Cesarean Section Environments Across Moves
 Note: We present risk-adjusted differences as well as the leave-out means which we take to regression equations.

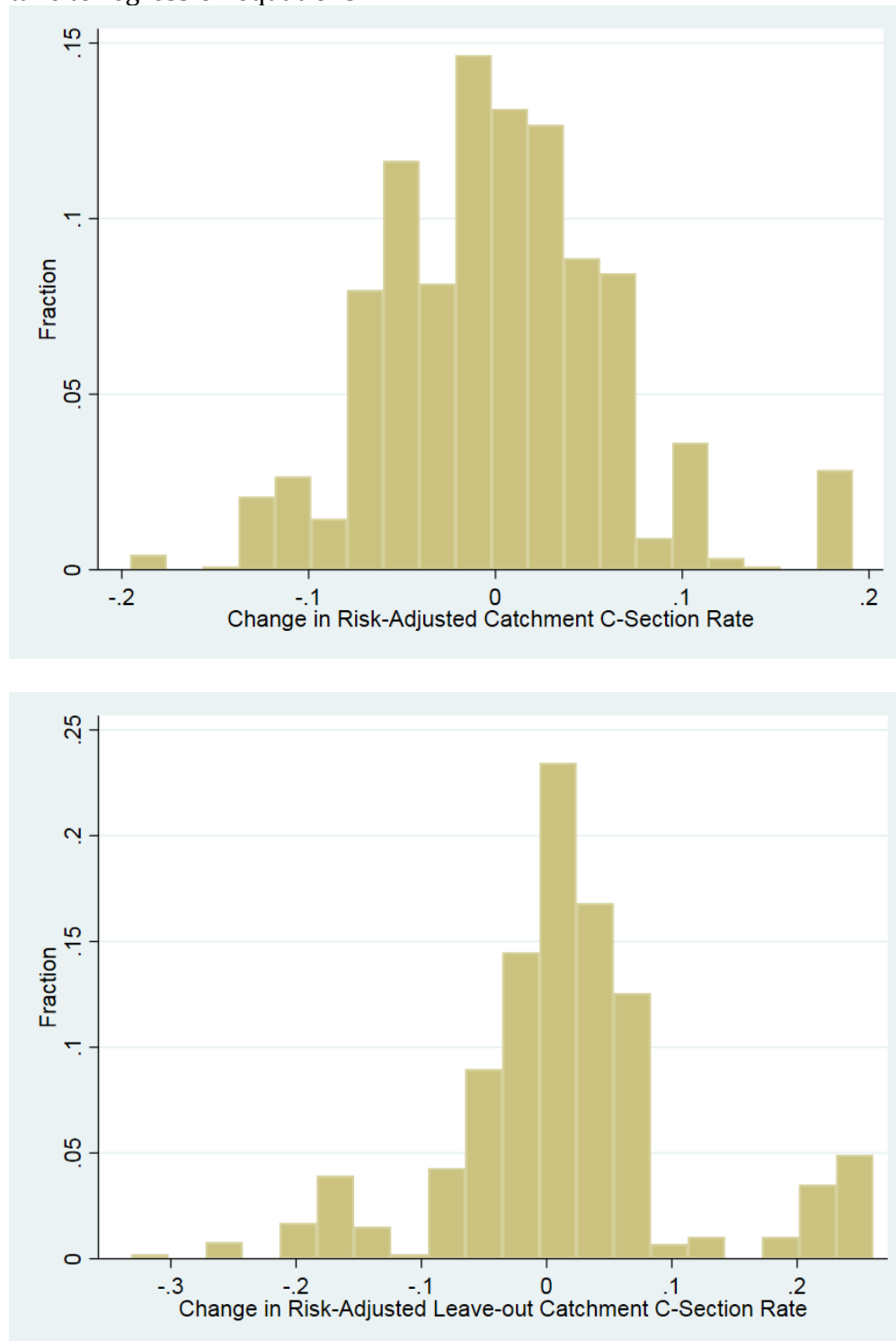


Table 10. Difference-in-Differences Analysis

Note: All equations contain calendar year fixed effects, year relative to move fixed effects, and a panel of patient risk adjusters used in the logit model. Each column represents a separate regression.

	All Movers	All Movers	All Movers	Less Experienced Movers (<10 yrs. of clinical exp.)	Experienced Movers (10+ yrs. of clinical exp.)
Δ_j	0.112		-0.051		
SE	(0.148)		(0.108)		
$\Delta_j \times T_{post-move}$	0.413	0.243	0.394	-0.166	0.9
SE	(0.262)	(0.268)	(0.262)	(0.215)	(0.417)
Origin Catchment FE					
Destination Catchment FE	X		X		
Physician FE		X		X	X
Additional Physician Controls			X		
Observations	2,498	2,498	2,498	1,291	1,207
Adjusted R-squared	0.3857	0.3882	0.3901	0.407	0.3732

Chapter 5: Conclusion and Discussion

BRIEF OVERVIEW OF MAJOR FINDINGS

In the first empirical paper, I use migrating patients to estimate variations in spending and utilization across TRICARE catchment areas. I find evidence of considerable variations across regions (up to 70% when comparing quintiles). Using a “mover” strategy, we can attribute 80-90 percent of the variation to place-specific factors. The remaining 10-20 percent is attributable to patient factors such as health status or patient preferences for care. Movers, when they relocate, almost immediately modify their health care utilization and spending to more closely resemble their environment. The event study graphically depicts this change across a variety of moves from high to low utilization or low to high utilization. Our results are also robust across the three main beneficiary categories of active duty, dependents, and retirees which show varying degrees of utilization differences but a persistently high place-share. The geographic variations in the MHS are striking and suggest that reducing spending to the level of the lowest quintile could potentially result in 32% savings system-wide. I also find that the variation persists when examining utilization instead of variation. This means that certain regions are providing more or less care. Future research should focus on measuring outcomes to understand and quantify returns to spending.

The second and third papers sought to better understand potential mechanisms that could cause differences across regions. I first focus on bariatric surgeries to better understand the wide magnitude in variations in surgical costs and complication rates. I study a sample of 38 uniformed surgeons who perform over 1,200 surgeries over 7 years and find that complication rates vary widely across facilities (49% CV) as do inpatient

costs (25.9% CV). Controlling for patient health explains some of the variations, but differences across facilities and individual surgeons persist. On average, complications lead to 32% higher inpatient costs than a complication-free surgery. The complication rate in this sample (15.5%) is comparable to rates among civilian centers (12.3-18.7%). These findings suggest that differences in surgical quality could at least in part explain spending differences across regions. Furthermore, efforts to improve surgical quality could lead to financial savings for the MHS while providing patients top-quality care and addressing the medical readiness of our surgeons.

The final paper studied physicians attending births to understand how the practice environment impacts their clinical decision-making using the case of Cesarean sections. Again using a mover strategy, I find that physicians modify their C-section rate in response to a new environment. I estimate that physician behavior changes .28-.37 percentage points for every 1% change in practice environment. This suggests that about a third of regional variations in practice patterns could be explained by environmental factors unique to place. I also find that more experienced physicians are more sensitive to changes in the practice environment and tend to change more than their less experienced peers.

Taken together, the results of this work suggest that despite the centrally administered, highly managed nature of the MHS large variations persist throughout the health system. Supply-side factors unique to place seem to determine much of the care provided for both patients and physicians. These factors may work through different mechanisms such as complication rates, procedure choice, physician beliefs about treatment, or differences in practice environments.

The efficiency and quality of facilities across the network vary substantially. Reliable differentiating between high- and low-performing hospitals can help allocate resources and interventions to improve quality, such as reducing surgical complications, or to decrease utilization of low-value services, such as unwarranted C-sections. Understanding what efficient hospitals are doing right is just as valuable as uncovering what low-performers could be doing better. Evidence using the case of bariatric surgery, however, suggests that reducing variations in care could be complementary to other operational goals. For instance, initiatives to improve surgical quality will not only reduce complication-related expenditures, but they will also lead to better trained surgeons who can provide better care to our beneficiary population. It is one of the rare instances when the “business case” aligns with broader objectives of the organization.

POLICY RECOMMENDATIONS

This research documents the magnitude of spending variations and the magnitude of possible savings to the MHS. Reducing expenditures to match the spending levels of regions in the lowest quintile will reduce MHS patient-care costs by 32 percent, roughly translating to approximately 8 billion dollars in savings. While this estimate does not account for quality of care, it does provide an upper-bound of the savings opportunities that could be realized through improvements to operational efficiency and the management of our beneficiary population. To realize these savings, policymakers should shift the focus of health care reforms to target supply-side determinants of variations. Not all determinants of spending are modifiable factors, nor can they be modified given the strategic mission of the MHS. For instance, higher spending in one catchment may be due to the nature of the work performed at the base (such as combat or specialty training

which may lead to injury) and not any inefficiency of the health system. However, many determinants can be changed to optimize spending and outcomes. Researchers should systematically evaluate hypothesized determinants of variations to help policymakers efficiently allocate resources and set operational goals.

Modifiable supply-side factors divide into two broad categories: provider-level factors and place-specific factors. At the provider-level, research suggests that health care providers play a significant role in determining a patient's care. Particularly under the highly-managed nature of the MHS, human capital factors such as provider beliefs about treatment, provider practice style, and coordination of care could play a sizeable role in driving variations across regions. Understanding these mechanisms can inform clinical practice by improving training and continuing education programs as well as clinical decision-making tools such as algorithms or standard operating procedures. These insights could also help MTF commanders to better understand how to optimize the management of their staff and facilities to meet peer benchmarks or organizational goals.

Given that the majority of providers in the MHS also rotate duty stations, place-specific determinants of variations are also an important consideration for policymakers. These factors could include organizational pressures, the distribution of human and physical capital, or the competitiveness of the local market. For instance, larger MTFs carry more overhead costs which are subsequently allocated to patient-level encounters. Policymakers could spread out overhead costs by recapturing volume to larger MTFs without sacrificing quality or access. The distribution of capital infrastructure should also be regularly evaluated to ensure that the system is maximizing the value of medical resources and efficiently serving the beneficiary population. Similarly, policymakers

could leverage purchasing power to ensure TRICARE support contractors are using data to optimize purchased-care provider networks to meet the needs of the local beneficiary population.

From a system perspective, as health care continues to increase in complexity, the MHS will need to continue to invest in data systems to support analytics and decision-making under the shared services model. Now, that health care has been centralized under the Defense Health Agency, operating in this joint environment will require an expansion of standardized and flexible data systems to answer critical policy questions. Policymakers should remove data siloes and invest in integrated systems to improve data integrity, safety, and access.

Future Research

This work points to the importance of supply-side determinants of variations. A comprehensive grasp of these factors will require further study incorporating outcomes and quality measures across facilities. While understanding the magnitude and determinants of variations is an important first step, policymakers will need to understand the mechanisms driving the variations in order to effective policy. The MHS, however, is uniquely positioned to isolate and empirically test many hypothesized causal mechanisms in a data-rich environment by using quasi-experimental designs. The most pressing question given this research is to understand whether, and why, health outcomes differ across regions, and whether such outcomes are related systematically to geographic variations in clinical inputs and quality of care.

For example, differences between the direct care and purchased care systems might allow researchers to begin to answer these questions. For instance, researchers

could test the physician induced demand hypothesis by comparing care across systems. Physician compensation in the direct care system is not tied in any way to the number or type of procedure performed, but civilian counterparts from the purchased care sector are reimbursed through a fee-for-service scheme indexed to Medicare reimbursement.

Furthermore, reimbursement under TRICARE Prime differs from a traditional HMO in that individual physicians do not face financial risk. One could ask how the variation patterns might differ if physicians or facilities did assume partial or full financial risk for their panel of patients. Second, researchers could investigate defensive medicine in treatment decision making, physicians practicing out of MTFs are free from medical malpractice liability while their civilian counterparts in the community are still subject to their state's medical tort law. Third, with additional fee-for-service data, it is possible to compare variations across differences in insurance type at a national level. While this research will need to overcome selection issues encompassing both patient insurance preferences and the choice of care setting (purchased care vs. direct care), it could provide valuable contributions to the variation literature.

Future variation research in the MHS should prioritize understanding quality and value of care. Most pressing, 2017 National Defense Authorization Act (NDAA) provisions call for the central management of MTFs under the Defense Health Agency. This work studies a time period before the DHA reached full operation. In the coming years, continued study of variations can provide valuable insight into the standardization and streamlining of operations across the enterprise and potentially justify the existence of the agency as a whole. Furthermore, as policymakers within the DoD consider incorporating alternative payment models such as value-based insurance design, bundled

payments, and accountable care organizations, further research should help clarify implementation of these programs in the MHS and indeed if savings and quality improvement are even attainable through these policy tools. Researchers should also test for provider characteristics that may be associated with the provision of high-value, high-quality care. Results for these studies would have direct implications for training and readiness in the MHS particularly for the large graduate medical education programs operating in our medical facilities. Understanding the relationship between provider human capital and patient outcomes is a burgeoning field in the literature due to the availability of novel data sources.

Finally, these questions can be answered only if datasets in the MHS are more freely available for this type of research. Researchers could work towards these goals by helping policymakers understand the value of the data and articulating the types of questions that can be answered. In sum, researchers are just now discovering the tip of the iceberg in these initial studies using the MHS. Much work remains to be done to truly understand the dynamics of delivering care to this unique population of beneficiaries and how best to deliver and pay for that care.

REFERENCES

1. Abe D, D. GS, H. ZS. 2018. Adjusting Health Expenditures for Inflation: A Review of Measures for Health Services Research in the United States. *Health Services Research* 53:175-96
2. Adesoye T, Kimsey LG, Lipsitz SR, Nguyen LL, Goodney P, et al. 2017. Geographic variation in Medicare and the military healthcare system. *The American journal of managed care* 23:e259-e64
3. Agha L, Frandsen B, Rebitzer J. 2017. Causes and Consequences of Fragmented Care Delivery: Theory, Evidence, and Public Policy. *NBER Working Paper*
4. Angrist JD. 2014. The perils of peer effects. *Labour Economics* 30:98-108
5. Aron DC, Harper DL, Shepardson LB, Rosenthal GE. 1998. Impact of risk-adjusting cesarean delivery rates when reporting hospital performance. *Jama* 279:1968-72
6. Baicker K, Buckles KS, Chandra A. 2006. Geographic Variation In The Appropriate Use Of Cesarean Delivery. *Health Affairs* 25:w355-w67
7. Baicker K, Fisher ES, Chandra A. 2007. Malpractice liability costs and the practice of medicine in the Medicare program. *Health Aff (Millwood)* 26:841-52
8. Baker LC, Bundorf MK, Kessler DP. 2014. Patients' Preferences Explain A Small But Significant Share Of Regional Variation In Medicare Spending. *Health affairs (Project Hope)* 33:957-63
9. Banka G, Woodard G, Hernandez-Boussard T, Morton JM. 2012. Laparoscopic vs open gastric bypass surgery: differences in patient demographics, safety, and outcomes. *Archives of surgery (Chicago, Ill. : 1960)* 147:550-6
10. Barnato AE, Herndon MB, Anthony DL, Gallagher PM, Skinner JS, et al. 2007. Are regional variations in end-of-life care intensity explained by patient preferences?: A Study of the US Medicare Population. *Medical care* 45:386-93
11. Bickett T, Cheng SL, Jaditz T, Miller R. 2010. Small Area Variation: Neonatal Intensive Care Unites, Hysterectomies, and Total Cost of Care for the Non-catchment Area Population. CNA
12. Birkmeyer JD, Gust C, Dimick JB, Birkmeyer NJ, Skinner JS. 2012. Hospital quality and the cost of inpatient surgery in the United States. *Annals of surgery* 255:1-5
13. Brevig H, Colosimo C, Jaditz T, Krauss R, Mandell K, et al. 2014. The Quality-Volume Relationship: Comparing Civilian and MHS Practice, CNA Analysis and Solutions
14. Bronnenberg BJ, Dubé J-PH, Gentzkow M. 2012. The Evolution of Brand Preferences: Evidence from Consumer Migration. *American Economic Review* 102:2472-508
15. Bureau of Labor Statistics. 2018. *Databases, Tables & Calculators by Subject*. <https://data.bls.gov/pdq/SurveyOutputServlet>
16. Chan D. 2016. Informational Frictions and Practice Variation: Evidence from Physicians in Training. *NBER Working Paper*
17. Chandra A, Skinner J. 2012. Technology Growth and Expenditure Growth in Health Care. *Journal of Economic Literature* 50:645-80

18. Chandra A, Staiger DO. 2007. PRODUCTIVITY SPILLOVERS IN HEALTHCARE: EVIDENCE FROM THE TREATMENT OF HEART ATTACKS. *The journal of political economy* 115:103-40
 19. Chernew ME, Sabik LM, Chandra A, Gibson TB, Newhouse JP. 2010. Geographic correlation between large-firm commercial spending and Medicare spending. *The American journal of managed care* 16:131-8
 20. Committee on Geographic Variation in Health Care S, Promotion of High-Value C, Board on Health Care S, Institute of M. 2013. In *Variation in Health Care Spending: Target Decision Making, Not Geography*, ed. JP Newhouse, AM Garber, RP Graham, MA McCoy, M Mancher, A Kibria. Washington (DC): National Academies Press (US)
- Copyright 2013 by the National Academy of Sciences. All rights reserved. Number of.
21. Cooper Z, Craig S, Gaynor M, Van Reenen J. 2015. The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured. *NBER Working Paper*
 22. Currie J, MacLeod WB. 2008. First Do No Harm? Tort Reform and Birth Outcomes*. *The Quarterly Journal of Economics* 123:795-830
 23. Currie J, MacLeod WB. 2016. Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians. *Journal of Labor Economics* 35:1-43
 24. Cutler D, Skinner J, Stern AD, Wennberg D. 2013. Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending. *NBER Working Paper* No 19320
 25. Defense Health Agency. 2014. Military Health System Review - Final Report, Defense Health Agency
 26. Department of Defense. 2016. Evaluation of the TRICARE Program Fiscal Year 2016 Report to Congress, Department of Defense
 27. Department of the Army. 2007. Army Regulation 600-8-11 Reassignment. Washington, DC: Department of the Army
 28. Dimick JB, Nicholas LH, Ryan AM, Thumma JR, Birkmeyer JD. 2013. Bariatric Surgery Complications Before vs. After Implementation of a National Policy Restricting Coverage to Centers of Excellence. *JAMA : the journal of the American Medical Association* 309:792-9
 29. Dimick JB, Osborne NH, Nicholas L, Birkmeyer JD. 2009. Identifying high-quality bariatric surgery centers: hospital volume or risk-adjusted outcomes? *J Am Coll Surg* 209:702-6
 30. Doyle J, Ewer S, Wagner T. 2010. Returns to Physician Human Capital: Evidence from Patients Randomized to Physician Teams *National Bureau of Economic Research*
 31. Epstein AJ, Nicholson S. 2009. The formation and evolution of physician treatment styles: An application to cesarean sections. *Journal of Health Economics* 28:1126-40
 32. Epstein AJ, Nicholson S. 2009. The formation and evolution of physician treatment styles: an application to cesarean sections. *J Health Econ* 28:1126-40

33. Finkelstein A, Gentzkow M, Williams H. 2016. Sources of Geographic Variation in Health Care: Evidence from Patient Migration. *National Bureau of Economic Research*
34. Finkelstein A, Gentzkow M, Williams H. 2016. Sources of Geographic Variation in Health Care: Evidence From Patient Migration*. *The Quarterly Journal of Economics* 131:1681-726
35. Fisher ES, Bynum JP, Skinner JS. 2009. Slowing the growth of health care costs--lessons from regional variation. *The New England journal of medicine* 360:849-52
36. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder É L. 2003. The implications of regional variations in medicare spending, part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine* 138:288-98
37. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder ÉL. 2003. The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care. *Annals of Internal Medicine* 138:273-87
38. Fisher ES, Wennberg JE, Stukel TA, Skinner JS, Sharp SM, et al. 2000. Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. *Health Services Research* 34:1351-62
39. Gottlieb DJ, Zhou W, Song Y, Andrews KG, Skinner JS, Sutherland JM. 2010. Prices Don't Drive Regional Medicare Spending Variations. *Health affairs (Project Hope)* 29:537-43
40. Gruber J, Kim J, Mayzlin D. 1999. Physician fees and procedure intensity: the case of cesarean delivery. *J Health Econ* 18:473-90
41. Gruber J, Owings M. 1996. Physician Financial Incentives and Cesarean Section Delivery. *The RAND Journal of Economics* 27:99-123
42. Healy MA, Mullard AJ, Campbell DA, Jr, Dimick JB. 2016. Hospital and payer costs associated with surgical complications. *JAMA Surgery* 151:823-30
43. Ibrahim AM, Ghaferi AA, Thumma JR, Dimick JB. 2017. Variation in Outcomes at Bariatric Surgery Centers of Excellence. *JAMA Surgery* 152:629-36
44. Jaditz T, Bickett T, Foster L, Farr D, Berens R, Morrow R. 2008. Examples of Practice Variation Across Navy MTFs, CNA
45. Jaditz T, Haglund E, LaBonte C, Miller R. 2009. Practice Variation Across Navy MTFs: Cesarean Sections, Hip Replacements, Knee Replacements, Ischemic Heart Disease, CNA
46. Johnson EM, Rehavi MM. 2016. Physicians Treating Physicians: Information and Incentives in Childbirth. *American Economic Journal: Economic Policy* 8:115-41
47. Khan S, Rock K, Baskara A, Qu W, Nazzal M, Ortiz J. 2016. Trends in bariatric surgery from 2008 to 2012. *American journal of surgery* 211:1041-6
48. Kimsey L, Olaiya S, Smith C, Hoburg A, Lipsitz SR, et al. 2017. Geographic variation within the military health system. *BMC Health Services Research* 17:271
49. Kizy S, Jahansouz C, Downey MC, Hevelone N, Ikramuddin S, Leslie D. 2017. National Trends in Bariatric Surgery 2012–2015: Demographics, Procedure Selection, Readmissions, and Cost. *Obesity Surgery* 27:2933-9

50. Kozhimannil KB, Law MR, Virnig BA. 2013. Cesarean Delivery Rates Vary 10-Fold Among US Hospitals; Reducing Variation May Address Quality, Cost Issues. *Health affairs (Project Hope)* 32:527-35
51. Krell RW, Finks JF, English WJ, Dimick JB. 2014. Profiling Hospitals on Bariatric Surgery Quality: Which Outcomes Are Most Reliable? *Journal of the American College of Surgeons* 219:725-34.e3
52. Lleras-Muney A. 2010. The Needs of the Army: Using Compulsory Relocation in the Military to Estimate the Effect of Air Pollutants on Children's Health. *Journal of Human Resources* 45:549-90
53. McKellar MR, Landrum MB, Gibson TB, Landon BE, Fendrick AM, Chernew ME. 2017. Geographic Variation in Quality of Care for Commercially Insured Patients. *Health Services Research* 52:849-62
54. McKellar MR, Naimier S, Landrum MB, Gibson TB, Chandra A, Chernew M. 2014. Insurer Market Structure and Variation in Commercial Health Care Spending. *Health Services Research* 49:878-92
55. Mihaylova B, Briggs A, O'Hagan A, Thompson SG. 2011. Review of Statistical Methods for Analysing Healthcare Resources and Costs. *Health Economics* 20:897-916
56. Molitor D. 2018. The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration(). *American economic journal. Economic policy* 10:326-56
57. Office of the Assistant Secretary of Defense. 2016. Memorandum for Under Secretary of Defense: Fiscal Year 2017 Direct Care inpatient Billing Rates.
58. Osborne NH, Nicholas LH, Ryan AM, Thumma JR, Dimick JB. 2015. Association of hospital participation in a quality reporting program with surgical outcomes and expenditures for Medicare beneficiaries. *Jama* 313:496-504
59. Osterman M, Martin J. 2014. Trends in Low-risk Cesarean Delivery in the United States, 1990-2013. *National Vital Statistics Report* 63
60. Papanicolaos I, Woskie LR, Jha AK. 2018. Health care spending in the united states and other high-income countries. *JAMA* 319:1024-39
61. Pradarelli JC, Healy MA, Osborne NH, Ghaferi AA, Dimick JB, Nathan H. 2016. Variation in Medicare Expenditures for Treating Perioperative Complications: The Cost of Rescue. *JAMA Surg* 151:e163340
62. Quan H, Li B, Couris CM, Fushimi K, Graham P, et al. 2011. Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries. *American Journal of Epidemiology* 173:676-82
63. Regenbogen SE, Gust C, Birkmeyer JD. 2012. Hospital surgical volume and cost of inpatient surgery in the elderly. *J Am Coll Surg* 215:758-65
64. Santry HP, Gillen DL, Lauderdale DS. 2005. Trends in bariatric surgical procedures. *Jama* 294:1909-17
65. Scally CP, Thumma JR, Birkmeyer JD, Dimick JB. 2015. Impact of Surgical Quality Improvement on Payments in Medicare Patients. *Annals of surgery* 262:249-52
66. Skinner J. 2012. Causes and Consequences of Regional Variations in Health Care. In *Handbook of Health Economics*, 2:45-86. Number of 45-86 pp.

67. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. 2010. Regional Variations in Diagnostic Practices. *New England Journal of Medicine* 363:45-53
68. The Joint Commission. 2018. Specifications Manual for Joint Commission National Quality Measures (v2018A). Version 2018A
69. Van Parys J. 2016. Variation in Physician Practice Styles within and across Emergency Departments. *PloS one* 11:e0159882
70. Wakeam E, Molina G, Shah N, Lipsitz SR, Chang DC, et al. 2017. Variation in the cost of 5 common operations in the United States. *Surgery* 162:592-604
71. Wennberg JE. 2002. Unwarranted variations in healthcare delivery: implications for academic medical centres. *BMJ : British Medical Journal* 325:961-4
72. Wennberg JE, Fisher ES, Skinner JS. 2002. Geography and the debate over Medicare reform. *Health Aff (Millwood)* Suppl Web Exclusives:W96-114
73. Zuckerman S, Waidmann T, Berenson R, Hadley J. 2010. Clarifying sources of geographic differences in Medicare spending. *The New England journal of medicine* 363:54-62

Supplemental Appendix

SUPPLEMENTAL APPENDIX – LARGE VARIATIONS FOUND AMONG MILITARY HMO ENROLLEES: A LOOK USING COMPULSORY PATIENT MIGRATION

Technical Description of Models

We want to decompose variation in log utilization across regions into a supply-side component captured by a place effect and a demand-side component attributable to patients. We do so through an additive decomposition of ln-utilization and by conducting an event study of movers. We use the following model to model an individual's healthcare expenditures. For a mover i moving in year t_i^* , we define the year relative to move to be $r(i,t)=t-t_i^*$ and the natural log of total healthcare spending, y_{ijt} , as follows:

$$y_{ijt} = \alpha_i + \gamma_j + \tau_t + \rho_{r(i,t)} + x_{i,t}\beta_1 + \epsilon_{i,j,t} \quad (1)$$

In this equation, α_i , γ_j , and τ_t are all fixed effects for patient, geographic area, and year respectively. $\rho_{r(i,t)}$ is a relative-year fixed effect term for movers. If an individual is a non-mover, then $\rho_{r(i,t)}$ is equal to zero. $x_{i,t}$ is the effect of measured time-varying patient characteristics. $\epsilon_{i,j,t}$ is the statistical residual with the assumptions that $E[\epsilon_{i,j,t} | \alpha_i, \gamma_j, \tau_t, \rho_{r(i,t)}, x_{i,t}] = 0$. We then denote the combined effect of patient characteristics through the following equation: $c_{it} = \alpha_i + \rho_{r(i,t)} + x_{i,t}\beta$. For our main model specification, $x_{i,t}$ consists of age bins. We also test alternate model specifications³.

Patient migration allow us to separate differences in patient fixed effects α_i , from location fixed effects γ_j . Using a simplified form of the model in which time-varying variables are suppressed to 0, and utilization depends only on patient and location fixed

³ In my original proposal, I had hoped to also include physician characteristics into the model through a physician fixed effect and time-variant characteristics such as clinical experience and additional training. However, due to data restraints, we were not able to include these variables into our model.

effects plus an error term, we can show that the difference between the average utilization in the years after the move and the years before the move is a constant estimator of location specific effects. When we observe samples of patients moving, we can calculate consistent estimates of the expected location fixed effects for each location in addition to average log utilization and can therefore calculate the average contribution of patient characteristics by differencing.

Identification in the full model is similar as time fixed effects and age coefficients do not rely on movers. Our relative year fixed effect allows for arbitrary changes in log utilization for movers pre- and post-move assuming that these changes are the same regardless of origin and destination. Identification of these effects occurs if we have movers with multiple origins and destinations. These fixed effects terms allow movers to differ from non-movers with respect to baseline utilization and trends in utilization over time.

Mathematically we can then describe the share of differences attributable to place and patients between areas j and j' to be:

$$S_{place}(j, j') = \frac{\gamma_j - \gamma_{j'}}{\bar{\gamma}_j - \bar{\gamma}_{j'}} \quad S_{patient}(j, j') = \frac{\bar{c}_j - \bar{c}_{j'}}{\bar{\gamma}_j - \bar{\gamma}_{j'}}$$

We then compare overall differences in utilization as well as differences in the average patient- and place- components of sets of catchments sorted by their average expenditures. We decompose expenditures in the tables for catchments with spending that falls above and below median, in the top and bottom quintile, and finally, in the top and bottom decile.

For our event study, following Finkelstein et al (33) and Bronnenberg et al (14), we scale y to provide a more informative estimate of the direction and magnitude of

changes resulting from a move. Scaling our outcome allows our model to be able to plot changes for movers across a spectrum of destinations and origins irrespective of the sign or magnitude of the change. After scaling, a value of 0 indicates that utilization is equal to the average utilization from the catchment of origin while a value of 1 indicates utilization equal to the average at this destination. We denote δ_i to represent the difference in average log expenditures between a beneficiary's destination $d(i)$ and origin $o(i)$ over the course of a move. We parameterize the interaction of δ_i with the relative year terms for movers. This results in the following equation:

$$y_{it} = \tilde{\alpha}_i + \theta_{r(i,t)} \hat{\delta}_i + \tau_t + x_{i,t} \beta_1 + \epsilon_{i,t} \quad (2)$$

In this equation, $\tilde{\alpha}_i$ is a patient fixed effect term which combines the individual patient fixed effect, α_i , with the catchment fixed effect at the mover's origin, $\gamma_{o(i)}$. τ_t and $x_{i,t}$ remain time fixed effects and age-bins respectively. The relative-year specific coefficients, $\theta_{r(i,t)}$, become the parameter of interest. They measure changes in y_{it} around years of the move scaled relative to δ_i .

Supplemental Tables

Table 11, Age and Sex-adjusted Spending by Census Region and Division (2007-2014)

Age and Sex Adjusted Annual Expenditures by Census Region		
Region	Mean Annual Expenditures	SD
Northeast	3,9821.48	560.24
Midwest	3,983.44	448.93
South	3,882.58	483.37
West	3,870.02	488.09

Age and Sex Adjusted Annual Expenditures by Census Division		
Division	Mean Annual Expenditures	SD
New England	3,969.96	476.91
Mid-Atlantic	3,718.99	619.99
East North Central	3,892.33	489.64
West North Central	4,036.59	425.10
South Atlantic	3,825.24	484.69
East South Central	3,950.57	507.41
West South Central	4,000.87	459.70
Mountain	4,016.46	385.62
Pacific	3,804.20	519.40

Table 12. Additional Characteristics of Movers

Additional Characteristics of Movers		
Origin Region	Counts	Percentage
Northeast	5,849	5.09
Midwest	8,970	7.81
South	67,020	58.33
West	33,057	28.77
Total	114,896	100
Destination Region		
Northeast	5,057	4.4
Midwest	9,391	8.17
South	69,932	60.87
West	30,516	26.56
Total	114,896	100
Year of Move		
2007	0	0
2008	17,550	15.27
2009	17,924	15.6
2010	16,120	14.03
2011	17,444	15.18
2012	14,703	12.8
2013	13,027	11.34
2014	18,128	15.78
Total	114,896	100
Move Distance		
Average (mi)	2,190.86	
SD	2069.11	

Table 13. Additive Decomposition of Expenditures and Utilization – Active Duty Sample
Note: Table is based upon estimation of our main model specification using either ln-transformed annual per-capita expenditures (Expenditures) or ln-transformed quantity-based expenditures (Utilization) as outcomes. The models both include patient, catchment, calendar year fixed effects, and relative year fixed effects for movers which are normalized to zero for non-movers. We also include time-variant MDR-generated age bins. The “expenditure” model had an adjusted R-squared value of .4671 and the “utilization” model’s adjusted R-squared was .4627. The sample includes all movers and non-movers. n=548,188 patient years

		Above & Below Median Catchments	Top & Bottom 25% Catchments	Top & Bottom 10% Catchments
Difference in Average Ln Expenditures				
	Overall	.5986	1.0974	1.4228
	Due to Place	.5799	1.0497	1.3872
	Due to Patients	.0188	.0478	.0356
Share of Difference				
	Patients	.0313	.0435	.0251
	Place	.9687	.9565	.9749
		Above & Below Median Catchments	Top & Bottom 25% Catchments	Top & Bottom 10% Catchments
Difference in Average Ln Utilization				
	Overall	.3736	.6948	.9164
	Due to Place	.3372	.6183	.8412
	Due to Patients	.0364	.0765	.0752
Share of Difference				
	Patients	.0975	.1101	.0821
	Place	.9025	.8899	.9179

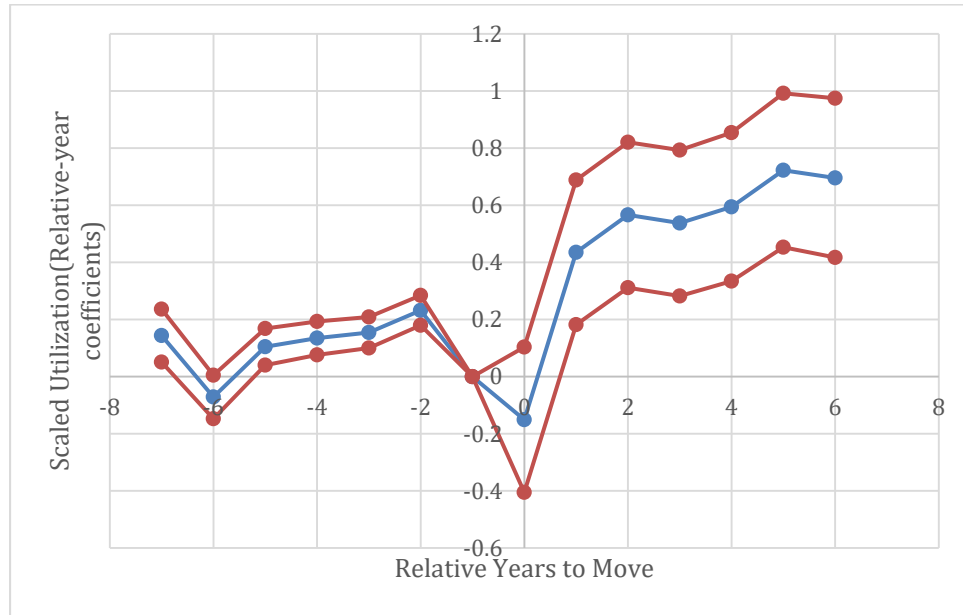
Table 14. Additive Decomposition of Expenditures and Utilization – Added Health Adjustment

Note: Table is based upon estimation of an alternate model specification using either ln-transformed annual per-capita expenditures (Expenditures) or ln-transformed quantity-based expenditures (Utilization) as outcomes. The models both include patient, catchment, calendar year fixed effects, and relative year fixed effects for movers which are normalized to zero for non-movers. We also include time-variant MDR-generated age bins, Charlson Comorbidity Score, and a MDR injury-flag. The MDR injury-flag is pulled from patient records if clinical staff record an acute injury during an encounter. These variables control for health status changes over time as well as acute shocks. The “expenditure” model had an adjusted R-squared value of .4013 and the “utilization” model’s adjusted R-squared was .3988. The sample includes all movers and non-movers. n=3,160,616 patient years.

		Above & Below Median Catchments	Top & Bottom 25% Catchments	Top & Bottom 10% Catchments
Difference in Average Ln Expenditures				
Share of Difference	Overall	.2652	.5097	.7182
	Due to Place	.2481	.4819	.6688
	Due to Patients	.0171	.0277	.0494
	Patients	.0644	.0544	.0687
	Place	.9356	.9456	.9313
		Above & Below Median Catchments	Top & Bottom 25% Catchments	Top & Bottom 10% Catchments
Difference in Average Ln Utilization				
Share of Difference	Overall	.1810	.3311	.4368
	Due to Place	.1320	.2588	.3269
	Due to Patients	.0490	.0724	.1098
	Patients	.2705	.2185	.2515
	Place	.7295	.7815	.7485

Figure 7. Event Study of Utilization

Note: The event study plots coefficients from our parameterized relative- year coefficients. The dependent variable is our ln-transformed quantity-based measure of utilization. The models include patient, catchment, calendar year fixed effects, and relative year fixed effects interacted with $\hat{\delta}_i$, the difference in average utilization between a beneficiary's destination and origin. We also include time-variant MDR-generated age bins. The coefficient for relative year -1 is normalized to zero. The orange lines represent 95% confidence intervals calculated from robust standard errors. The sample includes all movers. N= 919,168 patient years



Robustness Checks

Table 15. Traditional Split Sample

Note: Table 12 repeats our main model decompositions but with a traditional split-sample design. That is patients were randomly assigned to one of two sub-samples and the analysis was repeated. The sample is all movers and non-movers.

		Above & Below Median	Top & Bottom 20%	Top & Bottom 10%
<u>Sample 1</u>				
Difference in Average ln Expenditures				
	Overall	.3134	.5907	.7581
	Due to Place	.2806	.5257	.1003
	Due to Patients	.0329	.0650	.1003
Share of Difference				
	Patients	.1049	.1101	.1323
	Place	.8951	.8899	.8677
<u>Sample 2</u>				
Difference in Average ln Expenditures				
	Overall	.3093	.5751	.7918
	Due to Place	.2713	.4956	.6728
	Due to Patients	.0380	.0795	.1190
Share of Difference				
	Patients	.1229	.1382	.1503
	Place	.8771	.8618	.8497

Table 16. Split Sample by Year

Note: Here we repeat our main model decompositions but groups of catchments are sorted and grouped into median, quintiles, and deciles according to expenditure data from 2007-2009. We then use the groupings from the “early” period to run our decompositions on data for years 2010-2014. The sample is all movers and non-movers.

		Above & Below Median	Top & Bottom 20%	Top & Bottom 10%
Difference in Average ln Expenditures				
	Overall	.2906	.5563	.7395
	Due to Place	.2666	.5324	.6902
	Due to Patients	.0240	.0239	.0493
Share of Difference				
	Patients	.0826	.0430	.0666
	Place	.9174	.9570	.9334

Table 17. Split Sample by Year – Active Duty Sample

Note: Here we repeat our main model decompositions but groups of catchments are sorted and grouped into median, quintiles, and deciles according to expenditure data from 2007-2009. We then use the groupings from the “early” period to run our decompositions on data for years 2010-2014. The sample is all Active Duty movers and non-movers.

		Above & Below Median	Top & Bottom 20%	Top & Bottom 10%
Difference in Average ln Expenditures				
	Overall	.6053	1.1027	1.4255
	Due to Place	.5801	1.0570	1.3925
	Due to Patients	.0252	.0457	.0331
Share of Difference				
	Patients	.0417	.0414	.0232
	Place	.9583	.9586	.9768

Table 18. Alternate Outcomes – Level Spending

Note: Here we repeat our main model decompositions but instead of using log-transformed outcomes, we use spending in levels. The sample is all movers and non-movers.

		Above & Below Median	Top & Bottom 20%	Top & Bottom 10%
Difference in Average ln Expenditures				
	Overall	1,104.39	1,934.50	2404.09
	Due to Place	889.40	1571.32	1942.11
	Due to Patients	214.99	363.18	461.99
Share of Difference				
	Patients	.1947	.1877	.1922
	Place	.8053	.8123	.8078

Table 19. Alternate Outcomes- Log spending plus .01

Note: Here we repeat our main model decompositions but we instead use spending plus .01 and then natural log transform. The sample is all movers and non-movers.

		Above & Below Median	Top & Bottom 20%	Top & Bottom 10%
Difference in Average ln Expenditures				
	Overall	.3776	.7118	.9714
	Due to Place	.3522	.6707	.8962
	Due to Patients	.0254	.0411	.0752
Share of Difference				
	Patients	.0672	.0578	.0774
	Place	.9328	.9422	.9226

Table 20. Alternate Outcomes- Log spending plus 10

Note: Here we repeat our main model decompositions but we instead use spending plus 10 and then natural log transform. The sample is all movers and non-movers.

		Above & Below Median	Top & Bottom 20%	Top & Bottom 10%
Difference in Average ln Expenditures				
	Overall	.2912	.5193	.6628
	Due to Place	.2714	.4716	.6223
	Due to Patients	.0199	.0477	.0405
Share of Difference				
	Patients	.0683	.0919	.0611
	Place	.9317	.9081	.9389

Figure 8. Comparison of Adjusted Spending Across Hospital Referral Regions

Note: Figure 6 plots adjusted MHS spending adjusted Medicare spending stratified by census region. MHS expenditures are age and sex adjusted as race is commonly missing from the MDR. Medicare expenditures are age, sex, and race adjusted. TRICARE network claims in certain non-specific geographic catchments lacking zipcodes could not be matched to HRRs for this portion of the analysis due to the lack of precision with respect to location (i.e. Northern Idaho, Western Kentucky). Therefore we suppress approximately 27% of the sample when comparing to Hospital Referral Regions. We limit this portion of our analysis to the 112 HRRs with corresponding data from the MHS.

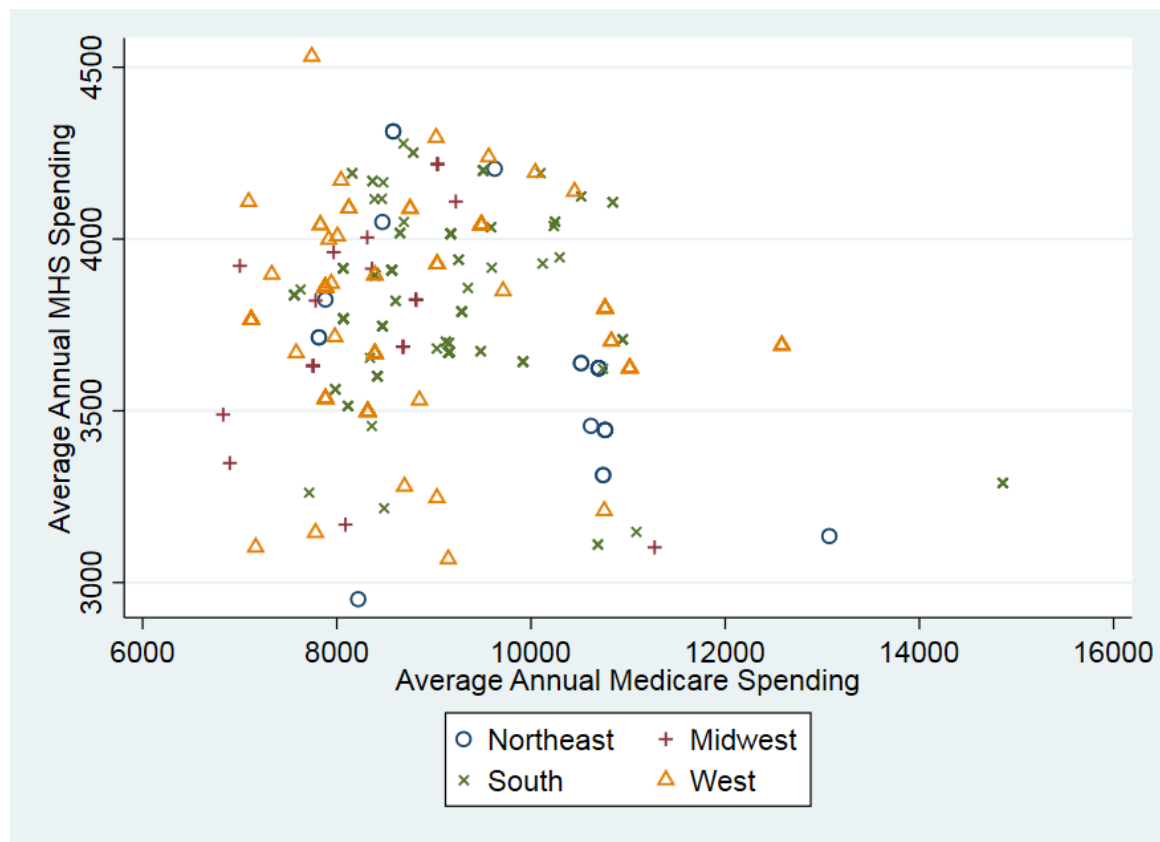


Figure 9. Comparison of Adjusted Spending Ranks across Hospital Referral Regions
 Note: Figure 7 plots MHS spending ranks against Medicare spending ranks stratified by census region. MHS expenditures are age and sex adjusted as race is commonly missing from the MDR. Medicare expenditures are age, sex, and race adjusted. TRICARE network claims in non-specific geographic catchments lacking zip codes could not be matched to HRRs for this portion of the analysis due to the lack of precision with respect to location (i.e. Northern Idaho, Western Kentucky). Therefore we suppress approximately 27% of the sample when comparing to Hospital Referral Regions. We limit this portion of our analysis to the 112 HRRs with corresponding data from the MHS. Therefore, our spending ranks may not correspond to ranks using all HRRs.

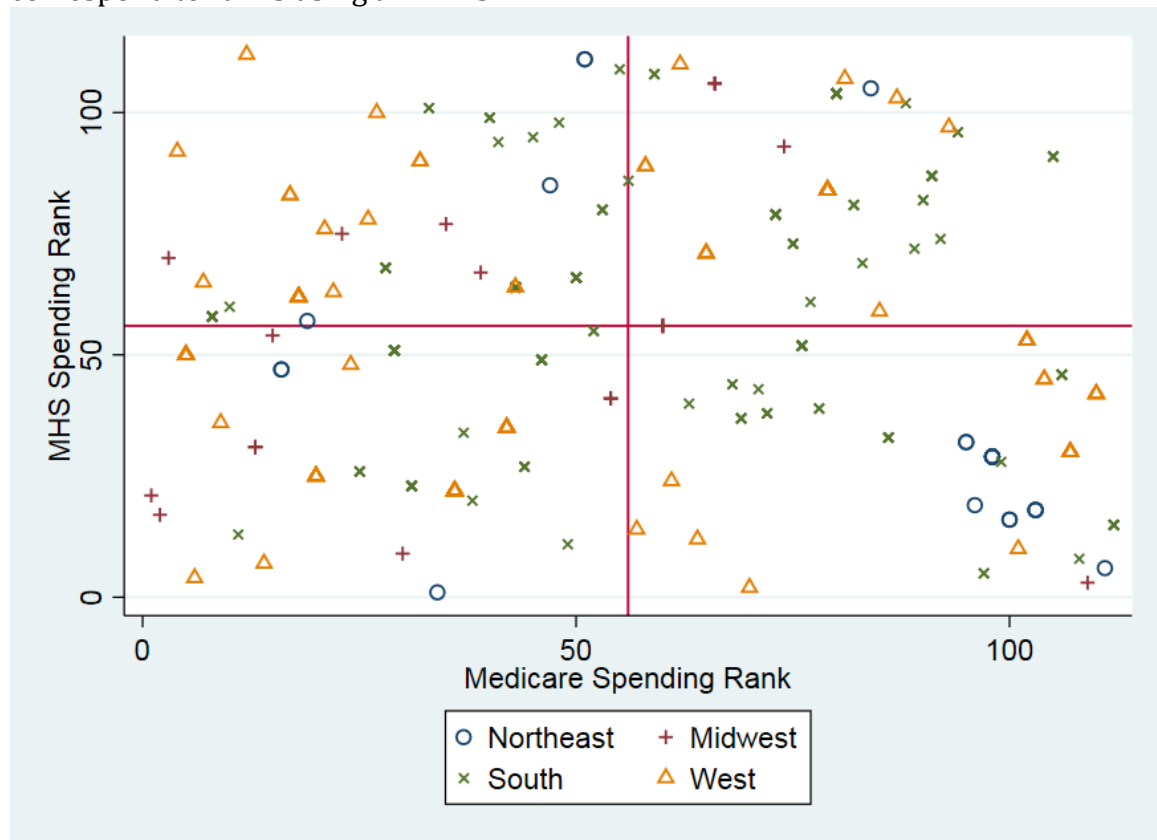


Table 21. Additive Decompositions Using Hospital Referral Regions

Note: Here we repeat our main model decompositions but use Hospital Referral Regions instead of TRICARE Catchments. TRICARE network claims in non-specific geographic catchments lacking zipcodes could not be matched to HRRs for this portion of the analysis due to the lack of precision with respect to location (i.e. Northern Idaho, Western Kentucky). Therefore we suppress approximately 27% of the sample when comparing to Hospital Referral Regions. We limit this portion of our analysis to the 112 HRRs with corresponding data from the MHS.

		Above & Below Median	Top & Bottom 20%	Top & Bottom 10%
Difference in Average In Expenditures				
	Overall	.2865	.5586	.7652
	Due to Place	.2643	.5194	.7464
	Due to Patients	.0222	.0391	.0187
Share of Difference				
	Patients	.0774	.0701	.0245
	Place	.9226	.9299	.9755
Difference in Average In Utilization				
	Overall	.1701	.3194	.4125
	Due to Place	.1441	.2662	.3560
	Due to Patients	.0260	.0532	.0565
Share of Difference				
	Patients	.1530	.1666	.1371
	Place	.8470	.8334	.8629

SUPPLEMENTAL APPENDIX – VARIATION IN COMPLICATION RATES AND BARIATRIC SURGERY EPISODE COSTS IN THE MILITARY HEALTH SYSTEM

International Classification of Disease Codes

Table 23. ICD-9-CM Complication Codes

Note: Adapted from Banka and Colleagues (2012)

Unexpected re-operations for surgical complications

Wound dehiscence Procedure: 054.61

Lysis of adhesions Procedure: 054.51, 054.59

Removal of foreign body Procedure: 054.92

Laparotomy Procedure: 054.12, 054.19

Anastomotic

Leak Diagnosis: 99.86

Percutaneous abdominal drainage Procedure: 054.91

Splenic

Injury Diagnosis: 86.5

Partial or complete splenectomy Procedure: 041.43, 041.5

Splenotomy Procedure: 041.2

Hemorrhagic

Intraoperative hemorrhage Diagnosis: 998.11

Postoperative hematoma Diagnosis: 998.12

Wound

Infection Diagnosis: 998.5, 998.51, 998.59

Seroma Diagnosis: 998.13

Dehiscence Diagnosis: 998.3

Intestine

Small bowel obstruction Diagnosis: 560

Injury to gastrointestinal tract Diagnosis: 863

Liver

Injury Diagnosis: 864

Other Post-Operative Complications

Other and Unspecified effects of external causes Diagnosis: 994, 995

Complications of Surgical and Medical Care, not elsewhere classified Diagnosis: 996, 997, 998, 999

Blood Transfusion 990

Pulmonary

Respiratory tract complications Diagnosis: 997.5

Pneumonia Diagnosis: 3.22, 20.3, 20.4, 20.5, 21.2, 22.1, 31.0, 39.1, 52.1, 55.1, 73.0, 83.0, 112.4, 114.0, 114.4, 114.5, 115.05, 115.15, 115.95, 130.4, 136.3, 480.0, 480.1, 480.2, 480.8, 480.9, 481, 482, 483, 484, 485, 486, 513.0, 517.1

Acute respiratory failure Diagnosis: 518.81

Other Pulmonary Complications

Other diseases of the Lung Diagnosis: 514, 518.4, 518.5, 518.6, 518.81, 518.82, 518.84

Tracheotomy Procedure: 31.1, 31.29

Cardiac

Cardiac complications, not elsewhere specified Diagnosis: 997.1

Cardiac Arrhythmias Diagnosis: 426, 427

Myocardial Infarction Diagnosis: 410

Neurologic

Central nervous system complications Diagnosis: 997.01, 997.02, 997.03

Acute cerebrovascular accident Diagnosis: 430, 431, 432.0, 432.1, 432.9, 433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434

Genitourinary tract

Urinary tract complications Diagnosis: 599.0, 997.5

Acute renal failure Diagnosis: 584

Acute dialysis Procedure: 38.95

Insertion of short-term dialysis catheter Procedure: 39.95

Thromboembolic

Acute Pulmonary Embolism Diagnosis: 415.1, 415.11, 415.19

Deep venous thrombosis Diagnosis: V12.51, V125.2, 451, 453.1, 453.2, 453.3, 453.8, 453.9

Postoperative

Shock Diagnosis: 998.0

Stroke Diagnosis: 430, 431, 432, 433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434, 436

Gastrointestinal Leak Diagnosis: 997.4

Abscess

Liver abscess and sequelae of chronic liver disease, other suppurative peritonitis,

Other postoperative infection Diagnosis: 572, 567.2, 998.59

Sepsis Diagnosis: .038-.038.9, 998.51, 998.59

Table 24. ICD-9-CM Codes for Bariatric Procedures

Procedure	Code
Vertical Banded Gastroplasty	44.68
Adjustable Gastric Band	44.69
Biliopancreatic Diversion with Duodenal Switch	45.91 43.89 45.51
Roux-en-Y Gastric Bypass	44.39
Biliopancreatic Diversion	43.7
Laparoscopic Roux-en-Y	44.38
Laparoscopic Sleeve Gastrectomy	43.82
Laparoscopic Adjustable Gastric Band	44.95

**SUPPLEMENTAL APPENDIX – PHYSICIAN COMPULSORY RELOCATION IN THE U.S.
MILITARY HEALTH SYSTEM: THE CASE OF C-SECTIONS**

International Classification of Disease Codes

Table 25. ICD-9-CM Codes for Deliveries

Procedure	Code
Delivery Single Live-Born	V27.0
Cesarean Section	74.*

Table 26. ICD-9-CM Codes for Complications

Diagnosis	Code
Previous C-section	654.2
Malpositioned Fetus	652*
Antepartum Hemorrhage	641*
Hypertension Complication	642*
Pre-term Labor	644*
Prolonged/Late Labor	645*
Excessive Edema	646.1
Maternal Infection	646.4 646.6
Diabetes	648.0
Cardiovascular Disorder	648.5 648.6
Multiple Gestation	651*
Disproportion of Fetus	653*
Soft Tissue Abnormality	654.0 654.7 654.8 654.9
Fetal Abnormality	678* 679.1
Polyhydramnios	65.7
Oligohydramnios	65.80
Membrane Rupture	665.0 665.1 658.1 658.2
Amniotic Infection	658.4
Advanced Maternal Age (>35)	659.5 659.6
Asthma	493*
Herpes	05.41